Extended Generative Lexicon

Sumiyo Nishiguchi

School of Management, Tokyo University of Science

500 Shimokiyoku, Kuki-city, Saitama 346-8512, Japan

nishiguchi@rs.tus.ac.jp

Abstract

This paper proposes an elaboration of the Generative Lexicon (GL) in Pustejovsky (1995) based on a survey of BC-CWJ (2009). I manually classified the Japanese NP_1-*no* NP_2 "NP_1's NP_2" construction in accordance with semantic relations between the two nominals. The result indicates the need for the expansion of GL for computing the meaning of the NP_1-*no* NP_2 construction by incorporating *referential module*, as I call, that predicates temporary location, time, and manner of the referent. For example, in *ima-no nihon* "the present Japan," *ima-no* modifies the time of the event argument in the referential module.

1 Generative Lexicon Theory

Generative Lexicon (GL) is a theory proposed in Pustejovsky (1995). GL reduces lexical ambiguity and avoids multiple lexical entries by allowing semantic type-shifts based on the detailed lexical information. For example, instead of considering *book* as lexically ambiguous, the Qualia Structure enables semantic type-shifting of *book*; this provides means for solving a type-mismatch between *finish* and *a book* in (1b).

(1) a. Sue finished reading a book.

b. Sue finished a book.

Most likely, the meanings of (1a) and (1b) are alike; (1a) expresses the action more explicitly than (1b), that is, Sue finished reading a book according to a highly probable reading. The correct interpretation of (1b) that Sue finished reading a book, rather than swallowing a book or something else, is obtained by means of the lexical knowledge that books are made to be read—the purpose or the telic role of the book is to have its readers. The reading activity contains an event argument inside and the agent of the event argument is realized as the sentential subject *Sue*.

Such "purpose" or TELIC role is encoded in the lexical knowledge in GL (Pustejovsky 1995). According to Pustejovsky who based his theory on Moravcsik (1975), the following four qualia that originate from Aristotle's concept of matters represent four inherent properties of the referent.

(2) **CONSTITUTIVE** part-whole relation, material, weight

FORMAL orientation, magnitude, shape, dimensionality, color, position, ontological category

TELIC purpose, function

AGENTIVE origin, creator, artifact, natural kind, causal chain

2 Problems with Deriving Possessive Relations

In formal semantics, Pustevjosky's qualia structure has been applied for deriving possessive relations by means of the type-shifting mechanism. Vikner and Jensen (2002) type-shift the possessor noun using one of the qualia roles to explain the meaning of the genitive phrases following Partee (1997).

Possessive relations are ambiguous in both English and Japanese. For example, there is more than one interpretation for *Tanaka-no hon* "Tanaka's book." *Tanaka's book* may refer to the book that *Tanaka* owns or the book that *Tanaka* wrote (Barker 1995, 87).

In view of such ambiguity, Langacker (1993) considers ownership to be the prototypical meaning of the possessive construction and other relations to be the instantiations. Partee (1997) assumes two syntactic types for *John's* depending on

whether or not the following noun is inherently relational.

According to Partee, if the following noun is a non-relational common noun (CN) such as *car*, *John's* composes with *car* which is a regular (e, t) type predicate, namely, a function from individuals to truth-values (Montague 1973), and the relation between *John* and *car* is contextually supplied as shown in (3a).

On the contrary, when *John* is followed by inherently relational nouns such as *brother*, *employee*, and *enemy*, which are $(e, (e, t))$ type with an extra argument slot (a function from individuals to another function from individuals to truth-values), the relation between *John* and his brother in *John's brother* inherits kinship from the two-place predicate *brother* in (3b). (4) exemplifies the computation related to another relational noun, *friend*.

