WHAT ARE GENERIC SENTENCES ABOUT?

Chungmin Lee clee@power1.snu.ac.kr
Seoul National University

This paper argues that generic sentences constitute Topic constructions in the sense that they talk about generic NPs involved, which are approximately definite. It further argues that a Topic construction shows some topical relation between its Topic operator and its related predicate. The topical relation is intensional, so the predicate shows a property or characteristics rather than a stage or episode. Thus, even a Proper Noun Topic, in its intensional sense, can form a generic sentence, with habitual aspect.

Generic sentences, sentences talking about generic NPs, are Topic constructions in all different languages. Generic NPs in such sentences become Topics easily and naturally1. It is because prototypical Topics are those NPs placed at the head of a sentence (in the Topic position, in other words) as NPs denoting things assumed by the speaker to be familiar to both speaker and hearer to talk about (definite, in this case)2. Naturally, generic NPs in this case, are argued to be semantic definites or approximated to definites, no matter whether they are syntactically definite or indefinite, mass or count, and singular or plural, language internally or cross–linguistically, as witnessed in the following generic sentences:

(1) a. The bird flies.
b. Birds fly.
c. A bird flies.
(2) a. Water is transparent.
(3) a. Les oiseaux volent.
b. L'eau est transparent.
(4) a. say –nun nal –n –ta
bird Top fly Pres Dec
"Bird' flies'
b. mul –un thwumyeng–ha –ta
water Top transparent Dec
'Water is transparent'

Generic NPs normally function as kind–referring NPs or default–set names. The names of kinds (/sub–kinds) and categories, represented by common nouns, are easily assumed to be familiar in common ground and easily accessed in any neutral context and can be talked about regarding their permanent or persistent properties or characteristics via information–giving predicates. Therefore, a generic statement is basically a Topic construction, involving resource situations and generalized quantifier or tripartite semantic structure. Hence, the pragmatic Topic–Comment

information structure. In this structure, the chosen Topic binds its Comment in the sense that the Comment is necessarily about the Topic NP. A Topic sentence has a topical relation between the Topic and the Comment. The assertion made in this structure is based on the speaker's cognitive 'topical judgment' (similar to, but broader than Kuroda's [1] 'categorical judgment'), whereas an episodic piece of information is based on the speaker's propositional act at the perceptual level and it is rendered in a different form in certain languages like Korean and Japanese, with the subject marker on the reference NP. Take a look:

(5) say –tul–i nal–a ka–ko iss –ta
bird Pl Nom fly go Prog Dec
'Birds are flying'
(6) say –tul–i nal–a ka–ko iss –ney!
bird Pl Nom fly go Prog Surprise
'Birds are flying'

Therefore, the Surprise epistemic marker *–ney* can be attatched to the end of an episodic sentence, as in (6), but not to the end of a Topic sentence, as follows:

(7) ?*say –nun nal–ney!
bird Top fly Surprise
'Birds fly!'

Then, let us consider why different languages show different generic NP forms, as in (1) through (4). The most typical generic NPs are kind–referring NPs and they are represented by singular common nouns in a majority of articleless languages such as Korean, Japanese and Chinese. They behave just like proper names, as argued by Carlson [2]. And in many languages Pl(ural)–marking is optional or lacking. Even in Pl–marking languages, Pl–making is acquired rather late. In English, it first occurs somewhere between one-and-a-half and two (Radford [3]) and in Russian, the distinction between mass and count nouns is reported to be stabilized only at age 8 (Slobin [4]). The prototypical generic NP form, in human languages, then, must be a singular common noun. When a kind is referred to, its category is well–defined and naturally can be assumed to be definite semantically. All the categories represented by generic NPs have their members, differently from Proper Nouns, and the plural number of members is reflected in Pl–marking. Then, the logically possible combination of markings for a common noun phrase must be the following:

(8) (I) COMMON NOUN (Singular), e.g. Korean, Japanese, Chinese, mass terms in English
(II) Def–marking + COMMON NOUN (Singular), e.g. English, mass terms in French
(III) COMMON NOUN + Pl–marking, e.g. bare Pl in English
(IV) Def–marking + COMMON NOUN + Pl–marking, e.g. French

(V) Indef–marking + COMMON NOUN (Singular),
e.g. English

English has all the possible forms (except type IV) and articleless and non–Pl–marking languages show the simplest form: COMMON NOUN (Singular) (type I). The bare Pl–form in English is definite when used as a generic NP in the Topic position (and possibly in the attitude or dispositional sentences)3. Only the existential reading of the form is indefinite. And the existential bare Pl cannot take the Topic position. It is closely attatched to the predicate (as in Diesing [5] and Kratzer [6]), even though it is at the head of a sentence on the surface.

