PROCEEDINGS OF THE THIRTY SIXTH ANNUAL MEETING OF THE
BERKELEY LINGUISTICS SOCIETY

February 6-7, 2010

General Session

Special Session
Language Isolates and Orphans

Parasession
Writing Systems and Orthography

Editors

Nicholas Rolle
Jeremy Steffman
John Sylak-Glassman

Berkeley Linguistics Society
Berkeley, CA, USA

Berkeley Linguistics Society
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA

Contents

iii

Acknowledgments

The editors of the 36th Annual Meeting of the Berkeley Linguistics Society are grateful to conference participants, our volunteers, session chairs, and the faculty, all of whom made the event an intellectually stimulating and enriching event. We would like to extend our sincere gratitude to the contributors of this volume for their professionalism, responsiveness, attention to detail, and patience in the editorial process, without which this would not have been possible.

Within our department, special thanks go to Paula Floro and Belén Flores for all their support with BLS 36 and this annual conference in general. The editors wish to also thank the executive committee of BLS 36 who organized and ran the conference, and Zachary O'Hagan for expertise in the final compilation using LaTeX.

Finally, we would like to thank the following organizations for their generous financial support:

Department of Linguistics
Graduate Assembly
Social Sciences Division
Student Opportunity Fund
International Computer Science Institute (ICSI)

Foreword

This monograph contains 34 of the 51 talks given at the 36th Annual Meeting of the Berkeley Linguistics Society (BLS 36), held in Berkeley, California, February 6-7, 2010. The conference included a General Session, one Special Session entitled *Language Isolates and Orphans*, and one Parasession entitled *Writing Systems and Orthography*. It was planned and run by the second-year graduate students in the Department of Linguistics at the University of California, Berkeley. The members of this executive committee were Jessica Cleary-Kemp, Clara Cohen, Stephanie Farmer, Melinda Fricke, Laura Kassner, and John Sylak-Glassman.

The papers contained herein were edited principally for style by the three editors Nicholas Rolle, Jeremy Steffman, and John Sylak-Glassman, and then given back to contributors to make changes. Nicholas Rolle took upon primary editorial responsibilities, Jeremy Steffman was an undergraduate editorial assistant, and John Sylak-Glassman helped to edit papers. Upon the final resubmission, the final versions of these papers were incorporated by Zachary O'Hagan and Nicholas Rolle into the monograph found here. Our goal has been the speedy publication of these proceedings, and as such, certain aspects – e.g., the complete unification of formatting – have been sacrificed. It is our belief that this does not detract from the final publication in any way.

Nicholas Rolle
Jeremy Steffman
John Sylak-Glassman

January 2016

A Quantitative Analysis of Nominative/Genitive Alternation in Japanese

SATOSHI NAMBU
University of Pennsylvania

Introduction

Since Harada (1971) first noted and analyzed it using a generative grammar framework, one of the most prominent case alternations in Japanese is Nominative/Genitive Alternation (henceforth NGA) which is known as Ga/No Conversion. In this case of alternation, the nominative particle *ga* is variably substituted for the genitive particle *no* in certain embedded clauses (1), but not in main clauses (2).

  (1) Ken-**ga/no**　　kayotta gakkoo
       Ken-Nom/Gen went　　school
       'the school where Ken went'

  (2) Ken-**ga/*no**　　hon-o　　kaita.
       Ken-Nom/Gen book-Acc write
       'Ken wrote the book.'

   Based on his grammaticality judgment test, Harada (1971) claimed that there is an ongoing change, whereby the speakers of Tokyo Japanese increasingly prefer *ga* to *no* in relevant environments. This hypothesis has been left untouched until Nambu (2007) and Nambu and Matsuda (2007) analyzed NGA quantitatively. The studies found that there exists an ongoing change in the variation, as described in section 2. They used a corpus 'the Minutes of the Japanese Diet' (hereafter MJD) (Matsuda 2004, 2008), which provides access to decades of speech data.
   In this paper, I present a quantitative analysis using the Corpus of Spontaneous Japanese (CSJ). The goal of this study is not only to validate the findings in Nambu (2007) and Nambu and Matsuda (2007), but also to give another perspective on the variation. As pointed out in Nambu (2007), using only the MJD corpus faces the caveat of a stylistic effect on NGA. Nakagawa (1987) indicates the possibility that stylistic difference might affect NGA. Comparing NGA in the two corpora, the

MJD and the CSJ, gives insight into the stylistic effect on NGA. In addition, the CSJ by itself can take into account the stylistic variation, investigating different speech types in the corpus. It provides a better understanding of the variation with respect to style effect.

