NOMINATIVE/GENITIVE CONVERSION IN JAPANESE
AND SYNTACTIC CLAUSE SHRINKING NOW IN PROGRESS

Yoshiki Ogawa, Keiyu Niikuni, Yuichi Wada

Tohoku University

Abstract.

In Japanese, Nominative Case on a subject in an adnominal clause can alternate with Genitive Case under limited syntactic and semantic conditions, a fact called "Nominative/Genitive Conversion" (NGC). Harada (1971) showed that the range of possible environments for NGC was narrowing and identified two different idiolects distinguished only in terms of the ages of the speakers. This article reviews corpus data from Ogawa (2016) to show that this diachronic change is still in progress. Next, within the minimalist framework, we argue that this ongoing change can be explained under the hypothesis that the syntactic size of a Genitive subject clause has been shrinking over the last 100 years, probably due to the decline in the frequency of such Genitive subjects and the Minimal Structure Principle (MSP) (Bošković 1997) operating during language acquisition. We make a prediction about the different acceptability ratings by native speakers of different age groups for adjectival and nominal predicates with a Genitive subject that is borne out in a large-scale Internet-based survey.

Keywords: generative syntax, Nominative/Genitive Conversion, Minimal Structure Principle, diachronic change, intergenerational variations, corpus, acceptability judgment

Language: Japanese

It is widely recognized that any child can learn its native language within three or four years after birth, and the basic grammatical components are uniform among native speakers of a particular language (in this case, Japanese). Such homogeneity can be attained without sufficient instruction or corrections by parents or sufficient positive evidence (or indeed its lack) necessary for a child to choose a particular grammar among many potential choices. Generative linguistics assumes that attainment of such uniformity of language acquisition under the "poverty of

stimulus" can only be explained if we assume that a large portion of the language faculty is genetically determined and that all children do in the course of language acquisition is to fix, through their limited linguistic experience, the value of each of a finite set of parameters that are underdetermined when they are born. Since the advent of the principles and parameters theory of language in Chomsky (1981) and its modification under the minimalist program of Chomsky (1995), such a program has succeeded in explicating what the syntax of natural languages is like and how a large portion of parametric language variations (including variations among dialects of a language) can be traced back to the morphological properties of the functional categories in the language to which even an infant can be sensitive (for examples, richness of inflection, presence or absence of verb raising, and head-initial/final nature, to name only a few; cf. Kayne 1994; Travis 1984). Moreover, the discovery of an outstanding cross-linguistic generalization has often led to proposals about language-universal syntactic mechanisms (Cinque 1999, 2006; Kayne 2000, among others). The same achievement has been made with the discovery of a common tendency of language change or a descriptive generalization about diachronic change (Maling 1983, Kemenade 1987; to name only a few).

Among many such language variations, in this article we will focus on a unique syntactic phenomenon in Japanese, "Nominative/Genitive Conversion" (NGC). More specifically, we will present a set of corpus data showing how the distribution of NGC has been changing over the last 100 years. This observation is a follow-up to Harada's (1971, 1976) illuminating observations. We then propose a syntactic hypothesis to explain this diachronic change in terms of the principles and parameters theory: clause shrinking forced by the Minimal Structure Principle. We confirm the validity of the prediction made from this hypothesis and certain auxiliary assumptions by administering a large-scale Web-based survey in which 180 participants from each of three different age groups were asked to answer the acceptability of Nominative- and Genitive-subjects in several different constructions. The corpus study and Web-based survey converge to the conclusion that a diachronic parametric change is still in progress in which speakers of younger generations tend to choose a smaller Genitive subject clause among the following four possibilities, CP, TP, vP, VP/AP, while the Nominative subject clause has been stable as CP during the same period.

1. NOMINATIVE/GENITIVE CONVERSION IN JAPANESE

1.1. NOMINATIVE/GENITIVE CONVERSION IN JAPANESE SYNCHRONICALLY   The subject of a clause in a Nominative-Accusative language is usually marked for the Nominative Case across

languages. While this holds true in Japanese as well, the subject of a certain subset of adnominal finite clauses can also be realized with the Genitive Case, as in (1). This phenomenon is called Nominative/Genitive Conversion (NGC):

(1) [$_{XP}$ Taro-ga/no   katta]   hon

         Taro-NOM/GEN bought   book

         'the book (that) Taro bought'