In English, the NP form of [Definite Article + Common Noun (Singular)] as a generic NP reflects the semantic definiteness of the noun that represents a singular category of relevant individual entities. The Definite Article applies to the unique kind/type meaning of the noun and not to the meaning of a unique token member of the kind/type. Then, what about the bare Pl form? The Pl–marking reflects the plural number of individual entities of the kind category involved, and when the number of member entities approaches the whole category/set, the bare Pl can be approximated to a definite NP. Even the form [Indefinite Article + Common Noun (Singular)] can get a generic sense in the Topic position, but not in the Object position, with its meaning of the representative of arbitrary choice from among the members of a category indicated by the common noun. Because of the meaning of representativehood and arbitrary choice from among all the members of a category, the [Indefinite Article + Common Noun] phrase can be said to be quasi–definite. The nature of arbitrary choice from among all the members of a category gives the impression of indefiniteness (cf. Heim 1982). And because of its explicit Indef–marking, this form cannot occur with a kind–predicate such as 'extinct' and a 'dynamic' (/'avant–garde') predicate such as 'set foot on' and 'reach,' even though it can occur with such a quantificational predicate as 'rare'. ('Extinct' can be said to be a particular case of quantificational predicate.) Transitive kind predicates such as 'extinguish' and 'invent' require the typical form in English [Definite Article + Common Noun (Singular)]. The notion of set and members alone cannot handle the case of the kind-generic NP applied to the predicate like 'invent.' Consider the following:

(9) a. ?*A dodo is extinct.
b. ?*A rat reached Australia in 1770.
[generically]
(10) A firefly is rare.
(11) Koreans invented the printer (??printers).

On the other hand, such articleless languages as Korean and Japanese employ the Topic marker *–nun* and *–wa*, respectively, and another articleless language Finnish uses the nominative marker in contrast to the partitive marker to express a Topic. Even in English, the Topic construction has an optional pause between the Topic and the rest of the sentence, whereas an episodic event–report does not have one (compare the glosses of (7) and (5)).

Then, why do I claim that generic sentences are Topic constructions? Consider the

following:

(12) Beavers build dams.
(13) a. Dams, beavers build them.
b. As for dams, beavers build them.
c. Dams are built by beavers.

In (12), 'beavers' is the Topic and because it is a Topic it encompasses the unique category of all the beavers normally, and because it is a Topic, it has a wide scope over another bare Pl in the object position. Consequently, 'dams' in (12) receives an existential reading. On the other hand, in (13) all the cases of 'dams' are Topics and this time 'dams' has a wide scope over 'beavers.' Then, all the sentences in (13) talk or make a generalization about 'dams,' not about 'beavers' this time (becoming factually false). Consequently, 'beavers' becomes existential. Even if it remains as a generic NP, it becomes a neutralized subject, not a Topic any longer. This situation also applies to embedded generic subject NPs.

As a footnote, the left–dislocation construction (13a) is analogous to a genuine case of Topic, but not the 'Topic' created by the so–called 'Topicalization,' which is identified with wh-movement by Chomsky [8]) and is closer to a contrastive element. This mistaken notion of Topic is being perpetuated by Chomsky although Gundel [9] indicated the relevant distinction. On the other hand, Chae [10] also tried to relate genericity to Topic, though not explicitly.