Another aspect of this study is to identify the language external/internal factors that affect NGA. Previous studies on NGA claim several factors, but they do not take into consideration the correlations between the factors, which mask the independent effects of each factor. Taking these problems into consideration, I adopt logistic regression to identify the factors from a statistical point of view.

Section 1 describes the CSJ corpus and the data for the analysis. To see the ongoing change, section 2 presents the observation of uses of the variation over the years. Section 3 includes a discussion of stylistic effect on NGA. Section 4 provides the procedure of logistic regression analysis, including a brief explanation of language external/internal factors for the analysis. Section 5 presents the results and discussion.

1   Data

1.1   Corpus

I obtained data from the CSJ corpus. The entire CSJ contains 661 hours of spontaneous speech collected from 1999 to 2003, which corresponds to about 7 million words (Maekawa 2003). It consists of spontaneous monologues (95% of the corpus) and dialogues and readings (5%). The speech of spontaneous monologues consists of academic presentation speech, simulated public speaking, and other speech. The CSJ has information about speakers' birth years, although at 5-year intervals, and their hometowns.

1.2   Subjects and speech data

To verify Harada's language change hypothesis, Nambu (2007) and Nambu and Matsuda (2007) extracted uses of the variants *ga* and *no* from speech by all Diet members, who are native speakers of Tokyo Japanese, from the MJD corpus (76 speakers, 100 tokens from each speaker, 7,600 tokens in total). From the CSJ corpus, I extracted data from speech by all native speakers of Tokyo Japanese available in the corpus (80 speakers). I obtained all tokens (4,945) of the variants available from the speakers, which includes 405 tokens (8.2%) of the genitive *no*.

1.3   Envelope of variation

I counted the frequencies of the nominative *ga* and the genitive *no* in relevant environments. Following Nambu (2007) and Nambu and Matsuda (2007), I delineate

the environments where the variation is possible. The canonical environment for NGA, adnominal clauses, is shown below.

(3) Adnominal clauses
   [Ken-wa musuko-ga/no yonda] hon-o   katazuketa.
   Ken-Top son-Nom/Gen read   book-Acc cleaned.up
   'Ken cleaned up the books which his son read.'

There has been an argument that NGA can be observed in clauses other than adnominal clauses (Hiraiwa 2001; Miyagawa to appear; Yoshimura and Nishina 2008). I included additional environments that have been raised in the literature to observe uses of the variation. One environment is subordinate clauses headed by *made* 'until' and *yori* 'than.' Watanabe (1996), Hiraiwa (2001, 2005), and Kikuta (2002) pointed out that NGA can occur in these clauses. Another environment that I included in the data is apposition clauses headed by *to-yuu/to-no*, discussed in Inoue (1976) and Ura (1993). The examples are as follows, showing uses of the nominative *ga* for variation.

(4) *Made* or *yori* subordinate clauses

   a. [Basu-ga kuru] **made** suwatte iyooka.
      bus-Nom come until   sit   be
      'Let's sit until the bus comes.'

   b. [kyaku-ga   kuru] **yori** hayaku nimotsu-ga   tsuita.
      customers-Nom come than early   luggage-Nom arrived
      'The luggage arrived before the customer came.'
      Kikuta (2002)

(5) *To-yuu* and *to-no* apposition clauses
   [karera-ga buzi-datta] **to yuu/to no** sirase
   they-Nom safe-were   COMP   news
   'the news that they were safe'
   Inoue (1976)

2   Ongoing Change in NGA

To see whether there exists a change in progress, counting frequencies of the variants in speech over the years is the most direct way. Figure 1 from Nambu and Matsuda (2007) presents the rate of *no* from 1870 to 1970 with respect to speakers' birth years. In the graph, each speaker is represented by a point which contains 100 tokens of the variants. The gradual decline of the overall trend is fairly clear, with speakers gradually switching from *no* to *ga*. The logistic regression line

(Y=1/(1+e(-(29.239-0.0162X))))) given in the graph supports this impression statistically. Thus, Harada's insight in the early 1970s was right, and NGA is indeed involved in a change in progress.

