

Words Matter:**Being // Non-being in Language(s) and Text(s)**

10.2478/abcsj-2023-0020

**Grammatical Number from an Ecological Perspective,
Focused on the “Here-Now-I-Real”****TOSHIYUKI SADANOBU**
Kyoto University, Japan**Abstract**

This article shows, using the example of number agreement, that an ecological perspective with a focus on the situation of the utterance (i.e. here-now-I-real) is an effective way to understand differences in grammar among individual languages and to identify commonalities above and beyond such differences. On the surface, English and Chinese would appear to be exact opposites in terms of whether or not they conform to the rule of number agreement, while the position of Japanese would appear to be ambivalent on this question. But in fact this is not the case. Numerous languages around the world are consistent in the way number distinction is more likely to arise the higher the subject's position on the animacy hierarchy. Chinese and Japanese are no exceptions. Leaving aside differences in individual circumstances, languages differ only in the boundary lines on the animacy hierarchy where number distinction becomes obligatory. In order to reach the above understanding, it is essential to see animacy from an ecological perspective. Furthermore, an ecological perspective is essential as a principle for explaining levels of not only animacy but also self-expressiveness, esteem, and transitivity. The prevailing view of language as separate from the situation of the utterance requires considerable revision on the subject of grammar.¹

Keywords: (dis)placement, situation of utterance, ecological perspective, number agreement, number distinction, attenuation of self-expressiveness,

discoloration of animacy/ esteem/ transitivity, animacy hierarchy, modifying-modified structure, appositive structure

Grammar may vary by language to a surprising degree. It is not unusual for two languages to be complete opposites in terms of grammar. One example of this phenomenon is number agreement. In English, number agreement would appear to be a correct rule that must be followed, while in Chinese, in contrast, it would appear to be an incorrect rule that, if followed, would make words, phrases, and sentences sound unnatural.

Number agreement means here grammatical concord of numbers between various kinds of lexemes. Combinations of quantifiers with nouns in English can be given as one well-known example. See Examples (1) and (2) below.

- (1) a. one child
b. * one children
- (2) a. * two child
b. two children

In Example (1), where the quantifier is singular, case (a), in which the noun phrase too is singular, sounds natural while case (b), with a plural noun phrase, does not. In Example (2), where the quantifier is plural, case (a), in which the noun phrase is singular, sounds unnatural while case (b), with a plural noun phrase, sounds natural.ⁱⁱ

But the situation is completely different in Chinese. In Example (3) below, the Standard Chinese equivalent of Example (2) in English, in contrast the singular noun phrase in case (a) sounds natural while the plural in (b) sounds unnatural.ⁱⁱⁱ

- (3) a. *liǎng -ge háizi*
two -CL child
(Lit.) “two child”
b. * *liǎng -ge háizi -men*
two -CL child -PL
(Lit.) “two children”

The first reason likely to occur to most readers to support the unnaturalness of Chinese Example (3b) might be the need to avoid redundancy. This is the argument that since adding “children” after already indicating the number “two” would be redundant, the statement should be simplified to eliminate a useless component. Applying this logic to English would lead to the view that English is a peculiar language that requires useless expressions.

Of course, English should be considered in light of its own logic. This is the logic of consistency, which demands that a speaker's message be internally consistent. Under this logic, when referring to multiple children the plural must be maintained not only in the modifier “two” but also in the modified “children.” Applying this logic to Chinese would imply that Chinese is a peculiar language that lacks consistency.

The situation appears to be even more complicated when we add Japanese to the mix. Consider Example (4) below.

- (4) a. *futa -ri =no kodomo*
two -CL =GEN child
(Lit.) “two child”
b. *futa -ri =no kodomo -tati*
two -CL =GEN child -PL
(Lit.) “two children”

In Japanese, “two” may modify either “child” as in Example (4a) or “children” as in Example (4b). As a result, number agreement would appear in Japanese to be a rule that may be either followed or ignored. From the perspective of Japanese, both English and Chinese would appear to be peculiarly inflexible languages tied to the logic of either consistency or avoidance of redundancy. In contrast, viewed from the English or Chinese perspectives Japanese would appear to be a peculiarly ad-hoc language.

When, in these ways, the grammatical rules of individual languages appear to differ substantially from one another, viewing one language in terms of the logic of another would make the former language appear somewhat irrational. The logic of one language is merely a logic that applies within that language alone. It would be incorrect to consider that it should apply to all languages in general, and view other languages in its terms.

English logic applies to English only, Chinese logic to Chinese only, and Japanese logic to Japanese only.

But surely this does not mean that there is no universal logic above and beyond the logic of each individual language. But what kind of universal logic could give rise to the individual logics of English, Chinese, and Japanese in their respective communities?

Functional Overkill and Its Moderation

Mark Durie proposed the concept of lessening functional overkill (1995) in order to understand such grammatical differences among languages and look for commonalities above and beyond these differences.

As used here, “overkill” refers metaphorically not simply to killing three birds when two would have done but to employing excessive destructive force, such as using a missile to kill a single ant. What is to be killed in this case is imprecision in linguistic expression. If we consider language to be used frequently to convey information in communication and thus believe that linguistic expression has a function of conveying information, then linguistic expression must be precise at all costs. To enable proper communication of a message or concept even if the listener is unable to hear some parts of the utterance, the speaker needs to make every effort at overkill to eliminate any imprecision in linguistic expression. This is the concept of functional overkill.

For example, generally it will be clear from the context whether a speaker is referring to one child or multiple children. The fact that, even so, languages such as English use redundant methods of expression such as indicating plurality by “two” at first and then indicating plurality again by using “children” instead of “child” does not seem unusual when considered from this perspective of functional overkill.

But since it applies an extra burden in the form of excessive destructive power, this functional overkill cannot be used unrestrictedly. Durie uses the allegory of fish living in a region stricken by drought. While some fish may acquire the ability to dig deep into the mud and others to lay eggs resistant to drying, no species will combine both the abilities to dig deep into the mud and to lay eggs resistant to drying. Although a population

will do anything it can to avoid the threat of extinction, the degree of such effort is balanced against the resulting burden. This is the concept of lessening functional overkill.