Since Harada's (1971) pioneering work analyzing the phenomenon with generative linguistics, a large number of articles studying the synchronic nature of the NGC have been published, which have largely contributed to the improvement of the theories of phrase structure and/or movement. Above all, Miyagawa (2011) has elucidated how the subject in an adnominal clause can be marked for Genitive under the phase theory of Chomsky (2001),[1] under the assumption that a normal finite clause projects up to Complementizer Phrase (CP), while the relevant adnominal clause only projects up to Tense Phrase (TP), which is selected by the functional category D(eterminer) instead of C(omplementizer). For a Nominative Case to be licensed in TP, the TP needs to be selected by the phase head C, as in (2a), but when the phase head C is replaced by another phase head D, the Genitive subject becomes possible, as in (2b):

(2) a. [$_{CP}$ [$_{TP}$ DP-Nom [$_{T'}$ [[$_{vP}$ … V]] T]] C]   (Nominative, when C selects TP)

    b. [$_{DP}$ [$_{TP}$ DP-Gen [$_{T'}$ [[$_{vP}$ … V]] T]] NP D]   (Genitive, when D selects TP)

1.2. NGC IN JAPANESE DIACHRONICALLY   However, Harada's (1971) original observation of the NGC was that we need to identify at least two different idiolects regarding the extent to which a native speaker admits a Genitive subject clause and that the two idiolects can be best

---

[1] In Chomsky (2001), a constituent XP, the construction of which is/can be followed by lexical access, is called a *phase*. Phases define impenetrable domains to the feature- checking operations of Agree and movement due to the following condition:
(i)   Phase Impenetrability Condition:
       In phase α with head H, the domain of H is not accessible to operations outside α, but only H and its edge.
Functional categories that form a propositional domain such as C, v, and D are defined as phase heads. Moreover, phases are divided into strong and weak phases, where strong phases are subject to (i).

identified as intergenerational variations, as there was no difference in terms of geographic or sociological divisions among his informants; all of what he calls Dialect A speakers were in their forties, while the majority of what he calls Dialect B speakers were in their twenties. From this fact, he concluded that there was an ongoing change in the range of possible occurrences of a Genitive subject in adnominal clauses as the result of a change in the grammar that can occur in the course of language acquisition. He suggests that the grammar acquired by a child can be slightly different from that of his or her parents because simplification of the grammar in the following sense could take place:

(3)   It seems that the cause of such a simplification lies in the fact that a child acquires his native language through "constructing the simplest (optimal) grammar capable of generating the set of utterances, of which the utterances heard by the child are a representative sample[2]." Notice that the set of data available for the child is inevitably restricted in size and often degenerate in quality. Since the child constructs the optimal grammar that is consistent only with the original data, the grammar he constructs needs not be identical to the grammar that adults have constructed.   (Harada 1971:36)

More specifically, he identified two different idiolects of Japanese, one of which allows a Genitive subject in adnominal clauses of the types in (4a) and (4b) among others, and the other of which does not allow either ("(?*)" with *no* in (4a,b) indicates that Dialect A speakers allow a Genitive subject here, while Dialect B speakers do not).

(4) a.   titioya-ga/(?*)no   dai-ongakka   de-atta   buturigakusya
          father-Non/Gen   great musician   was physicist
          'a physicist whose father was a great musician'

    b.   me-ga/(?*)no   nakanaka   de-nai   sakura-no-ki
          sprout-Nom/Gen be-slow-to   come.out-not cherry-tree
          'a cherry tree which is slow to sprout'

---

[2] Halle (1962:344).

(4a) is a case of Genitive subject of a nominal copula sentence, and (4b) is a case of Genitive subject and a predicate with intervening adverb(s). Harada (1971:35) concluded that "Dialect A is on the edge of losing its status as the majority dialect, and the newcomer, Dialect B, is spreading among the speakers of the Tokyo dialect."