Figure 1: Scatterplot of the rate of *no* in the MJD corpus (Nambu and Matsuda 2007)
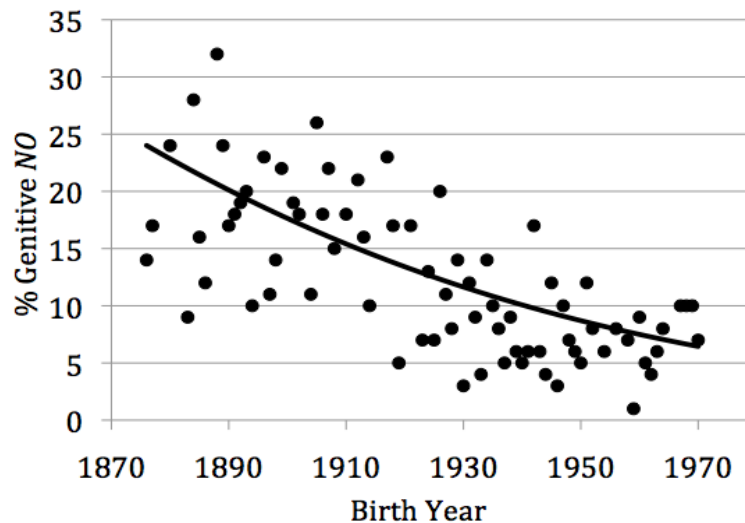
The data from the CSJ corpus is given in Figure 2. Because of the information available in the corpus, the speakers' birth years are given at 5-year intervals. The figure has a different distribution from Figure 1. This is mainly because all tokens of the variants from the 80 speakers were obtained and each dot in the figure contains different tokens from each other. Thus, it is difficult to compare the results from the two corpora from scatter plots. Introduced in section 4, a logistic regression analysis can take into account this difference in tokens per speaker in the CSJ corpus. In addition to the problem of inconsistent frequencies with respect to birth years, I need to take into consideration effects of language internal/external factors that previous studies omitted, as described in section 4.

3   Stylistic Effect

Although he does not specify the detail of the effect, Nakagawa (1987) states that there is a style difference between *ga* and *no*, and it is intuitively assumed that written language and formal speech promote the use of *no* rather than *ga*. Although a precise effect of style on NGA remains unclear, Nambu (2007) uses the MJD corpus and considers the stylistic effect with respect to the type of session and House type as a pilot study. However, the result of logistic regression analysis

Figure 2: Scatterplot of the rate of *no* in the CSJ corpus
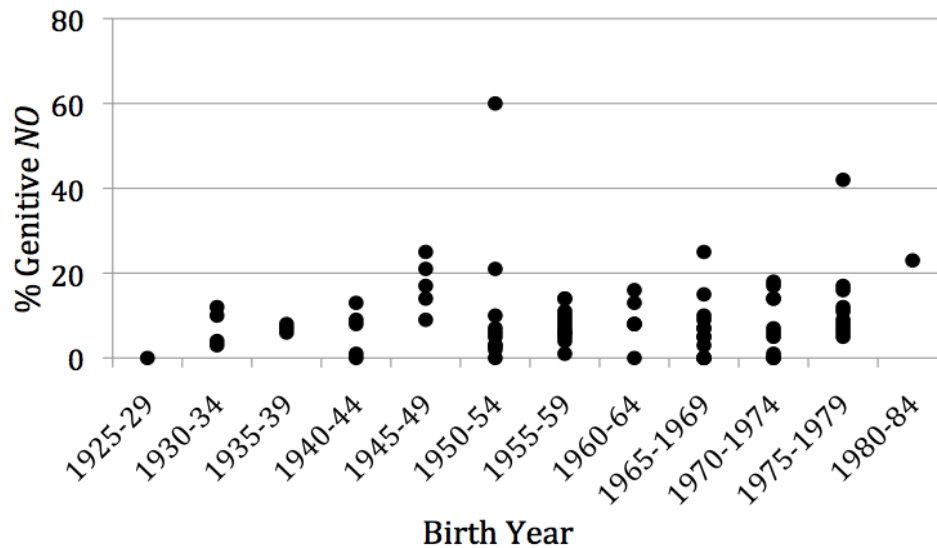
in Nambu (2007) does not show a statistically significant effect. In this paper, I compare the frequencies of *ga* and *no* between the two corpora and also in speech types within the CSJ corpus.