Similarly, ambiguity arises in a sentence with a definite but not an indefinite Locative, as follows:

(14) Hurricanes arise in this part of the Pacific.
(15) a. Uragany voznikajut v etoj casti Tixogo Okeana
Hurricanes arise in this part of the Pacific Ocean.
b. V etoj casti Tixogo Okeana voznikajut uragany
In this part of the Pacific Ocean arise hurricanes.
(Nakhimovsky [11])

The ambiguity is resolved in another articleless language Russian, with the unmarked, fixed Topic + Comment order, as in (15) above. A Topic in the adverbial or Object position is markedly possible in English, but in this case the definiteness must be speicified even syntactically. Frequently, a spatio–temporal location becomes a background Topic to make the following kind of habitual sentence acceptable:

(16) (??)It rains. (Carlson [12])

By (16), 'here and recently' as a background Topic or some other restrictive term (in the tri-partite semantic structure) such as a conditional (like 'when there are black clouds') or some quantifier plus restrictor expression (like 'every day') must be understood. Otherwise, (16) becomes odd.

Let us turn to the issue of the dual nature of the kind–token relation. As far as I can see, the generic NP shows a dual relation between the kind/category and its token individual members. In the case of kind–predicates like 'extinct,' the

kind–level and a part–whole relation seems to be involved in expressing the process of 'getting extinct' or 'dying out'. When the part or subset of living things becomes smaller, we apply those expressions. Atomic individuals do not particularly count. In the following case of collective property, the sum individual matters. Take a look:

(17) Germans/The German bought 83,000 BMWs last year. (Krifka)

On the other hand, the following example shows a habitual distributive activity on the part of atomic individuals:

(18) The average American family eats at McDonald's twice a week. (Carlson [12])

Therefore, I can attempt to show the following kind of Topic operation:

(19) TOPIC $\left[\begin{array}{cc} X & X \\ \; ; & \\ x & x \end{array}\right]$ ([birds' (Xx), [fly' (Xx)])

Here, the Topic operator binds the variables, relating the antecedent and the consequent. The upper level capital letter variables represent kind–level and the lower–level small letter variables represent token members. If both capital and small letter variables show up in the restrictor–nuclear scope part, then it must mean a generically distributive event. If the NP refers to the kind only, the capital letter variables alone must appear, perhaps as in the case of invention predicate. And Topic inherently has a conditional/near-necessity meaning. Let me further try to show the translations of (12) and (13a).

(20) TOPIC $\left[\begin{array}{cc} X & X \\ \; ; & \\ x & x,\ y \end{array}\right]$ ([beavers' (Xx)], [dams' (y) & build' (Xx, y)])

(21) TOPIC $\left[\begin{array}{cc} X & X \\ \; ; & \\ x\ x, & y \end{array}\right]$ ([dams' (Xx)], [beavers' (y) & build' (y, Xx)])

The dual relation between kind and token is reflected in colloquial French. Observe:

(22) Les chats, c'est gentil. (Auger)

Even though the Topic is in the Def Pl, the pronominal ca shows Singular agreement. The definite article semantically applies to the kind category but the Pl applies to its token members, and then the pronominal *ca* again applies to the semantic kind category. We can employ different syntactic devices to mark the definiteness of a kind, the plurality of its members (this part gives the impression

of indefiniteness), or both, or alternatively as here. As Cooper [13] argued, there must be a resource situation distinct from a described situation to determine the range of quantification for the use of an NP in different contexts. And the idea must be extended and modified to be applied to different uses of bare common nouns in articleless languages, to mass terms and bare plurals in article languages, and to proper names and definite descriptions involving generic predicates as well. To handle the latter cases, he might need an abstract null operator. However, the Topic operator is a general operator proposed to cover all the different cases. Cooper's proposal of a generic situation type S for the restrictor/antecedent argument seems to be a good start, although we have much left to do. For instance, we must reconsider treating genericity as 'universal' quantification.

The issue of a small subset reading of generic sentences is not easily solved, of course. Consider:

(23) Finns are excellent skiers.