In the subsequent history of generative syntax, the fact that (4a) and (4b) are ill-formed (for native speakers of the contemporary Japanese) has been repeatedly cited. Strangely enough, however, Harada's important discovery of the ongoing language change and the question of how to define the grammar of Dialect A speakers was ignored for the subsequent 30 years or so before Satoshi Nambu, a Japanese psycholinguist, presented a corpus-based study of the diachronic change in the use of a Genitive subject in his 2007 article and several later articles. Nambu investigated a Hansard Corpus that compiled official speeches of the 100 members of the Japanese Diet who were Tokyo dialect speakers and whose years of birth ranged from 1870 to 1970, and showed that among these speakers the ratio of the occurrences of Genitive subjects in adnominal clauses to all the adnominal clauses with a subject decreased from approximately 25% to 7% in the last 100 years. He argued that the gradual decline of the overall trend is fairly clear, with speakers gradually switching from the Genitive *NO* to the Nominative *GA*. He also investigated how often a Genitive subject co-occurs with adjectival predicates, verbal predicates, nominal predicates, or an overt complementizer, and concluded that there is a significant difference among predicate types in terms of the frequency of NGC, and that among these the occurrence of a Genitive subject with an overt complementizer is extremely rare. However, he did not investigate whether there was diachronic change in the frequency of the Genitive subject coupled with the various types of syntactic constructions he selected.

Ogawa (2016) was the first to present a corpus-based observation of diachronic change among predicate types. Ogawa collected about 3500 examples of Genitive subject adnominal clauses and 3800 examples of Nominative subject ones from 28 books (including 7 fictional and 21 nonfictional) published between 1904 and 2014, showing that examples of a Genitive subject clause headed by an eventive predicate as in (1) have been increasingly less likely to occur, and that the ratio of stative predicates co-occurring with a Genitive subject has been increasing in the written texts published in the last 110 years.

More specifically, Ogawa (2016) sorted the 3800 Nominative subject adnominal clauses and the 3500 Genitive subject adnominal clauses in terms of which of the following six types of predicates a subject co-occurs with: (a) an adjective, (b) a stative verb, (c) a change-of-state verb

that implies a resultant state, (d) a verb that denotes repetition or habitual activity, (e) a verb that denotes a semelfactive event, or (f) a predicative nominal (i.e., the copulative identificational sentence). Specific examples of the six types of predicate are illustrated in (5a–f), respectively:

(5) a. adjective:

Hige-ga/no   koi dansei-wa josei-ni amari sukarenai.

beard-Nom/Gen dark male-Top female-by so   loved-by

'A man with dark beard is not so loved by women.'

   b. state verb:

Mehanadachi-ga/no sikkarisita   hito-wa   saisyo-no insyo-ga   yoi.

features-Nom/Gen clear-is   man-Top first impression-Nom good

'A handsome man gives good impression to everyone who meets him the first time.'

   c. verbs denoting habituality or repetition:

Sakura-ga/no saku   jiki-wa   minna-ga tanosimi-ni   siteiru

cherry blossom-Nom/Gen season-Top all-Nom looking.forward.to is

'Everyone is looking forward to the season in which cherry blossoms.'

   d. verbs denoting result state of a changing event:

Mado-ga/no   ware-ta   heya-wa tachiiri-ga   kinsi-sarete-iru.

window-Nom/Gen broken-was room-Top going.into-Nom prohibit-Pass-is

'You are prohibited from entering the room whose windows are broken.'

   e. verbs denoting a semelfactive event:

Nimotu-ga/no   todoita jikoku-wa yuugata-no   yoji   datta

package-Nom/Gen arrived   time-Top evening-Gen   four-o'clock was

'The time at which the package arrived was 4 o'clock in the afternoon.'

   f. nominal copula sentences (cf. (4a)):

Keiba-ga/no   syumi-dearu oji-wa   kyuujitu-ni-wa   ie-ni inai.

horse.race-Nom/Gem hobby-is   uncle-Top holiday-in   home-at   is-not

'My uncle whose hobby is betting on horse races is out on vacation.'

Ogawa (2016) shows that, although the ratio of Genitive subject adnominal clauses to all the adnominal clauses with a subject that are headed by any one of the six types of predicates has been gradually decreasing in the last 100 years, the more eventive predicates such as (5c–e) have

declined more steeply than the purely stative ones, such as (5a) and (5b), and nominal copular sentences as in (4a) and (5f) ceased to co-occur with a Genitive subject in the 1960s.

On the basis of Ogawa's (2016) observations, Niikuni, Wada, and Ogawa (2017) administered a large-scale Internet-based survey that imposed an acceptability rating task for Nominative- and Genitive subject clauses like (5a–f) on those who were born and raised in the Tokyo metropolitan area, and concluded that the diachronic language change that Harada (1971, 1976) identified some forty years ago is still in progress.

In their experiment, participants were shown six pairs of a Genitive subject clause and a Nominative subject clause for each of the six types of predicates and were asked to rate the acceptability of each item. Each participant belonged to one of the following three age groups: (i) 20–29, (ii) 40–49, and (iii) 65–74 years old. The result of the experiment was surprisingly parallel to what Ogawa (2016) identified from the corpus study. In summary, the following two results were obtained:

(i) The younger age groups are less likely to accept Nominative/Genitive Conversion.