Table 1 gives the frequencies of *ga* and *no* in the two corpora. The Pearson's chi-square test shows that the difference between the two corpora is statistically significant. The speech in the MJD are uttered by Diet members in sessions that may have been broadcasted on TV, and thus speech in the CSJ are described in section 1.1. These token uses can be considered to be more informal. This result supports Nakagawa (1987), in addition to the intuition of native speakers of Japanese which indicate that when the speech is more formal, the use of the genitive *no* increases.

Table 1: Frequency of *ga* and *no* in MJD and CSJ

	MJD	CSJ
ga	87.7% (6,662/7,600)	91.8%(4,540/4,945)
no	12.3% (938/7,600)	8.2%(405/4,945)

X^2=54.03, d.f.=1, $p < 0.001$

Table 2 gives the frequencies of NGA for each speech type in the spontaneous monologues.[1] As described in section 1.1, the spontaneous monologues in the CSJ

---

[1] The data only contains the spontaneous monologues. I excluded the dialogues from the data,

Table 2: Speech Type in CSJ (SPS = simulated public speaking, APS = academic presentation speech, OS = other speech)

	SPS	APS	OS
ga	90.9%(3213/3,533)	92.6%(838/905)	96.4%(489/507)
no	9.1%(320/3,533)	7.4%(67/905)	3.6%(18/507)

X^2=18.8, d.f.=2, $p < 0.001$

consist of academic presentation speech (APS), simulated public speaking (SPS), and other speech as lecture (OS). The result of the chi-square test shows a statistically significant difference in Table 2. I will discuss the effect of stylistic difference within the CSJ more in detail based on the results from the logistic regression analysis in section 5.

Thus, the comparison between the MJD and the CSJ provides clear evidence that there is a stylistic effect on NGA and that the more formal speech style contains a higher frequency of the genitive *no*. In addition, different speech types in the CSJ show a different frequency of *no*, indicating an effect of style on NGA.

4   Logistic Regression Analysis

I conduct a logistic regression analysis in order to verify whether there exists an ongoing change in NGA, excluding effects of other language external/internal factors that I will introduce in this section. As Rietveld and van Hout (1993) explain, a logistic regression model consists of several independent variables and one dependent variable. In the case of NGA, the dependent variable has two values, *ga* and *no*, and the independent variables correspond to language external/internal factors that affect NGA. The logistic regression model can detect which independent variable has an effect on NGA and compute their values.

As mentioned in section 2, several factors have been proposed for NGA in the literature. In the following sections, I briefly describe the language internal/external factors. After glancing over the factors, I explicate the procedure of logistic regression analysis.

4.1   Language External Factors

Birth year and style are considered as language external factors for this analysis. Including birth year as an independent variable, logistic regression computes its

---

because they do not contain enough data for this analysis. Also, I excluded the reading speech from the data, because there are very few uses of the variation in the readings, and in addition, as for NGA, all of the readings are exactly the same as the reading material.

effect independently, and based on the result of this analysis, we can confirm the ongoing change. I converted the speakers' birth years into numerical values with 5-year intervals. As I mentioned earlier, the CSJ corpus only provides information on birth years with 5-year intervals, thus the conversion of birth years makes it possible to compare the result with the one from the MJD corpus. For style, the values for the logistic regression analysis are APS, SPS, and OS.

4.2 Language Internal Factors

4.2.1 Adjacency

Harada (1971) argued that disrupting adjacency by having intervening elements between the subject NP and the predicate affects the acceptability of NGA, as in (6). He insisted that if there exist intervening elements, the genitive *no* cannot appear as a subject marker. To consider the effect of adjacency which is about the subject NP and the predicate, I categorize NGA into two environments: adjacent and non-adjacent.

(6) kodomotati-ga/*no minna-de ikioiyoku kakenobotta kaidan
     children-Nom/Gen all-with   swiftly   run.up   stairs
     'the stairs which all children run up swiftly'

4.2.2 Transitivity Restriction

Watanabe (1996) proposes a transitivity restriction. The claim is that if a direct object exists as an argument of the predicate in the embedded clause, as in (7), the genitive *no* cannot appear in the same embedded clause as a subject marker. Comparing two environments, one with a direct object and the other without, I consider the transitivity restriction as an independent variable.