Only a very small subset, perhaps less than 3% of Finns may be 'excellent skiers' but (23) is an excellent generic sentence. Why is it so? I would argue that even if we start with the total set of normal Finns, based on nonmonotonic reasoning, because of the predicate expression, the relevant portion of Finns is the subset of skiers. In the predicate expression, the focal, asserted part is 'excellent' and the part 'skiers' can be accommodated as a presupposition such as Finns, if they are involved in the event of skiing, they are excellent skiers. 'Finns' is a Topic, binding the rest of the sentence. There seems to be a subtle interplay involved between the critical ratio arising from the Topic generalization and the critical mass coming from the predicate presupposition. Presupposition accommodation is thus proposed to solve the problem of a poor subset interpretation. Then, it will be incorporated into the restrictive term in the tripartite semantic structure of Quantifier–Restrictor–Nuclear Scope. However, if we have some explicit modification attatched to NPs, then there arises an intersectional reading and the consequence is that the modified NP can hardly constitute a generic NP. If the modified NP is an indefinite NP, it can be all right. But this is the least typical generic, with its contrastive meaning because of the modifier, if it is a generic at all. Accommodation, which can be accommodated into 'resource situations,' can be extended to the problem of an 'extrinsic' Topic such as the following in Korean. Take a look:

(24) Coffee-nun cam -i an o-a.
Top sleep Nom not come
'Lit.: Coffee, sleep does not come.
/we cannot sleep.' (Ahn, Kim, & Lee [14])

The accommodated presupposition could be 'if we drink coffee, the event of drinking it causes....' A generic situation type is involved here. Otherwise, it is impossible to get a coherent predicational, topical relation. This kind of somewhat far-fetched accommodation can occur when there can be a certain plausible causal/logical link between the extrinsic Topic and the associated predicate.

How about a proper name Topic as in (32) below. It can be regarded as an intensional object consisting of different properties and the sentence is a

generalization about or characterization of 'John.' It can be assumed to be familiar to the hearer and to be present in the domain of discourse. A pronoun can also serve as a Topic. But it typically refers to an already introduced entity and must go back to the referent to represent its properties.

Middles in English and some other Indo-European languages are also a type of Topic construction as well as generic statement; they do not report time-bound events but attribute properties to or characterize objects (Fagan [15]). Observe:

(25) This book sells/reads well.

This nature of middles is made possible by the simple present tense for timeless propositions and the implicit generic human agents involved, which change otherwise stage-level predicates to individual-level predicate interpretations. The Topic NP 'this book' is interpreted as referring to a subkind (title) rather than a particular copy of book. So, only those verbs that can show the property of the Topic can be employed in the construction. The 'tough' construction is another kind of Topic construction, generalizing over the NP in the Topic position.

On the other hand, a common noun stands for a commonly known category of things in the world and can be relatively easily referred to when necessary. Children might perceive categories as masses first and individuate their members gradually (Lee [16]). Particularly in articleless languages, bare common nouns can be rather easily employed as Topic to generalize over their members or their categories, by means of Topic devices.

The givenness hierarchy (given by Gundel et al [17]) shows 'the N' but not in its generic sense and we cannot be quite sure about the status of the generic Topic common noun in the hierarchy.

The embedded cases or listing cases of generic statements cannot be prototypical instances of either Topic sentences or generic statements. The embedded cases are ones that are suppressed as neutral information and the listing cases (such as 'Flowers bloom, birds sing, and children play') are description of a situation in which those categories are contrasted.

Last of all, I would like to make it clear that a Topic cannot be topically related to its Comment predicate if the latter is simply for an episodic event–report, based on transitory perception. A real Topic sentence must be based on the topical judgment of high–level cognition. In that sense, topical relation or operation is mutual intensionality (as hinted by Carlson [12]). Consequently, the following cannot constitute Topics:

(26) ???Birds | are flying
(27) ???say –nun nal–a ka –ko iss –ta
bird Topic fly go Prog Dec
"Bird' is flying'

At most, those NPs can be contrastive, with certain contrast items in the speaker's mind. Thus, Diesing's [5] analysis regarding the position of the subject of a stage-level predicate in LF does not sound correct. Then, what about the dynamic/avant–garde predicates, which may pose a problem to Krifka and probably to Takami [19] also? Consider:

(28) Man set foot on the moon in 1969.
(29) The rat/Rats reached Australia in 1770.

These predicates are used in recording the relevant events as significant milestones from the perspective of the whole diachrony of the respective explorations/spreadings of the kinds. So, they are differently interpreted from their 'stage–level' or episodic readings for particular or existential events. In (29), the bare plural cannot be existential in this reading and necessarily forms a Topic. These are not simply 'specificational.' Even if an individual like John becomes a Topic, it must be interpreted as an intensional being, as a complex of different stages/phases, as already indicated. Thus, the following can hardly constitute a Topic sentence:

(30) (??) John | is eating.
(31) (??) John | fell.