The degree at which functional overkill is necessary – that is, the degree to which it will be lessened – is determined by each individual community. In light of this way of thinking, it would appear reasonable that in expressing number languages such as Chinese differ from English by focusing mainly on pragmatic means through which it is understood from the context whether an expression concerns one or more children, and thus their functional overkill in this case is relatively subdued. The fact that in some situations Japanese resembles Chinese in this regard also would appear reasonable under this thinking.

In this way, based on the concept of lessening functional overkill, English, Chinese, and Japanese grammar would each appear to be the result of the struggle between the two inverse motives of ensuring consistency through overkill of imprecision in linguistic expression and avoiding redundancy in order to lessen overkill. Thus, differences in grammar among individual languages can be understood as differences in the motive prioritized as a result of this struggle. From the perspective of overkill, it would appear that English grammar prioritizes the motive of ensuring consistency, while Chinese grammar gives priority to avoiding redundancy to lessen overkill and in Japanese grammar the motives of ensuring consistency and avoiding redundancy are rivals, neither of which can be said clearly to have priority.

Certainly, understanding and ascertaining grammar in this way through the concepts of overkill and lessening its burden would appear reasonable because it is consistent with our generally accepted ideas that language is used frequently to convey information in communication and that in such a case linguistic expression has the function of conveying information. But actually, understanding and ascertaining grammar in this way is not enough. As one example, number agreement may occur frequently, although not always, in Chinese when used in appositive expressions containing personal pronouns, as in Example (5) below. (This will be discussed in detail in the section *Animacy and syntactic structure*.)

- (5) a. *wǒ -men liǎng -ge*
 1sg -PL two -CL
 “we two”
 b. *nǐ -men liǎng -ge*
 2sg -PL two -CL
 “you two”
 c. *tā -men liǎng -ge*
 3sg -PL two -CL
 “they two”

In Example (5a), the plural suffix *-men* is attached to the first-person pronoun *wǒ* in agreement with the plurality of the number of persons referred to, two. Similarly, in Examples (5b) and (5c), *-men* is attached to the second-person pronoun *nǐ* and the third-person pronoun *tā*, respectively. In this way, in Chinese, in which one would expect number agreement to be avoided, it actually does occur in the limited case of appositive expressions with personal pronouns. This fact cannot be explained by acceptance of the concepts of overkill and lessening of its burden alone.

So how should we understand such diversity among languages? The following sections will present an analysis based on the ecological view of language, which considers language without separating it from the situation of the utterance.

The Ecological View of Language and Image Attenuation/Discoloration

The ecological view of language, a holistic view that emphasizes the close connection between language and situation of the utterance (“here-now-I-real”), is advocated by Talmy Givón and Paul J. Hopper for example as in (6) and (7) below.

- (6) I will describe a number of recurring themes in diachronic syntax. I would like to suggest that all of them represent processes by which loose, paratactic, “pragmatic” discourse structures develop-over time-into tight, “grammaticalized” syntactic structures. (Givón 208)
- (7) Because grammar is always emergent but never present, it could be said that it never exists as such, but is always coming into being. There is, in other words, no “grammar” but only “grammaticization” – movements

toward structure which are often characterizable in typical ways.
(Hopper 148)

This view is the inverse of the prevailing view. The prevailing view considers language as separate from the situation of an utterance. For example, see the passage from Charles F. Hockett under (8) below.

- (8) Man is apparently almost unique in being able to talk about things that are remote in space or time (or both) from where the talking goes on. This feature – “displacement” – seems to be definitely lacking in the vocal signaling of man’s closest relatives, though it does occur in bee-dancing. (Hockett 90)

It is true that a displaced view of language would appear to be appropriate for the bulk of the lexicon. For example, the word “apple” refers to a large deciduous tree of the family Rosaceae and its fruit no matter when, where, and by whom it is spoken. Perhaps most of the lexicon has this aspect of displacement. Only a few words are deictic, or change in meaning with the situation in which they are uttered.

However, this displaced view can explain only one side of language. This is because if we turn from vocabulary to grammar, or regularity among words, we can observe in language a strong situational quality. This section will demonstrate this fact by observing expressions of person, animacy, transitivity, and esteem in Japanese.

In the world of Japanese-language expression, as the thing under discussion is removed from the situation of an utterance, its self-expressive quality attenuates and its animacy and esteem, as well as the transitivity of an event, also undergo fading and discoloration. This also is apparent in the structure of the words and phrases used to express the thing (that is, the grammar). These phenomena will be observed below, in order.

In discussing the naturalness of Japanese-language utterances below, this article will refer where appropriate to the results of a survey of about 100 native Japanese speakers.

Attenuation of Self-Expressiveness

First, we will consider some expressions of the current internal aspects (thoughts and feelings) of the speaker and others. See Examples (9) and (10).

- (9) a. *watasi =wa kokyō =o koisiku omoi -masu.*
 1sg =TOP hometown=ACC miss -POL
 “I miss my hometown.”
 [Survey 1, mean: 4.61, variation: 0.38, SD: 0.62, median: 5]^{iv,v}
- b.?? *kare =wa kokyō =o koisiku omoi -masu.*
 3sg.M =TOP hometown=ACC miss -POL
 “He misses his hometown.”
 [Survey 1, mean: 2.51, variation: 1.68, SD: 1.30, median: 2]
- (10) a. *watasi =wa kaeru =wa yōseirui =da =to omoi -masu.*
 1sg =TOP frog =TOP amphibian=COP=QUOT think -POL
 “I think a frog is an amphibian.”
 [Survey 2, mean 4.50, variation: 0.70, SD: 0.83, median: 5]^{vi}
- b.?? *kare =wa kaeru =wa ryōseirui =da =to omoi -masu.*
 3sg.M=TOP frog=TOP amphibian=COP=QUOT think -POL
 “He thinks a frog is an amphibian.”
 [Survey 2, mean: 1.55, variation: 0.85, SD: 0.93, median: 1]

Here, the figures after each example (9a, b) (10a, b) indicate the mean, variation, standard deviation (SD), and median values in evaluation of how natural the expressions sounded based on survey results. In the questions on the degree of naturalness of expressions in this survey, respondents were asked to grade expressions in terms of how natural they sounded, on a scale of 1 to 5 (1 pt.: very unnatural – 5 pts.: very natural). For example, the scores “Survey 1, mean 4.61, median 5” shown at the end of (9a) indicate that a majority of respondents in Survey 1 chose the answer “very natural” for expression (9a). Single (“?”) or double question marks (“??”) are appended before examples considered by majorities of respondents to sound unnatural, as in Ex. (9b), to indicate unnaturalness of expressions (and the degree thereof). The same indications will be used for (10) and later examples.