(3) a. Free R type:

　　Syntax: [John's]$_{NP/CN}$

　　Semantics: $\lambda Q \lambda P[NP'(\lambda z\ [\exists x[\forall y[[Q(y) \wedge R(y)(z)] \leftrightarrow y = x] \wedge P(x)]])]$

　　b. Inherent relation type: inherited from relational nouns:

　　Syntax: [John's]$_{NP/TCN}$ (TCN: transitive common noun)

　　Semantics: $\lambda R \lambda P[NP'(\lambda z\ [\exists x[\forall y[R(z)(y) \leftrightarrow y = x] \wedge P(x)]])]$

(4) Syntax: [[John's]$_{NP/TCN}$[friend]$_{TCN}$]$_{NP}$

　　Semantics: $\lambda R \lambda P[\text{John'}(\lambda z.\exists x[\forall y[R(z)(y) \leftrightarrow y = x] \wedge P(x)]](\text{friend-of'})$

　　$= \lambda P[\text{John's}(\lambda z.\exists x[\forall y[\text{friend-of'}(z)(y) \leftrightarrow y = x] \wedge P(x)]]$

If we apply Partee's theory to Japanese examples, most of the possessive relations are unpredictable, and there is no way to disambiguate the contextually supplied relation R.

Vikner and Jensen (2002) apply the qualia structure of the possessee noun and type-shift the possessee noun into a relational noun. For example, *John's poem*, that is, a possessive + CN, can be interpreted as the poem that John composed because the internal semantic structure of *poem* contains an *author-of* relation, which is the agentive role. According to Vikner and Jensen (2002), the meaning-shifting operator Q_A raises a one-place

holder *poem* in (5a) into a two-place holder as in (5b). The type-shifted *poem* can now combine with the possessive NP, which has a uniform type $((e, (e, t)), ((e, t), t))$, so that the authorship relation is inherited from NP *poem*, and R is no longer a free variable.

(5) a. $[\![poem]\!] = \lambda x.[\text{poem'}(x)]$

　　b. $Q_A(\text{poem}) = \lambda x \lambda y[\text{poem'}(x) \wedge \text{compose'}(x)(y)]$

Similarly, *the girl's teacher* can be explained by their mechanism. The purpose of teachers is to teach; therefore, the TELIC role of teachers is to teach someone. Now, the telic quale in the qualia structure of *teacher* raises the semantic type of a common noun *teacher* into the one of a relational noun as given in (6). *Teacher* is always someone's teacher so that *teacher* is a function from individuals to another function from individuals to truth-values.

(6) a. $[\![teacher]\!] = \lambda x.\text{teacher'}(x)$

　　b. $Q_T(\text{teacher}) = \lambda x \lambda y[\text{teacher'}(x) \wedge \text{teach'}(y)(x)]$

Such a mechanism has dramatically reduced the ambiguity of possessive relations.

3　Limit to GL

Table 1 manually classifies the 3030 examples containing the NP_1-*no* NP_2 "NP_1-GEN NP_2" construction in Japanese, such as *Fuji-no rendora* "a soap opera by Fuji TV," according to the semantic relations between the two noun phrases. The examples were sorted out of the core data of the *Yahoo! Chiebukuro* portion of BCCWJ (2009) by using ChaKi.NET 1.2β.

The survey indicates that the qualia structure plays an important role in disambiguating the meaning of the genitive marker *no* in Japanese. 29% of all instances are examples that NP_1 *selectively binds*, or modifies the qualia structure of the lexical meaning of the NP_2. For example, *Fuji-no rendora* "a soap opera by Fuji TV" is a soap opera *created* by Fuji TV, i.e., the agentive relation between the Fuji TV and a soap opera substitutes the relation between the two. In *windows-no CM* "TV commercial for the Windows," the CM is for the Windows; therefore, the meaning of *no* inherits the telic role of CM. In *Gandamu-no kao* "the