Compare them with (32) below, which is a Topic.

(32) John | smokes. (cf. Krifka [20])

In conclusion, I can state that generic sentences are Topic sentences, and prototypical Topic sentences are generic sentences. In other words, they are weakly equivalent.

Endnotes

1. I would like to express my gratitude to Benjamin Tsou and others for their comments at the 8th Asian Conference on Language, Information and Computation held at Kyoto in 1994.
2. Gundel (1974) also alludes to the definiteness of a generic Topic (bare Pl) NP.
3. NPs such as 'the lions' are usually interpreted as taxonomic, definite subkinds by native speakers.

References

[1] Kuroda, S.-Y. The Categorical and Thetic Judgement. *Foundations of Language* 9. 153-85, 1972.
[2] Carlson, G. *Reference to Kinds in English*. New York: Garland, 1980.
[3] Radford, A. *Syntactic Theory and the Acquisition of English Syntax*. Oxford: Blackwell, 1990.
[4] Slobin, D.I. The Acquisition of Russian as a Native Language. In M. Smith and G.A. Miller(eds.) *The Genesis of Language*. Cambridge MA: MIT Press, 1966.
[5] Diesing, M. *Indefinites*. Cambridge MA: MIT Press, 1992.
[6] Kratzer, A. Stage-Level and Individual-Level Predicates, ms U. of Mass. at Amherst.
[7] Heim, I. The Semantics of Definite and Indefinite Noun Phrases. University of Massachusetts Dissertation, 1982.
[8] Chomsky, N. On Wh-movement. In P. Culicover et al (eds.) *Formal Syntax*. New York: Academic Press, 1977.
[9] Gundel, J. The Role of Topic and Comment in Linguistic Theory, University of Texas Austin Dissertation, 1974.
[10] Chae, Wan Topic and Genericity, Specificity, and Definiteness. (in Korean) In *Festschrift for Dr. Yu Changgyun on his 60th birthday*, 1984.
[11] Nakhimovsky, A. Pragmatic Factors in the Interpretation of Generics. NELS 20 vol. 2, 1990.
[12] Carlson, G. 1989 English Generic Sentences. In G. Chierchia et al (eds.) *Properties, Types and Meaning* II. Dordrecht: Kluwer.
[13] Cooper, Robin Generalized Quantifiers and Resource Situations. ms. U. of Edinburgh, 1991.
[14] Ahn, Sungho, Kwangsup Kim and Chungmin Lee Constraints on Presupposition accommodation. In Chungmin Lee and Beom-mo Kang (eds.) *Language, Information and Computation*. Seoul: Thaehaksa, 1993.
[15] Fagan, Sarah The English Middle. *Linguistic Inquiry* 19.2. 181-203, 1988.
[16] Lee, Chungmin Conditional Constructions in Korean. *Proceed's of the 13th Int'l Congress of Linguistics*, Tokyo, 1983.
(In)definites, Case Markers, Classifiers and Quantifiers in Korean. *Harvard Studies in Korean Linguistics* III, 1989.Dept. of Linguistics, Harvard University.
[17] Gundel, J., N.Hedberg and R. Zacharsky. Cognitive Status and the Form of Referring Expressions in Discourse. *Language* 69. 274-309, 1993.
[19] Takami, Ken-ichi Abstract of "Characterizing information in Linguistic Phenomena." In *4th Int'l Pragmatics Conference: Abstracts*. Kobe, Japan, 1993.
[20] Krifka, M. et al An Outline of Genericity. ms. University of Tuebingen, 1987.
Genericity: An Introduction. ms. University of Tuebingen, 1990.

[0] Barwise, J. and R. Cooper Simple Situation Theory and its Graphical Representation. ms. 1991.
[0] Devlin, K. A Mathematical Analysis of Definite descriptions. ms. St. Mary's College, 1991.
[0] ter Meulen, Alice Semantic Constraints on Type-shifting Anaphora. In G. Carlson & J. Pelletier(eds.) *The Generic Book*. U. of Chicago Press, 1994.
(13) M. Krifka [20]
(22) J. Auger *Meeting Handbook*, LSA 1993