(7) a. Ken-ga   hon-o   katta   mise
        Ken-Nom book-Acc bought store
        'the store where Ken bought a book'
     b. hon-o   Ken-ga   katta   mise
        book-Acc Ken-Nom bought store
        'the store where Ken bought a book'

4.2.3 Stativity of Predicate

Horie and Kang (2000) argue that the acceptability of NGA follows a hierarchy of stativity: Verb > Existential predicate, Adjective > Copula (the frequency of *no* from highest to lowest). Their data, however, only consists of the frequency of the

genitive *no* as opposed to having both frequencies of *ga* and *no*. Thus the frequency of *no* is completely dependent on the frequency of the predicate type. Their claim is still worth checking with the data in this paper. I add adjectival verb as one of the categories for predicate and consider the effect of stativity on NGA.

4.2.4 *To-yuu/To-no* Clauses

As shown in (5) in section 1.3, Inoue (1976) observes that NGA cannot occur in *to-yuu* and *to-no* clauses. To confirm the claim, the existence of *to-yuu* and *to-no* is considered as independent variable.

4.2.5 Subject NP Type

The variable accusative case markers (*o*/zero) in the colloquial speech of Tokyo Japanese is partially determined by the object NP type (Matsuda 1995). The use of the zero-form increases in the order of the hierarchy, wh-word > lexical noun > pronoun > nominal clause. It can be inferred that NGA would also have preference for the subject NP type. I follow this categorization of NP type to consider the effect on NGA.

4.2.6 Nominal Head

The nominal head of the embedded clauses where NGA occurs can be categorized as either a grammatical noun (8) or other nouns (9).

(8) Grammatical noun *koto* as nominal head
[Taro-ga   kita] **koto**
Taro-Nom came thing
'the thing that Taro came'

(9) Other nouns as nominal head
[Taro-ga   kaita] **hon**
Taro-Nom wrote book
'the book that Taro wrote'

Masuoka and Takubo (1992) state that the meaning of grammatical nouns is bleached and their main roles shifted to a grammatical function. If the grammatical nouns do not have enough characteristics as noun, the expected result is that the use of the genitive *no* with grammatical nouns is lower than that with other nouns. This is because the adnominal clause should be a crucial environment for NGA (Hiraiwa 2001; Yoshimura and Nishina 2008). The nouns *koto*, *no*, *tokoro*, *wake*, *yoo*, and *mono* are considered as grammatical nouns in this paper, following Nambu (2007).

In addition, it should be also considered whether the clause has a nominal head or not. As shown in (4) in section 1.3, Watanabe (1996) and Kikuta (2002) claim that *made* and *yori* clauses, which do not have a nominal head, can be environments for NGA. Thus, the logistic regression analysis contains these environments as independent variables.

4.2.7 Animacy

Croft (1990) discusses the effect of animacy on grammatical phenomena. From cross-linguistic observation, he gives the extended animacy hierarchy in (10) that is for number distinction. In his discussion, more highly ranked items on the hierarchy are more likely to show number distinction.

(10) Extended Animacy Hierarchy in Croft (1990)
      first, second-person pronouns > third-person pronoun > proper names >
      human common noun > nonhuman animate common noun > inanimate
      common noun

Although there has been nothing claimed for NGA in terms of animacy, it might be the case that NP ranked higher in the hierarchy should be more likely to allow the use of the genitive *no*, assuming that the higher ranked items can occur with grammatical elements more in general.

4.2.8 *No* **Precedence Environment**

"*No* precedence environment" refers to an instance when *ga/no* follows an NP that has *no* at the end of the word, as in (11) below. If the NP is marked by the genitive *no*, the sound concatenation becomes *no-no*. To avoid the sequence of identical sounds *no-no*, the alternative *ga* may be preferred.

(11) ookii mono-ga/no   aru mise
      big   thing-Nom/Gen exist shop
      'the shop which has big things'

4.3   Procedure

Table 3 is a summary of candidates for independent variable in a logistic regression analysis, showing categories/values of each variable.