(ii) The more stative predicates are more likely to accept Nominative/Genitive Conversion, where stativity is aligned in the following order:

(5a) > (5b) > (5c) = (5d) > (5e) > (5f)

2. THE PROPOSED SYNTACTIC HYPOTHESIS

2.1. A MODIFICATION OF MIYAGAWA (2011)   On the basis of Ogawa's (2016) corpus-based observation and the result of Niikuni, Wada, and Ogawa's (2017) experiment, we will propose a set of syntactic hypotheses that can explain the diachronic change still in progress. We will lay out the hypothesis on the basis of Miyagawa's (2011) (language-universal) proposal about Genitive Case licensing in finite clauses.

Recall that Miyagawa (2011) proposes that while a Nominative subject clause always has a CP structure as in (6a), a Genitive subject clause always has a CP-less TP structure (cf. (2a,b)). More specifically, he argues that the head of the TP is defective in the sense that it is unable to assign Nominative Case and assimilates such a TP to that in the Exceptional Case-Marking (ECM) complement clause, as in (6b):

(6) a. John believes [$_{CP}$ that (=C) [$_{TP}$ Mary/she is guilty]].

      b. John believes [$_{TP}$ Mary/her to be guilty]].

The defectiveness of an ECM complement is not only syntactic but also semantic: Ormazabal (1995) points out that a purely eventive predicate that denotes a temporally independent event that, in a finite complement, could be paired with a future modality or past tense, as in (7a), is excluded in an ECM complement, as in (7b):

(7) a. John believed [that Mary {would bring / brought} the beer].   (27)
    b.?*John believed [Mary to bring the beer].   (90)

If an assertion of a temporally independent eventive predication needs CP, and if an ECM complement cannot be larger than TP, then it is naturally explained that an ECM complement cannot host such a purely eventive predication. If we also assume that a Nominative Case-licensing requires a CP projection (Chomsky 2001; Miyagawa 2011), it is also naturally expected that the subject of an ECM complement resists Nominative Case licensing and needs to be assigned Accusative Case by the matrix verb. This is the nature of the so-called "exceptional" Case-marking.

Returning to the NGC, Miyagawa (2011) points out correctly that a Genitive subject clause also resists a purely eventive predication, so that there is reason to assume that a Genitive subject clause should be TP without CP. The only differences between the ECM complement and the NGC clause are that while the former is a non-finite clause, the latter is a finite one, and that the subject of the latter lacks a matrix verb to assign it an Accusative Case. Instead, as it is located in an adnominal clause, which is selected by D, it can be assigned a Genitive Case from D. If an adnominal clause happens to be CP, its subject is assigned a Nominative Case, while if it happens to be TP its subject is assigned a Genitive Case. In other words, the TP of the NGC clause is defective in the sense that it resists a pure eventive predication and does not license a Nominative Case, with both properties deriving from the fact that it does not project up to CP.

While we take over the essential component of Miyagawa's proposals, we will modify his assumptions for the following two reasons: (i) it is not the case that all ECM complement clauses lack a CP universally; (ii) it is not the case that all subordinate clauses whose subjects are assigned an Accusative Case by the matrix verb do not have to be as large as TP; rather, they can be small clauses lacking TP, such as vP, AspectP, VP, or AP. These two facts enable us to suppose that there are several different types of NGC clauses, both synchronically and

diachronically. The first point is illustrated by the ECM construction in Japanese, as in (8b), which is minimally different from a finite construction in (8a) in the type of Case assigned to the embedded subject, and both complement clauses have a CP projection:

(8) a. John-wa [$_{CP}$ [$_{TP}$ Tom-**ga**   kasikoi] to]   omotteiru.

     John-Top   Tom-Nom smart   Comp   think

     'John believes that Tom is smart.'

   b. John-wa [$_{CP}$ [$_{TP}$ Tom-**o**   kasikoi] to]   omotteiru.

     John-Top   Tom-Acc smart   Comp   think

     'John believes Tom to be smart.'