The survey results show that the speaker’s current internal thoughts or feelings can be expressed using the form *=to omoi =masu* (“to think –”) (Exs. (9a) (10a)). Dividing the self into the <feeling self> and <thinking self> for ease of explanation, expressions in the form *=to omoi =masu* sound highly natural in the cases of both the <feeling self> (Ex. (9a)) and the <thinking self> (Ex. (9a)). The form *=to omoi=masu* is a form of expression for use in referring to oneself only. This form sounds unnatural

when used to express the internal thoughts or feelings of another person (Exs. (9b) (10b)).

However, in expressions of the past, the <thinking self> may be handled largely in the same way as another person (Exs. (11) (12)).

(11) a. *kekkon=suru =made =wa, watasi =mo kokyō =o*
 marriage=do =till =TOP 1sg =also hometown=ACC

koisiku omoi -masi =ta
 miss -POL =PST

“I too missed my hometown until I got married.”

[Survey 1, mean: 4.12, variation: 1.42, SD: 1.19, median: 5]

b. *kekkon =suru =made =wa, kare =mo kokyō =o*
 marriage=do =till =TOP 3sg.M =also hometown=ACC

koisiku omoi -masi =ta
 miss -POL =PST

“He too missed his hometown until he got married.”

[Survey 1, mean: 2.65, variation: 2.17, SD: 1.47, median: 2]

(12) a. ?? *tyūgakusei =ni naru =made =wa, watasi =wa*
 middle school student =into become =till =TOP 1sg =TOP

kaeru =wa hatyūru =da =to omoi =masi =ta
 Frog =TOP reptile =COP =QUOT think =POL =PST

“Until middle school, I thought a frog was a reptile.”

[Survey 2, mean: 2.42, variation: 2.04, SD: 1.43, median: 2]

b. ?? *tyūgakusei =ni naru =made =wa, kare =wa*
 middle school student =into become =till =TOP 3sg.M =TOP

kaeru =wa hatyūru =da =to omoi =masi =ta
 Frog =TOP reptile =COP =QUOT think =POL =PST

“Until middle school, he thought a frog was a reptile.”

[Survey 2, mean: 1.92, variation: 1.49, SD: 1.22, median: 1.5]

Among the two types of self under consideration, the <feeling self> is treated in the same way as a present-tense expression. That is, (Ex. (11b)), which would sound unnatural if used for another person, can be expressed using the form *omoi-masi=ta* that is used to refer to oneself (Ex. (11a)). However, this form would be unnatural in an expression referring to the <thinking self> (Ex. (12a)). Ex. (12a) was less natural sounding than Ex. (10a) above by a statistically significant degree ($p < 0.01$).^{vii}

In this way, whether the <thinking self> can be expressed in a self-expressive way depends on the situation of the utterance. When speaking of the past, which is removed from the situation of the utterance (the actual reality here and now), the self appearing in the utterance is expressed as oneself if it is the <feeling self> but is treated as another person in the case of the <thinking self>. Just as the sound of clapping hands attenuates and disappears in an instant, self-expressiveness also attenuates when referring to the past.

Discoloration of Animacy

Basically the same phenomenon seen in self-expressiveness can be observed in the case of animacy as well.

It should be noted first of all that animacy is not a concept that concerns the thing referred to itself. Animacy is a concept concerning the image of the thing created by the speaker's words (i.e., the degree to which the speaker expresses the thing as a living image). William Croft's cross-linguistic observation that animacy tends to be related to person, with persons expressed in the first (e.g., "I") and second persons (e.g., "you") tending to have greater animacy than those expressed in the third person (e.g., "s/he") (Croft 112-113), can probably be considered to be premised on animacy's being a concept that concerns the world created through words.

But why does a person expressed in the first or second person have greater animacy than one expressed in the third person? While Croft does not comment on this question, the answer probably is because the degree of animacy is judged based on the situation of the utterance. The speaker "I" (first person) is speaking in the situation of the utterance and has a feeling of reality. The listener "you" (second person) also is in the situation of the utterance, at least in a case of in-person verbal communication, which is the most basic form of communication, and also has a feeling of reality. In contrast, "he" or "she" appearing in the utterance (third person) is not a person expected to be present in the situation of the utterance and thus does not have such a feeling of reality. Animacy is considered to be high when there is a feeling of reality, and the animacy of a thing undergoes fading

and discoloration in an expression that does not have this feeling of reality. (See section *Animacy and number distinction* for details.)

The same can be said concerning Croft's description of animacy as tending to relate not only to person but also to nominal forms in various languages – that is, something expressed using a personal pronoun tends to have greater animacy than something expressed using a noun (Croft 112-113). As seen, for example, in the way it is difficult to refer as *kare* (“he”) to a person whose actual existence is doubted (Ex. (13)), a pronoun is more of a real expression than a noun (Takubo), and for this reason alone it is likely to be considered to have greater animacy.