Before conducting a logistic regression, I need to exclude "knockout factors" from the analysis. As discussed in Paolillo (2002), a knockout factor is an independent variable that contains a value yielding only one fixed value for the dependent variable. Such a variable results in a problematic calculation. The transitivity restriction is a knockout factor, because there are not any examples of the genitive

Table 3: Candidates for Independent Variables

Independent Variables	Categories/values
Birth Year	numerical (5-year intervals)
Style for CSJ	SPS, APS, OS
*Transitivity Restriction	with a direct object, without a direct object
Adjacency	adjacent, non-adjacent
Predicate type	adjective, nominal adjective, verb, existential verb, *copula
Animacy	first/second-person pronouns, third-person pronoun, proper names, human common noun, nonhuman animate common noun, inanimate common noun
Subject NP Type	*wh-word, lexical noun, pronoun, nominal clause
To-yuu/to-no Clauses	with *to-yuu/*to-no*, without *to-yuu/*to-no*
Nominal head (Grammatical Noun)	*no*, *koto*, *tokoro*, *mono*, *wake*, *yoo*, others, *non-existence of nominal head (*made/yori*)
No Precedence	subject NP with *no* ending, without *no* ending

no with a direct object in the data (0/179). This supports the claim by Watanabe (1996). *No* precedence environment is also a knockout factor; there is no use of *no* when the subject NP ends with *no* (0/390). This follows the prediction that the sequence of the same sound should be avoided if there is an alternative.

Furthermore, there are no examples of *no* in clauses with a nominal head *wake* (0/209) and *tokoro* (0/10), with a copula as a predicate (0/288)[2], with a wh-phrase as a subject (0/40), in *to-no* (0/0) clauses, *made* clauses (0/14), and *yori* clauses (0/4). Therefore I do not include examples of the environments to avoid any undesirable effects on the analysis. They are marked with * in Table 3.

In addition, dependency between independent variables is excluded from the analysis. Dependency here means that one value of an independent variable is dependent on values of other variables. In the data, dependency is observed with three variables: subject NP type, animacy, and *no* precedence environment. They all have a subject NP that is in the clause with NGA. I excluded *no* precedence environment as a knockout factor. Instead of using the two separate variables, subject NP type and animacy, I combined them into one factor "SA", following Nambu (2007) to take into account dependency. For SA, the subject NP is divided into categories: inanimate common noun, human pronoun, clause, human common noun, non-human animate common noun, and proper name. Because the category proper

---

[2] The copula does not occur with the genitive *no* for an independent reason, which is "NP-*no*-NP" constraint by Shibatani (1975). The constraint mentions that the genitive *no* as a subject marker cannot occur when it can be interpreted as possessive in that position.

name does not occur with *no* in the data, I excluded it from the analysis.

In this paper, I used SPSS following Nambu (2007). Computing the regression model, the SPSS setting for logistic regression was taken as a stepwise analysis. The stepwise analysis excluded factors that do not affect NGA, as reported from the regression model.

5   Results and Discussion

Table 4 shows the results of the logistic regression analysis, giving the regression coefficient (B), the standard error for the Wald test (S.E.), Wald test for regression coefficients (Wald), the degree of freedom for the Wald test (d.f.), the statistically significant difference as probability computed by the Wald test (P), odds ratio (Exp(B)), and the confidence intervals of Exp(B). The value of Exp(B) shows an effect of each variable on NGA, excluding effects of other variables. In the case of the categorical variables, the value of Exp(B) is calculated by comparing two or more categories of each variable, one of them serving as the reference for Exp(B). To clarify the values of Exp(B), I provide the reference category for each variable. For example, the category non-adjacent environments is a reference for the variable adjacency. The value 40.6 of Exp(B) indicates that the rate of *no* in the adjacent environments is 40.6 times more than that in the non-adjacent environments. This demonstrates a strong effect of adjacency on NGA.

The rest of the variables in Table 4 are as follows. The variable *to-yuu* apposition clause compares *to-yuu* apposition clauses with clauses without *to-yuu*, having its reference as clauses without *to-yuu*. The value of Exp(B) shows that *to-yuu* has a negative effect on the use of the genitive *no*. For the variable predicate type, verb is a reference, and the value of Exp(B) shows that the frequency of *no* is adjective > existential verb > nominal adjective > verb (from highest to lowest). This result shows an effect of stativity, the more stative the predicate, the higher the use of genitive *no*. The reference for SA is non-human animate common nouns. The value of Exp(B) shows that the frequency of *no* is human pronoun > inanimate common noun > human common noun > non-human animate common noun > clause (from highest to lowest). This reflects an effect of subject NP type where the frequency of *no* is pronouns > lexical nouns > clause (from highest to lowest). Unlike Nambu (2007), animacy does not show a consistent effect on NGA in the data. The noun *mono* is a reference for the variable nominal head. The value of Exp(B) shows that the frequency of *no* with grammatical nouns is lower than non-grammatical nouns. This supports the idea that grammatical nouns are less nominal than non-grammatical nouns and that NGA is less likely to occur in clauses with grammatical nouns that are different from canonical adnominal clauses. The stylistic effect shows the frequency of *no* as OS > APS > SPS (from highest to lowest), which is different from Table 2. As described in the manual for the CSJ corpus, the APS variant is used by specialists in the academic audience and SPS is