This means that even if there is a CP projection, a matrix verb can assign an Accusative Case to the embedded subject if the CP is "defective" in some sense (not necessarily identical to that of Miyagawa 2011). For example, we could assume that the CP in question can be a weak phase (Chomsky 2001). The same should also apply to the NGC, implying that there can in principle be a NGC construction with a "defective" CP, as in (9a,b):

(9) a. Guntai-no mesi-**no**   ikani mazui   **ka** nado-o   hanashi-at-te, ...

     army-Gen meal-Gen   how   unsavory **Q** etc.-Acc talk.over

     '(we) talked about how unsavory the meals served in the army are"

                          (Naoya Shiga (1910), *Kamisori*:66)

   b. Yo-**no**   akeru   **toiu**   koto-ga … soo   hayaku   kite-wa

     night-Gen open   **Comp**   fact-Nom so   early   come-Top

     naranai   …

     must not

     'the situation in which a day breaks must not come so early, …'

                          (Ryunosuke Akutagawa (1916), *Imogayu*:16)

   Ogawa (2016) observes that such an example of co-occurrence between a Genitive subject and an overt complementizer is sporadically found in writings published between the 1900s and 1960s, though its frequency was monotonically decreasing and its occurrence completely disappeared by the year 1970. It would not be accidental that the period in which such an

expression was disappearing overlaps with the period in which, according to Harada (1971), the transition from Dialect A to Dialect B was in progress. Hence, we propose that there are at least two types of NGC constructions that are structurally distinguished from each other, as in (10a,b):

(10)   a. [$_{DP}$ [$_{NP}$ [$_{CP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ DP-**Gen** V/A/N] v] T] C] N] D]   (Harada's Dialect A)

          (with a nominal predicate or an overt complementizer)

        b. [$_{DP}$ [$_{NP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ DP-**Gen** V/A] v] T] N] D]   (Harada's Dialect B)

          (without a nominal predicate or an overt complementizer)

The second point, or the fact that there are subordinate clauses whose subject is assigned an Accusative Case by the matrix verb but whose syntactic size is smaller than TP, are illustrated by the following English examples:

(11) a. John found [$_{AP}$ Mary/her out].   (cf. Stowell 1983)

     b. John saw [$_{AspP}$ Bill/him running]   (cf. Felser 1999)

     c. John had [$_{vP}$ Bill/him cut his hair]   (cf. Ritter and Rosen 1993)

     d. John made [$_{TP}$ Bill/him cut his hair]   (cf. Ritter and Rosen 1993)

Given this fact, there can in principle be a NGC construction without a TP, which can be vP or VP/AP. Hence, our proposal is that in addition to the two versions of NGC constructions in (10a,b), we also have two different versions as in (12a,b1,b2). As (12b1) and (12b2) are non-distinct in the sense of only containing a lexical category projection within NP, our claim is tantamount to the claim that there are at least four different types of NGC constructions, which we distinguish as Dialects A to D:

(12)   a. [$_{DP}$ [$_{NP}$ [$_{vP}$ [$_{VP}$ DP-**Gen** V] v] N] D]   (Dialect C)

        b1. [$_{DP}$ [$_{NP}$ [$_{VP}$ DP-**Gen** V] N] D]   (Dialect D1)

        b2. [$_{DP}$ [$_{NP}$ [$_{AP}$ DP-**Gen** A] N] D]   (Dialect D2)

We claim that one native speaker cannot have more than one choice among the four possibilities and that the unmarked clausal size for a native Japanese speaker has been shrinking from (10a) to (10b) to (12a) to (12b1,b2) in the last 100 years or so.

Why does a language learner who lives in a different time period acquire a different type of subordinate clause? Our answer to this question is that (i) a diachronic change regarding the kinds of predicates that can co-occur with a Genitive subject frequently enough in their linguistic community to be observed by learners and (ii) the general principle in (13) are relevant for each language learner to fix the value of the parameter for the unmarked syntactic size of an adnominal Genitive subject clause:

(13) Minimal Structure Principle (MSP) (Bošković 1997:25):

Provided that lexical requirements of relevant elements are satisfied, if two representations have the same lexical structure and serve the same function, then the representation that has fewer projections is to be chosen as the syntactic representation serving that function.

Bošković's (1997) application of the MSP to linguistic data is as follows: when a language user analyzes the syntactic structure of, for example, (14a), both (14b) and (14c) are logically possible; however, the MSP prefers (14b) to (14c) because they share the same lexical structure (i.e., the phonetically empty complementizer does not contribute to phonology) and serve the same function (i.e., the empty complementizer is void of semantic function), and the former has fewer projections than the latter:

(14)   a. I think he is guilty.
       b. I think [$_{TP}$ he is guilty].
       c. I think [$_{CP}$ C(φ) [$_{TP}$ he is guilty]].