- (13) [On hearing that a mysterious youth supposedly haunts a forest]
- a. *somosomo seinen -tte hontoni iru =no?*
 in the first place youth -QUOT really exist =IM
 “But does the youth even really exist in the first place?”
 [Survey 1, mean: 4.28, variance: 1.29, SD: 1.14, median: 5]
- b. ?? *somosomo kare -tte hontoni iru =no?*
 in the first place 3sg.M -QUOT really exist =IM
 “But does he even really exist in the first place?”
 [Survey 1, mean: 2.39, variance: 2.09, SD: 1.45, median: 2]

This difference in animacy due to differences in realism can also be observed in the differences between the Japanese-language verb expressions of existence: *iru* and *aru* (“exist”). Since things that have high levels of animacy are expressed using *iru* and those with low levels of animacy with *aru*, a thing's degree of animacy can be discussed in terms of the naturalness of using an *iru* or *aru* expression. Consider Examples (14) and (15) below, expressions concerning the presence of applicants.

- (14) a. *ima koko =ni kibōsya =ga iru.*
 now here =LOC applicant =NOM exist
 “An applicant is here now.”
 [Survey 2, mean: 4.33, variation: 1.08, SD: 1.04, median: 5]
- b. ?? *ima koko =ni kibōsya =ga aru.*
 now here =LOC applicant =NOM exist
 “An applicant is here now.”
 [Survey 2, mean: 1.52, variation: 0.76, SD: 0.87, median: 1]
- (15) a. *mosi kibōsya =ga ire -ba sono kazu =o*
 if applicant =NOM exist -COND that number =ACC

- mi -te taiō =o kentō =si -masu.*
 see -CONN response =ACC consider=do -POL
 “We will consider our response in light of the number of applicants, if any.”
 [Survey 2, mean: 4.73, variation: 0.40, SD: 0.63, median: 5]
- b. ? *mosi kibōsya =ga are -ba sono kazu =o*
 if applicant=NOM exist -COND that number=ACC
- mi -te taiō =o kentō =si -masu.*
 see -CONN response =ACC consider=do -POL
 “We will consider our response in light of the number of applicants, if any.”
 [Survey 2, mean: 2.69, variation: 2.11, SD: 1.45, median: 3]

Since an applicant refers to a person who wants to apply for something – that is, a living human being – it has a very high level of animacy. The presence of an applicant is likely to be expressed using *iru* (Ex. (14a)), but not using *aru* (Ex. (14b)). But if the presence of applicants is merely a hypothetical expressed using the word *mosi* (“if”) (Ex. (15)), then the applicants lose their reality and are more likely to be imagined without very high degrees of animacy. Comparison of the naturalness of examples (14b) and (15b), in which the presence of applicants is expressed using *aru*, shows that Ex. (15b) was considered more natural to a statistically significant degree ($p < 0.01$). It would seem that it does not sound very unnatural to refer to the applicants using *aru* in this case because their animacy has undergone discoloration as a result of being considered as a hypothetical.

Furthermore, a similar phenomenon to that seen with a hypothetical can be seen in cases of negation as well (Ex. (16)).

- (16) a. ?? *kesa, yamazoi =no kokudō =de*
 This morning mountainous =GEN national road =LOC
- rakusekijiko =ga ari, makikom -are =ta keganin =ga*
 rockslide =NOM exist I nvolve -PASS =PST injured person =NOM
- futa -ri a =tta.*
 two -CL exist =PST
 “There was a rockslide on a mountain road this morning that injured two people.”
 [Survey 3, mean: 2.43, variation: 1.22, SD: 1.10, median: 2]^{viii}

b. ? *kesa*, *yamazoi* =*no* *kokudō* =*de*
 this morning mountainous =GEN national road =LOC

rakusekijiko =*ga* *a-tta=ga*, *makikom-are* =*ta*
 rockslide =NOM exist-PST=CONJ involve -PASS =PST

keganin =*wa* *naka* =*ta*.
 injured person =TOP nonexist =PST

“There was a rockslide on a mountain road this morning, but nobody was injured.”

[Survey 3, mean: 3.06, variation: 1.42, SD: 1.19, median: 3]

Example (16a), in which the presence of two injured persons is expressed using *aru*, is more likely to be considered to sound unnatural. But when their number is reduced to zero, and no injured persons exist in fact, their animacy undergoes discoloration. Example (16b), in which these nonexistent persons are expressed using *nai* (i.e. negative form of *aru*) instead of *i=nai* (i.e. negative form of *iru*), sounds a little less unnatural. The difference between the two examples is statistically significant ($p < 0.01$).

In this way, when discussing hypotheticals or unreality, which are removed from the situation of the utterance, the animacy of the things appearing in the utterance undergoes fading and discoloration. Like self-expressiveness (section *Attenuation of self-expressiveness*), the degree of animacy too is determined based on the situation of the utterance.

Discoloration of Esteem

The esteem with which a person deserving of respect is referred to also decreases due to discoloration when discussing a negated event (Ex. (17)).

(17) a. *ano kata* =*wa* *robī* =*ni* *iras: yaru*.
 that person (HON) =TOP lobby =LOC exist (HON)
 “That person is in the lobby.”

[Survey 2, mean: 4.06, variation: 1.53, SD: 1.24, median 5]

b. *ano kata* =*wa* *robī* =*ni* *iras:yara -nai*.
 that person (HON) =TOP lobby =LOC exist (HON)-NEG
 “That person is not in the lobby.”

[Survey 2, mean: 3.61, variation: 1.94, SD: 1.39, median: 4]

(19)	HIGH	LOW
A. PARTICIPANTS	2 or more participants, A and O	1 participant
B. KINESIS	action	non-action
C. ASPECT	telic	atelic
D. PUNCTUALITY	punctual	non-punctual
E. VOLITIONALITY	volitional	non-volitional
F. AFFIRMATION	affirmative	negative
G. MODE	realis	irrealis
H. AGENCY	A high in potency	A low in potency
I. AFFECTEDNESS OF O	O totally affected	O not affected
J. INDIVIDUATION OF O	O highly individuated	O non-individuated

(Hopper and Thompson 252)

Regarding “affirmation/negation,” they described expressions affirming the realization of an event as having higher levels of transitivity than expressions negating the realization of an event. Regarding mood, they described expressions concerning real events as having higher levels of transitivity than expressions concerning unrealized events. It does not seem out of place to see the standard of the situation of the utterance as being at the root of their observations. This is because, as we saw above when considering the animacy of things, the degree of reality differs between affirmative and negative expressions, and the concepts of reality and unreality are expressions of this difference.