Table 4: The output of logistic regression for the CSJ data

	B	S.E.	Wald	d.f.	P	Exp(B)	Confidence interval of Exp(B)	
							Lower	Upper
Style			6.38	2	.04			
(APS)	-.27	.17	2.37	1	.12	.77	.55	1.08
(SPS)	-.62	.29	4.69	1	.03	.54	.31	.94
Adjacency	3.70	.59	39.98	1	.000	40.60	12.88	128.00
To-yuu clause	-4.78	1.01	22.45	1	.000	.008	.001	.061
Predicate Type			100.73	3	.000			
(Adjective)	1.52	.16	91.54	1	.000	4.56	3.34	6.23
(Existential Verb)	1.20	.21	33.08	1	.000	3.31	2.20	4.97
(Nominal Adjective)	.87	.31	7.80	1	.005	2.38	1.30	4.39
Nominal Head			119.80	4	.000			
(*Yoo*)	-1.26	.43	8.68	1	.003	.28	.12	.66
(*Koto*)	-.24	.33	.52	1	.47	.79	.42	1.50
(*No*)	-4.57	.56	67.53	1	.000	.010	.004	.031
(Others)	.41	.25	2.70	1	.10	1.51	.92	2.46
SA			37.81	4	.000			
a	1.20	1.05	1.31	1	.25	3.32	.42	26.03
b	1.68	1.06	2.53	1	.11	5.37	.68	42.62
c	-.62	1.11	.31	1	.58	.54	.06	4.70
d	.94	1.06	.78	1	.38	2.56	.32	20.49
Constant	-6.42	1.22	27.82	1	.000	.002		

For a category of SA, a = inanimate common noun, b = human pronoun, c = clause, d = human common noun

used by non-specialists to a public audience; and thus APS is more formal than SPS. This is reflected in the frequency of *no*; formal speech prefers *no* to *ga*, as discussed in section 3. OS is used by specialists to a public audience, and also has the highest frequency of *no*. As OS contains a variety of speech type and has a idiosyncratic property, it is difficult to attribute the high frequency of *no* to one specific factor. The logistic model was judged as statistically good (Hosmer-Lemeshow test, χ^2 = 8.28, d.f.= 8, and p = .407). The value of Nagelkerke's R^2 is 0.495 and the probability of prediction of NGA is 91.1%.

In Table 4, the independent variable birth year is not included. This is because the stepwise method statistically judged that the variable does not affect NGA and therefore it was excluded from the regression model. This result is different from the logistic regression analysis using the MJD data in Nambu (2007) that showed a significant effect of birth year, indicating an ongoing change in NGA as seen in Figure 1 in section 2. This discrepancy between the two corpora could be attributed to an effect of individual speakers' variation which this analysis and Nambu (2007)

did not take into consideration. As explicated in Johnson (2008), the analysis considering inter-speaker variation might provide a more precise effect of language external factors such as birth year and resolve the inconsistency observed in a comparison of the CSJ with the MJD, and this is an issue worth investigating in future research.

With respect to the factors adjacency, stativity, and the transitivity restriction, the quantitative analysis shows that they have significant effects on NGA. Adjacency requires that the genitive *no* be used for a subject that is adjacent to its predicate. Stativity indicates that the genitive *no* is likely to be used with a stative predicate. The transitivity restriction restricts the occurrence of the genitive *no* to a clause without a direct object. Each factor affects NGA independently, but this does not imply that their effects derive from independent or distinct reasons. In fact, they suggest that the occurrence of NGA is restricted to stative/nominal environments rather than verbal/clausal environments. In addition to that, the effect of grammatical nouns demonstrates that NGA most often occurs in adnominal clauses. Thus, the analysis in this paper shows that the effect of each factor indicates the NGA's preference for stative/nominal environments.