Table 1: Distribution of Semantic Patterns of NP_1-*no* NP_2 Construction

selective binding of qualia in NP_2	886	0.292409241
NP_2 is a relational noun	777	0.256435644
NP_2 is a deverbal noun	445	0.146864686
NP_1 is adjectival property	395	0.130363036
referential module modification of NP_2	244	0.080528053
NP_1 is a quantifier	152	0.050165017
possession	45	0.014851485
demonstratives	32	0.010561056
NP_1 is a deverbal noun	24	0.007590759
NP_1 is theme of deadjectival NP_2	23	0.007306226
adverb	6	0.001980198
selective binding of qualia in NP_1	1	0.000330033
total	3030	1

face of Gundam," the face is part of the Gundam robot (constitutive quale). *Shikaku* in *shikaku-no katachi* "square shape" describes the shape (formal role modification).

(7) a. Fuji-no   rendora
       Fuji TV-GEN soap
       "the soap opera by Fuji TV"

   b. $[\![Fuji - no\_rendora]\!] = \lambda$e,x[soap(x) & AGENTIVE = [make_act(e) & agent(e) = FujiTV & theme(e) = x]]

Crucially, the survey demonstrated that the GL needs to be expanded to include not only inherent properties but also referential descriptions, because 8% of the data involved the modification of the temporary elements, such as location, time, and manner of the referent of *NP₂* (e.g., *Operaza-no Kaijin* "Phantom of the Opera", that is, Phantom *in* the Opera) (Nishiguchi 2012) . As the relation between the Phantom and the Opera does not involve any of the inherent qualia structure—Phantom of the Opera was not born in the Opera (agentive), the Phantom is not made for the Opera (telic), the Phantom is not any part of the Opera (constitutive), or does not form any shape of the Opera (formal), none of the relations among the qualia structure Pustejovsky (1995) cannot substitute for the relation between the two.

4 Extended GL

Even though Pustejovsky's four qualia express inherent properties of referents, I propose supplementing lexical semantics with information about the referents. Besides type, argument, event, and

qualia structures in GL (cf. Johnston and Busa 1996, 79), the referential module (REF) has subcategories of TIME, LOC, and MANNER roles.

(8) Original GL Template

$$
\begin{bmatrix}
\alpha \\
\text{TYPESTR} = \begin{bmatrix} \text{ARG1} = \text{THE TYPE OF } \alpha \end{bmatrix} \\
\text{ARGSTR} = \begin{bmatrix} \text{D-ARG1} = \text{OTHER ARGUMENTS IN THE QUALIA} \end{bmatrix} \\
\text{EVENTSTR} = \begin{bmatrix} \text{E1} = \text{EVENTS IN THE QUALIA} \end{bmatrix} \\
\text{QUALIA} = \begin{bmatrix} \text{FORMAL} = \text{ISA-RELATION} \\ \text{CONST} = \text{PARTS OF } \alpha \\ \text{TELIC} = \text{PURPOSE OF } \alpha \\ \text{AGENT} = \text{HOW } \alpha \text{ IS BROUGHT ABOUT} \end{bmatrix}
\end{bmatrix}
$$

(Johnston and Busa 1996, 79)

(9) Template for Extended GL

$$
\begin{bmatrix}
\alpha \\
\text{TYPESTR} = \begin{bmatrix} \text{ARG1} = \text{THE TYPE OF } \alpha \end{bmatrix} \\
\text{ARGSTR} = \begin{bmatrix} \text{D-ARG1} = \text{OTHER ARGUMENTS IN THE QUALIA} \end{bmatrix} \\
\text{EVENTSTR} = \begin{bmatrix} \text{E1} = \text{EVENTS IN THE QUALIA} \end{bmatrix} \\
\text{QUALIA} = \begin{bmatrix} \text{FORMAL} = \text{ISA-RELATION} \\ \text{CONST} = \text{PARTS OF } \alpha \\ \text{TELIC} = \text{PURPOSE OF } \alpha \\ \text{AGENT} = \text{HOW } \alpha \text{ IS BROUGHT ABOUT} \end{bmatrix} \\
\text{REF} = \begin{bmatrix} \text{LOC} = \text{IN}(e2, x, l) \\ \text{TIME} = \text{AT}(e2, x, t) \\ \text{MANNER} = \text{WITH}(e2, x, y) \end{bmatrix}
\end{bmatrix}
$$