It is possible based on this idea to understand the accusative case marking of potential verbs in Japanese in a crosslinguistic context. For example, the verb *nomu* (“drink”), as a verb that expresses an event that has a strong effect on a liquid by ingesting it orally into the body, has a high level of transitivity, and its object is expressed using the accusative marker *o* exclusively (Ex. (20)). However, for the verb *nom-eru* (“be able to drink”), which differs from *nomu* in expressing the unreality of potential, the accusative case marking is not obligatory (Ex. (21)).

- (20) a. *sake =o nomu.*
 sake =ACC drink
 “(A person) drinks sake.”
 [Survey 3, mean: 4.52, variation: 0.59, SD: 0.77, median: 5]

b.?? *sake* =*ga* *nomu*.
sake =NOM drink
 “(A person) drinks sake.”
 [Survey 3, mean: 1.35, variation: 0.71, SD: 0.84, median 1]

(21) a. *sake* =*o* *nom* -*eru*.
sake=ACC drink -possible
 “(A person) can drink sake.”
 [Survey 3, mean: 3.67, variation: 1.66, SD: 1.29, median 4]

b. *sake* =*ga* *nom* -*eru*.
sake =NOM drink -possible
 “(A person) can drink sake.”
 [Survey 3, mean: 4.25, variation: 0.83, SD: 0.91, median: 4]

In Ex. (20), while case (a), which indicates the noun *sake* using the accusative marker *o*, sounds highly natural, case (b), in which it is indicated using the nominative marker *ga*, does not sound very natural. The difference between the two is statistically significant ($p < 0.01$). In Ex. (21), however, it is not only the accusative in case (a) but also the nominative in case (b) that sounds natural, and in fact the nominative in case (b) is the more natural sounding of the two ($p < 0.01$).

In this way, unrealized events do not seem very real, and their transitivity is lower as a result. Thus, the degree of transitivity is determined based on the speaker present in the actual situation of the utterance.

Summary of Attenuation and Discoloration

In this section I have shown how in Japanese the self-expressiveness of a thing tends to attenuate and its animacy, esteem, and transitivity tend to undergo fading and discoloration as they are removed from the situation of the utterance (the actual reality here and now). This tendency has been seen in the degree of naturalness of combinations of the first-person pronoun *watasi* (“I”) with the verb *omou* (“think”) (Section *Attenuation of self-expressiveness*), the degree of naturalness of combinations of the nouns *kibōsyā* (“applicant”) and *keganin* (“injured person”) with the verbs *iru* and *aru* (“exist”) (Section *Discoloration of animacy*), the degree of naturalness of combinations of the verbs *iru* and *iras: yaru* (“exist”) with the negative *nai* (Section *Discoloration of esteem*), and the degree of naturalness of combinations of the noun *sake* with the nominative particle *ga* (Section

Discoloration of transitivity). In these ways, grammar (i.e. the structure of combinations of words) can be seen to have a situational quality.

Animacy and Number Distinction

A look at languages around the world shows a tendency for number distinctions to be more likely with regard to things that have higher degrees of animacy. While probably there is a need to probe further into the reason for this highly interesting tendency, and the human inclinations from which it arises, my study here will be based on this attested tendency.

In introducing this tendency, Croft (1990), while mentioning Silverstein and Dixon with regard to ergativity or agentivity, referred to the animacy hierarchy. He argued that languages tend to draw a boundary line at some point on this hierarchy. For a thing that has an animacy level that is above the boundary line, singular and plural forms will necessarily be distinguished, but a thing with a lower degree of animacy than the boundary line will not necessarily involve such distinctions.

As already mentioned (Section *Discoloration of animacy*), it must be noted here that Croft's animacy hierarchy, intended to match the actual state of various languages, is a general concept that incorporates not only animacy proper but also person animacy and NP-type animacy, as shown in (22).

- (22) Animacy: first, second-person pronouns < third-person pronoun < proper names < human common noun < nonhuman animate common noun < inanimate common noun

This combination of features has been named the animacy hierarchy. The most common representation of the animacy hierarchy is the one given above, found in Dixon (1979: 85), though its first modern description is found in Silverstein (1976). The animacy hierarchy actually involves several distinct but related grammatical dimensions. The first is a person hierarchy, in which first and second person outrank third person. The second is an NP-type hierarchy, in which pronouns outrank common nouns (there is some evidence that proper names occupy an intermediate position on this hierarchy). Finally, there is the animacy hierarchy in proper, in which humans outrank nonhuman animates, which in turn outrank inanimates. (Croft 112-113)

Animacy proper reflects the image of the thing itself. Inherently, human beings are viewed as more animate than animals, while animals are more animate than plants.

In contrast, person animacy and NP-type animacy reflect the image that arises from the way a thing is expressed. Specifically, person animacy concerns the image arising from differences in the grammatical person used to express something. The same thing will be expressed with higher degrees of animacy in the first and second persons than in the third person. NP-type animacy concerns the image arising from differences in the NP type (noun or pronoun) used to express a thing. The same thing will be expressed with a higher degree of animacy with a pronoun than with a noun.

Illustration (23) below depicts these three types of animacy in a form partially modified (through further division into parts (c) and (d)) by the author.

- (23) a. human being expressed using first, second-person pronoun
- b. human being expressed using third-person pronoun
- c. human being expressed using noun 1: high specificity
- d. human being expressed using noun 2: low specificity
- e. animal
- f. other

(Cf. Croft 111ff)

A language that places the boundary line between (a) and (b) in (1) is Guarani, spoken chiefly in South America. In this language, the singular and plural form differs only for human beings represented by first- and second-person pronouns. Specifically, in the first person the singular form is *šé* and the plural form *yané* or *oré*, in the second person the singular is *né* and the plural *peé*, but in the third person *haʔé* is used for both the singular and plural form. Chinese is a language that draws the boundary between (b) and (c), while in Japanese the boundary line lies between (c) and (d). As used here, “high specificity” refers to specification of the referent, as in *honjin* and *tōnin* in the Japanese examples (24) and (25).