6   Conclusion

I considered the effects of language internal/external factors reported in previous studies, focusing on the effect of stylistic difference on NGA. By comparing the two corpora, this study considered stylistic effects from two perspectives. The comparison between the two corpora provides evidence that more formal speech shows a high frequency of the genitive *no*. In addition, the stylistic difference in the CSJ also shows the effect of formality on NGA. Thus, this study offers evidence that there is a stylistic effect on NGA. Based on the findings from the logistic regression analysis, the effects of language internal factors indicate that NGA prefers stative/nominal environments. Additional analyses are needed to investigate whether or not language change is occurring in NGA, specifically with respect to inter-speaker variation.

References

Harada, Shin-ichi. 1971. Ga-No Conversion and Idiolectal Variations in Japanese. *Gengo Kenkyu* 60:25–38.

Hiraiwa, Ken. 2001. On Nominative-Genitive Conversion. In Elena Guerzoni and Ora Matushansky, eds., *MIT Working Papers in Linguistics*, volume 39, 66–125.

Hiraiwa, Ken. 2005. Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture. Ph.D. thesis, MIT.

Inoue, Kazuko. 1976. *Henkei Bunpoo-to Nihongo [Transformational Grammar and Japanese]*. Tokyo: Taishukan.

Johnson, Daniel Ezra. 2008. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3(1):359–383.

Kikuta, Chiharu. 2002. Ga-no kootai gensho-no hihaseitekibunseki: Jutsugorentaikei-no meishisei [Ga-No Conversation in a New Light: A Non-Derivational, Mixed Category Analysis]. *Doshisha Studies in English* 74:93–136.

Maekawa, Kikuo. 2003. Corpus of Spontaneous Japanese: Its Design and Evaluation. In *Proceedings of the ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition*.

Matsuda, Kenjiro. 2004. "Gengo Shiryo-to shiteno Kokkai Kaigiroku Kensaku Shisutemu [The On-line Full-text Database of the Minutes of the Diet: Its Potentials and Limitations]". *Theoretical and Applied Linguistics at Kobe Shoin* 7:55–82.

Matsuda, Kenjiro. 2008. *Kokkaikaigiroku-o tukatta nihongo kenkyu [Research on Japanese using the Minutes of Japanese Diet]*. Hitsuji Shobo.

Miyagawa, Shigeru. to appear. Genitive Subjects in Altaic and Specification of Phase. *Special Lingua volume edited by Jaklin Kornfilt and John Whitman* .

Nakagawa, Yoshio. 1987. Ga-No Conversion in Japanese Particles. *Bulletin of the Kyoto University of Foreign Studies* 28:309–318.

Nambu, Satoshi. 2007. Reconsideration of Ga/No Conversion Based on a Quantitative Analysis. *Gengo Kenkyu* 131:115–149.

Nambu, Satoshi and Kenjiro Matsuda. 2007. Change and Variation in Ga/No Conversion in Tokyo Japanese. In Joseph C. Salmons and Shannon Dubenion-Smith, eds., *Historical LInguistics 2005: Selected Papers from the 17th International Conference on Historical Linguistics, Madison, Wisconsin*, 119–132, John Benjamins Publishing Company.

Rietveld, Toni and Roeland van Hout. 1993. *Statistical Techniques for the Study of Language and Language Behaviour*. Berlin: Mouton de Gruyter.

Shibatani, Masayoshi. 1975. Perceptual Strategies and the Phenomena of Particles Conversion in Japanese. In *Papers from the Parasession on Functionalism*, 469–480, Chicago: Chicago Linguistics Society.

Ura, Hiroyuki. 1993. L-Relatedness and Its Parametric Variation. In Colin Phillips, ed., *MIT Working Papers in Linguistics*, volume 19, 377–399.

Watanabe, Akira. 1996. Nominative-Genitive Conversion and Agreement in Japanese: A Cross-Linguistic Perspective. *Journal of East Asian Linguistics* 5:373–410.

Yoshimura, Noriko and Akira Nishina. 2008. Ga-no kootai-to karadaimeesi-ni tuite [Ga-No Conversion and Empty Pro]. In *6th International Conference on Practical Linguistics of Japanese*.

Satoshi Nambu
University of Pennsylvania
Department of Linguistics
619 Williams Hall
Philadelphia, PA 19104-6305

satoshin@babel.ling.upenn.edu