For example, *Operaza-no* "of the Opera" in *operaza-no kaijin* "the Phantom of the Opera" in (10a) and *mayonaka-no* "midnight" in *mayonaka-no kaigan* "the midnight beach/the beach in midnight" in (11a) modify referential modules of the Phantom and the beach. In *baiku-no karera* "those

on scooters" in (12a), scooter-riding is one of the temporary properties of the referents, so that it is MANNER role modification.

As a result, selective binding not only applies to qualia structure but also to a referential module, which enables the computation of the meaning of the *NP₁-no NP₂* construction. For example, *Operaza-no* "of the Opera" specifies the location of the Phantom as the Opera, *mayonaka-no* "midnight" modifies time and *baiku-no* "on scooters" fills the manner role as shown in (10b), (11b) and (12b).

(10) a. Operaza-no   kaijin
The Opera-GEN phantom
"The Phantom of the Opera

     b. ⟦*The_Phantom_of_the_Opera*⟧ = λx[phantom(x) ∧ [REF = ∃e[be-phantom(e) & theme(e) = x ∧ location(e) = The Opera]]]

(11) a. Mayonaka-no   kaigan-e   it-te
midnight-GEN   beach-GOAL   go-and
sakende-kudasai.
shout-IMP.HON
"Go to beach during midnight and shout there."

                 (BCCWJ 2011, oc 104343)

     b. ⟦*midnight_beach*⟧ = λx[beach(x) & [REF = ∃e[be-beach(e) ∧ theme(e) = x ∧ time(e) = midnight]]]

(12) a. Baiku-no   karera-mo
scooter-GEN   they-also
kekkona   ritsu-de   te-o
high   freuency-by   hand-ACC
agete-kure-ta
raise-BENEF-PAST
"Those on scooters also raised their hands often."

                 (BCCWJ 2011, oc 56711)

     b. ⟦*those_on_scooters*⟧^{g} = λx[g(1) = x & [REF = ∃e[born(e) & manner(e) = with-scooter]]]

(13) a. kinjo-no
neighborhood-GEN
seikeigeka-ni-wa   iki-mashi-ta
orthopidics-DAT-TOP go-HON-PAST

"I visited the orthopedics in neighborhood."

                 (BCCWJ 2011, oc 97196)

     b. ⟦*neighborhood* − GEN_*orthopedics*⟧ = λx[orthopedics(x) ∧ [REF = ∃e[location(e) = neighborhood ∧ theme(e) = x]]]

Kinjo-no "in the neighborhood" in (13a) and *mayonaka-no* "midnight" in (11a) represent the temporary location and time of the referents of *seikeigeka* "orthopedic clinic" and *kaigan* "beach."

Therefore, I propose the addition of a referential module to the lexical meaning in GL, for incorporating temporary location, time, manner and others of referents, in addition to the qualia structure. The possessive or genitive phrases *NP₁-no* in these examples modify the referential modules of *NP₂* which cannot be captured within the framework of the already existing GL.

5   EGL Database

I have made a small database of fifty lexical items taken from BCCWJ (2009) in the format of the Extended GL.

6   Conclusion

A quantitative survey of the meaning of the *NP₁-no NP₂* construction in Japanese revealed the need for the expansion of the GL for the computation of the meaning, although many examples were of the qualia structure modification in GL.

References

Barker, C. (1995). *Possessive Descriptions*. Stanford: CSLI Publications.

BCCWJ (2009). *Balanced Corpus of Contemporary Written Japanese, BCCWJ2009 edition*. The National Institute of Japanese Language.

BCCWJ (2011). *Balanced Corpus of Contemporary Written Japanese*. The National Institute of Japanese Language.