- | | | | |
|---------------------------|-------------|-------------|-------------|
| (24) <i>honjin</i> | = <i>mo</i> | <i>sōki</i> | = <i>no</i> |
| person being talked about | =also | early time | =GEN |

person, who appears within the utterance, does not necessarily exist. This can be seen in the way a third person can be referred to hypothetically using the expression "if any." Even a third person who does exist may not necessarily be present for the utterance (a typical example is gossip) and thus would not have a high degree of reality. For this reason, the animacy of a third person is not high.

NP type has a similar effect. A pronoun has more reality than a noun. As one example, in a Japanese utterance that involves doubt concerning something's existence it would sound unnatural to refer to that thing using a pronoun but would not sound unnatural to refer to it using a noun. This can be seen in Example (13) above, which is reproduced as Example (26) below.

- (26) = (13)
- [On hearing that a mysterious youth supposedly haunts a forest]
- a. *somosomo seinen -tte hontoni iru =no?*
 in the first place youth -QUOT really exist =IM
 "But does the youth even really exist in the first place?"
 [Survey 1, mean: 4.28, variance: 1.29, SD: 1.14, median: 5]
- b. ?? *somosomo kare -tte hontoni iru =no?*
 in the first place 3sg.M -QUOT really exist =IM
 "But does he even really exist in the first place?"
 [Survey 1, mean: 2.39, variance: 2.09, SD: 1.45, median: 2]

In an utterance that questions the existence of the mysterious youth under discussion, Example (a), which refers to the youth using the common noun *seinen*, sounds natural while Example (b), which refers to him using the pronoun *kare*, sounds unnatural. This example shows how a pronoun has greater realism, and thus greater animacy, than a noun.

Here, I apply the same categorization of specificity (differentiation between (c) and (d)) to the animacy hierarchy (23). The higher the degree of specificity, the greater the realism and the higher the animacy.

Animacy and Syntactic Structure

As seen above, person animacy refers to how the same thing is expressed with higher degrees of animacy in words in the first and second persons than in the third person. NP-type animacy refers to the way the same thing

can be expressed with a higher degree of animacy using a pronoun than with a noun. These can be considered animacy hierarchies that concern the forms of words used to express things. Now let us consider another animacy hierarchy, one concerning syntax, or the combinations of words used to express things. Under this hierarchy, a thing is expressed with a higher degree of animacy in an appositive structure than in a noun-modifier structure.

This syntactic animacy can be backed by observation of the following two points concerning the Japanese language. First, in considering word formation using the plural suffix *-tati*, when introducing a new person to the conversation an expression in an appositive structure tends to sound less natural than one in a modifier-modified structure. See Examples (27) and (28) below.

- (27) a. *aru hi, totuzen, mura =ni, san -nin =no*
 one day suddenly village =DAT three -CL =GEN
seizika -tati =ga yatteki =ta.
 politician -PL =NOM come =PST
 “One day, suddenly, three politicians came to the village.”
 [Survey 4, mean: 4.13, variance: 1.04, SD: 1.02, median: 4]^x
- b. *aru hi, totuzen, mura =ni, seizika -tati*
 one day suddenly village =DAT politician -PL
san -nin =ga yatteki =ta.
 three -CL =NOM come =PST
 “One day, suddenly, three politicians came to the village.”
 [Survey 4, mean: 3.59, variance: 1.31, SD: 1.14, median 4]
- (28) a. *mura =no sisan =ni me =o hikar-ase*
 village=GEN property =DAT eye -ACC light-CAUS
-teiru =no =wa, sontyō =o nozoke =ba,
 -CONT=GEN =TOP village headman =ACC exclude=COND
reino san-nin =no seizika -tati
 in question three-CL =GEN politician -PL
-gurai =dearu.
 -only =COP
 “Excluding the village headman, it was only the three politicians in question who keep a watchful eye on the village's property.”
 [Survey 4, mean: 3.98, variance: 1.29, SD: 1.13, median: 4]

b. <i>mura</i>	= <i>no</i>	<i>sisan</i>	= <i>ni</i>	<i>me</i>	= <i>o</i>	<i>hikar-ase</i>
village	=GEN	property	=DAT	eye	-ACC	light-CAUS
<i>-teiru</i>	= <i>no</i>	= <i>wa,</i>	<i>sonyō</i>		= <i>o</i>	<i>nozoke =ba,</i>
-CONT	=GEN	=TOP	village headman	=ACC	exclude	=COND
<i>reino</i>	<i>seizika</i>	<i>-tati</i>	<i>san</i>	<i>-nin</i>	<i>-gurai</i>	= <i>dearu.</i>
in question	politician	-PL	three	-CL	-only	=COP
“Excluding the village headman, it was only the three politicians in question who keep a watchful eye on the village’s property.”						
[Survey 4, mean: 3.84, variance: 1.32, SD: 1.15, median: 4]						

In Example (27), the three politicians are introduced newly to the discussion, while in Example (28) they are referred to as characters already introduced. In each of these examples, the three politicians are referred to using a modifier-modified structure in which “three” modifies “politicians” in (a) and using an appositive structure in which “politicians” and “three” appear in apposition in (b). According to our survey, none of these four sentences was evaluated to be so lacking in naturalness as to be described as unnatural. However, comparison of the four shows that in Example (27) the naturalness of (b), which has an appositive structure, appeared quite low. Specifically, there was no statistically significant difference in naturalness between (a) and (b) in Example (28) ($p = 0.197$) or between (27a) and (28a) ($p = 0.090$), but the naturalness of (27b) was lower to a statistically significant degree than that of both (28b) ($p < 0.05$) and (27a) ($p < 0.01$). This shows that a thing in an appositive structure requires a high level of specificity, and by extension a high degree of reality. At the stage at which they are introduced as a new element in the conversation, the “politicians” are not yet specified and do not seem very real.

The second point concerns the fact that while the plural form of a common noun such as *musume-tati* (“young women”) is not particularly unnatural sounding whether it appears in an appositive expression (Example (29a)) or in a modifier-modified expression (Example (29b)), the plural form of a personal pronoun such as *kanozō-tati* (“they”), which has a higher degree of animacy, is more limited in meaning when it appears within a modifier-modified expression (Example (30a)) than in an appositive expression (Example (30a)). This is the appositive structural meaning of “three females” as seen in Example (31).

- (29) a. *musume* *-tati* *san* *-nin*
 young woman -PL three -CL
 “three young women”
- b. *san* *-nin* =*no* *musume* *-tati*
 three -CL =GEN young woman -PL
 “three young women”
- (30) a. *kanozyo* *-tati* *san* *-nin*
 3sg.F -PL three -CL
 “three females”
- b. *san* *-nin* =*no* *kanozyo* *-tati*
 three -CL =GEN 3sg.F -PL
 “three females”
- (31) *hito -ri fueru -gotoni sore =dake =no tsuityōkin =o*
 one -CL increase -everytime that =amount =GEN surcharge =ACC
- tor -areru =nara, anata -gata =wa futa -ri =dakara*
 take -PASS =COND 2sg -PL =TOP two -CL =since
- madaïi* =*tositemo*, *san* *-nin* =*no*
 still no problem =even if three -CL =GEN
- kanojo* *-tati* =*wa* *taihen* =*darō*.
 3sg.F -PL =TOP terrible =INFR
 “If there’s a surcharge for each additional person, then even if it would be no
 problem for the two of you, it probably would be quite expensive for those
 women who are three in number.”

Now let us consider the Chinese language. In fact, Chinese is not thoroughly averse to number agreement. Number agreement sounds natural for personal pronouns referring to human beings, at the highest level on the animacy hierarchy, when used in appositive expressions that can be considered to have high degrees of animacy. See Examples (32) and (33) below.

- (32) a. *wǒ -men liǎng -ge*
 1sg -PL two -CL
 “we two”
- b. *nǐ -men liǎng -ge*
 2sg -PL two -CL
 “you two”
- c. *tā -men liǎng -ge*
 3sg -PL two -CL
 “they two”

- (33) * *nǚháir* -*men* *liǎng* -*ge*
 girl -PL two -CL
 (Lit.) “the girls two”

In Example (32), in each of the cases (a), employing the first-person pronoun *wǒ* (“I”), (b), employing the second-person pronoun *nǐ* (“you”), and (c), employing the third-person pronoun *tā*, “s/he”), personal pronouns representing human beings can be made plural with the plural suffix *-men*, to agree in number with the following *liǎng-ge*, “two persons.” Unlike pronouns, the noun *nǚháir* (“young girl”) referring to a human being cannot be made plural with *-men* to agree in number with *liǎng-ge*. As shown by the asterisk in front of it, Example (33) sounds unnatural.

To be perfectly clear, this is not simply a consequence of the fact that *wǒ*, *nǐ*, and *tā* here are pronouns representing individual people (Section *Animacy and number distinction*), necessarily distinct from *wǒ-men*, *nǐ-men*, and *tā-men*. This is not an issue of common sense but one of language. There are two reasons for considering it in this way. The first is because number agreement is not obligatory for *wǒ*, *nǐ*, and *tā* either. Real examples (34), (35), and (36) below, elicited from famous novels, illustrate the first, second, and third persons, respectively.

- (34) *wǒ liǎng -ge méiyǒu jièshàoxìn néng bù néng dēngjì?*
 1sg two -CL nonexist letter of reference can NEG can submit
 “May the two of us submit a marriage notice even though we do not have a
 letter of recommendation?”
 [Dēngjì, by Shùlǐ Zhào, 1950]
- (35) *nǐ liǎng-ge dào hěn tóujī =ma.*
 2sg two -CL unexpectedly very hit it off =SFP
 “The two of you have hit it off surprisingly well, haven’t you?”
 [Dōngcángjì, by Pú Zōng, 2000]
- (36) *zhè -zhi kāi wǎng- zhōngguó de chuán*
 this -CL go for China LM ship

sìhu jiù zhǐ děnghòu tā liǎng -ge.
 seem ADV only wait 3sg two -CL
 “It appears that this boat for China will wait for the two of them only.”
 [Ālísī Zhōngguó Yóujì, by Cóngwén Shěn, 1928]

The second reason is because the naturalness of number agreement varies with the linguistic structure. While Examples (34)-(36) above have appositive structures, even the cases of *liǎng-ge wǒ*, “the two of us,” *liǎng-ge nǐ*, “the two of you,” and *liǎng-ge tā*, “the two of them,” which have modifier-modified structures instead of appositive ones, would sound unnatural unless interpreted in special ways, such as references to multiple personalities.

As seen above, number agreement is not completely unheard of in Chinese. Rather, it is more likely in an appositive structure than in a modifier-modified structure. This should be considered a result of the fact that since an appositive structure is a realistic structure with a high degree of specificity the things expressed in such a structure have high levels of animacy, and thus number distinction is used not only for the quantifiers but for the personal pronouns as well.

Summary

This article has shown, using the example of number agreement, that an ecological perspective with a focus on the situation of the utterance (here-now-I-real) is an effective way to understand differences in grammar among individual languages and to identify commonalities above and beyond such differences. On the surface, English and Chinese would appear to be exact opposites in terms of whether or not they conform to the rule of number agreement, while the position of Japanese would appear to be ambivalent on this question. But in fact this is not the case. Chinese, Japanese, and English do conform to number agreement in similar ways. They do not either conform to or violate number agreement capriciously. Animacy is the key to understanding this issue. Numerous languages around the world are consistent in the way number distinction is more likely to arise the higher the subject's position on the animacy hierarchy. Chinese and Japanese are no exceptions. Leaving aside differences in individual circumstances, languages differ only in the boundary lines on the animacy hierarchy where number distinction becomes obligatory.