Johnston, M. and F. Busa (1996). Qualia structure and the compositional interpretation of compounds. In *Proceedings of the ACL SIGLEX Workshop on Breadth and Depth of Semantic Lexicons*, pp. 77–88. Kluwer.

$$
\begin{bmatrix}
\text{ASA-NO KOEN "PARK"} \\[4pt]
\text{TYPESTR} = \begin{bmatrix} \text{ARG1} = \boxed{y}\ \text{OUTDOOR'S\_LOCATION} \end{bmatrix} \\[4pt]
\text{ARGSTR} = \begin{bmatrix}
\text{D-ARG1} = \boxed{w}\ \text{HUMAN} \\
\text{D-ARG2} = \boxed{z}\ \text{HUMAN} \\
\text{D-ARG3} = \boxed{l}\ \text{LOCATION} \\
\text{D-ARG4} = \boxed{t}\ \text{TIME} \\
\text{D-E1} = \boxed{e1}\ \text{TRANSITION} \\
\text{D-E2} = \boxed{e2}\ \text{STATE} \\
\text{D-E3} = \boxed{e3}\ \text{PROCESS}
\end{bmatrix} \\[4pt]
\text{QUALIA} = \begin{bmatrix}
\text{FORMAL} = \boxed{x} \\
\text{CONST} = \{ \text{LAWN, BENCH, FOUNTAIN,...} \} \\
\text{TELIC} = \text{RECREATIONAL\_ACTIVITY}\left(\boxed{e3}, \boxed{w}, \boxed{y} \right) \\
\text{AGENTIVE} = \text{MAKE\_ACT}\left(\boxed{e1}, \boxed{z}, \boxed{y} \right)
\end{bmatrix} \\[4pt]
\text{EXT} = \begin{bmatrix}
\text{LOCATION} = \text{IN}\left(\boxed{e2}, \boxed{y}, \boxed{l} \right) \\
\text{TIME} = \text{AT}\left(\boxed{e2}, \boxed{x}, \begin{bmatrix}
\text{ASA "MORNING"} \\
\text{TYPESTR} = \begin{bmatrix} \text{ARG1} = \boxed{x}\ \text{TIME} \end{bmatrix} \\
\text{ARGSTR} = \begin{bmatrix} \text{D-ARG1} = \boxed{t}\ \text{DATE} \\ \text{D-ARG2} = \boxed{t1}\ \text{TIME} \end{bmatrix} \\
\text{QUALIA} = \begin{bmatrix} \text{FORMAL} = \boxed{x} \\ \text{CONST} = \text{PART-OF}\left(\boxed{t}, \boxed{x} \right) \wedge \text{PART-OF}\left(\boxed{t1}, \boxed{x} \right) \end{bmatrix}
\end{bmatrix} \right)
\end{bmatrix}
\end{bmatrix}
$$

Langacker, R. W. (1993). Reference-point constructions. *Cognitive Linguistics 4:1*, 1–38.

Montague, R. (1973). The proper treatment of quantification in ordinary english. In K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes (Eds.), *Approaches to Natural Language : Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, pp. 221–242. Dordrecht: Reidel.

Moravcsik, J. M. (1975). Aitia as generative factor in aristotle's philosophy. *Dialogue 14*, 622–636.

Nishiguchi, S. (2012). *Disambiguation of Possessives: The Extended Generative Lexicon.* Saarbrücken: Lambert Academic Publishing.

Partee, B. H. (1983, 1997). Genitives: A case study. In J. van Benthem and A. ter Meulen (Eds.), *Handbook of Logic and Language*, pp. 464–470. Amsterdam: Elsevier.

Pustejovsky, J. (1995). *The Generative Lexicon.* Cambridge: MIT Press.

Vikner, C. and P. A. Jensen (2002). A semantic analysis of the english genitive. interaction of lexical and formal semantics. *Studia Linguistica 56*, 191–226.