To reach the above understanding, it is essential to see animacy from an ecological perspective. To the speaker, whose frame of reference is

centered on the situation of the utterance, things with a high feeling of reality have high animacy while other things have low animacy. This is why person and NP type affect the level of animacy, and the same is true of the effect of the difference between a modifying-modified structure and an appositive structure. An ecological perspective is essential as a principle for explaining levels of not only animacy but also self-expressiveness, esteem, and transitivity. The prevailing view of language as separate from the situation of the utterance certainly appears to require considerable revision on the subject of grammar.

Notes:

ⁱ We thank our colleagues from the Graduate School of Letters, Kyoto University, especially Rui-jia Zhang, for their help. This study summarizes and further builds on Sadanobu (“Hokusūsei to riaritī”) and Sadanobu (“Bumpō no bamensei”).

This study was partially supported by the Ministry of Education, Culture, Sports, Science and Technology Grant-in-Aid for Scientific Research (S) 20H05630.

ⁱⁱ The asterisks before (1b) and (2a) indicate that the expressions are extremely unnatural sounding.

ⁱⁱⁱ Abbreviations used are as follows: 1sg (first person singular), 2sg (second person singular), 3sg (third person singular), ACC (accusative), ADV (adverb), CL (classifier), COND (conditional), CONJ (conjunctive), CONN (connective), CONT (continuous), COP (copula), DAT (dative), F (feminine), GEN (genitive), HON (honorific), HUM (humble), IM (interrogative marker), INFR (inferential mood), LM (linking morph), LOC (locative), M (masculine), NEG (negative), NOM (nominative), PASS (passive), PL (plural), POL (polite), PST (past), QUOT (quotation), SFP (sentence-final particle), and TOP (topic).

^{iv} Survey 1 of 113 native Japanese speakers was conducted online in April 2021. Its respondents consisted of 36 males, 74 females, and three unknown, in the following age ranges: three persons aged 10–19, 21 aged 20–29, 31 aged 30–39, 31 aged 40–49, 13 aged 50–59, seven aged 60–69, four aged 70 or older, and three unknown. However, no major differences were apparent by gender or age. (The same hereinafter.) The respondents could play the utterances as many times as they liked in any order, with no time limit for their responses.

^v Hereinafter, means, variance, and SD shown are rounded off to two decimal points.

^{vi} Survey 2 was conducted online in May-June 2021. Its respondents, 100 native Japanese speakers, consisted of 51 males and 49 females, in the following age ranges: four persons aged 10–19, 36 aged 20–29, 24 aged 30–39, 17 aged 40–49, three aged 50–59, 14 aged 60–69, and two aged 70 or older.

^{vii} Testing of differences in naturalness hereinafter employs the (two-tailed) Wilcoxon signed-rank test.

^{viii} Survey 3 was conducted online in June-July 2022. Its respondents, 143 native Japanese speakers, consisted of 72 males and 71 females, in the following age ranges: twenty-two persons aged 20–29, 20 aged 30–39, 25 aged 40–49, 25 aged 50–59, 25 aged 60–69, 25 aged 70 or older, and one unknown.

^{ix} Survey 4 was conducted online in August 2023. Its respondents, 100 native Japanese speakers, consisted of 51 males, 47 females, and two unknown in the following age ranges: fifteen persons aged 20–29, 15 aged 30–39, 17 aged 40–49, 18 aged 50–59, 18 aged 60–69, 14 aged 70 or older, and three unknown.

Works Cited

- Croft, William. *Typology and Universals*. Cambridge UP, 1990.
- Dixon, R. M. W. "Ergativity." *Language*, vol. 55, no. 1, Mar. 1979, pp. 59-138.
- Durie, Mark. "Towards an Understanding of Linguistic Evolution and the Notion 'X Has a Function Y'." *Discourse Grammar and Typology: Papers in Honor of John W.M. Verhaar*, edited by Werner Amraham, Talmy Givón, and Sandra A. Thompson, John Benjamins, 1995, pp. 275-308.
- Givón, Talmy. *On Understanding Grammar*. Academic Press, 1979.
- Hockett, Charles F. "The Origin of Speech." *Scientific American*, vol. 203, no. 3, Sept. 1960, pp. 89–97.
- Hopper, Paul J. "Emergent Grammar." *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society, 1987, pp. 139-157.
- Hopper, Paul J., and Sandra A. Thompson. "Transitivity in Grammar and Discourse." *Language*, vol. 56, no. 2, June. 1980, pp. 251-299.
- Sadanobu, Toshiyuki. "Hokusūsei to Riaritī (Plurality and Reality)." *Nihon Jinbun Syakai Kenkyū*, no. 2, Nov. 2019, pp. 1-13.
- . "Bumpō no Bamensei (The Placement of Grammar)." *Papers from the 23rd National Conference of the Japanese Cognitive Linguistic Association*, Japanese Cognitive Linguistic Association, 2023, pp. 22-34.
- Shibuya, Katsumi, Yoko Ieiri, and Hiroyuki Takada. "Rekisisyakaigengogaku no Kisotisiki (Basic Knowledge of Historical-Socio Linguistics)." *Rekisisyakaigengogaku Nyūmon: Syakai kara Yomitoku Kotoba no Uturikawari (An Introduction to Historical-Socio Linguistics: Analysis of Language Change from Social Perspective)*, edited by Hiroyuki Takada, Katsumi Shibuya, and Yoko Ieiri, Taishukan, 2015, pp. 5-42.
- Silverstein, Michael. "Hierarchy of Features and Ergativity." *Grammatical Categories in Australian Languages*, edited by R. M. W. Dixon, Australian Institute of Aboriginal Studies, 1976, pp. 112-71.
- Takubo, Yukinori. "Meisiku no Modaritī (Modality of Noun Phrases)." *Nihongo no Modaritī (Japanese Modality)*, edited by Yoshio Nitta and Takashi Masuoka, Kurosio Publishers, 1989, pp. 211-233.