We will reanalyze the MSP as a principle that works when a language learner determines the syntactic size of a specific construction. While a genitive subject could in principle co-occur with any one of the six types of predicates as illustrated in (5) or with an overt complementizer in (9a,b), suppose that the positive evidence available for a child is almost always an example of a Genitive subject co-occurring with a stative verb as in (15a). (15a) could in principle have any of the four structures in (15b–e), where all the three functional heads C, T, and *v* are phonetically empty:

(15) a. Taro-**no**　iru　heya

Taro-Gen is   room

'the room in which Taro is'

b. [$_{DP}$ [$_{CP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ Taro-no iru] v (φ)] T(φ)] C(φ)] NP(heya) D]

c. [$_{DP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ Taro-no iru] v (φ)] T(φ)] NP(heya) D]

d. [$_{DP}$ [$_{vP}$ [$_{VP}$ Taro-no iru] v (φ)] NP(heya) D]

e. [$_{DP}$ [$_{VP}$ Taro-no iru] NP(heya) D]

As all four structures have the same lexical structure and serve the same function and (15d) has the fewest projections, the MSP forces a language learner to choose (15e) as his or her unmarked structure, unless positive evidence forcing a larger structure is (sufficiently) included in their input.[3] The same principle also applies when (15d) is chosen over (15c), when (15c) is chosen over (15b), or when (15b) is chosen over (15a), although the specific positive evidence for each choice differs in the four cases. As for the types of positive evidence, it is possible to argue that exposure to a Genitive subject co-occurring with an overt complementizer as in (9) suffices for them to fix their parameter value as (10a), that exposure to a Genitive subject co-occurring with an eventive predicate in the past tense (and referring to a past semelfactive event; cf. Niikuni, Wada, and Ogawa 2017) suffices for them to fix their parameter value as (10b), and that exposure to a Genitive subject co-occurring with a verb with a passive voice morpheme *-rare* suffices for them to fix their parameter value as (10c), and so on. In fact, Ogawa's (2016) corpus data show that the frequencies of the four types of positive evidence have been decreasing in this order and one of them disappeared in the 1960s. It is reasonable to assume that such a diachronic change, along with the economy principle (13), has triggered the clause shrinking in the way we claim. In fact, this can be identified with an updated version of Harada's (1971) simplicity measure as stated in (3), since it also argues that a simpler grammar is preferred to a more complex one and that a micro-step language change can take place even between two adjacent generations if the positive evidence available for each generation during language acquisition is different.

---

[3] We have no definite idea about the minimum frequency of tokens of a certain construction that suffices for it to function as positive evidence for a specific parameter setting, but see Goldberg (2006) and references therein for relevant arguments.

2.2. TWO TYPES OF ADJECTIVAL PREDICATES AND THE NGC   Suppose that all we have said so far is correct. Suppose also that there are at least two different subtypes of adjectival predicates as illustrated by (16a) and (16b), respectively:

(16) a.   se-ga/no   takai   hito
           stature-Nom/Gen   high   person
           'a man who is tall/a tall man'
     b.   kono yama-ga/no   takai   koto
           this mountain-Nom/Gen   high   fact
           '(the fact) that this mountain is high'

In both (16a) and (16b) the same adjective *takai* 'high' is used; in the two sentences, however, the syntactic relation between the head noun and the adnominal clause differs: the former is a relative clause, while the latter is an appositive clause. More specifically, in the (a) sentences, the subject of the adnominal clause is combined with the adjectival predicate to form a complex predicate that is predicated of the head noun (*hito* 'man'), and there is a relation of inalienable possession between the head noun and the subject of the adnominal clause. Arguably, the subject and the adjectival head form a complex adjectival predicate. For example, the Japanese expression *se-ga takai* in (16a) corresponds to the English word *tall*, *kaoiro-ga warui* 'face color-Nom bad' corresponds to the English word *pale*, and so on. For these reasons, we will refer to the complex predicate as the "possessional adjective (PA) predicate." By contrast, in (16b), the head of the noun phrase is a formal noun *koto* 'fact', to which the adnominal clause is in apposition. The subject of the adnominal clause is the external argument of an individual-level adjectival predicate that describes a permanent property of the subject (Diesing 1992; Kratzer 1996). As such, we will refer to the adjectival predicate in (16b) as the "property-describing adjective (PDA) predicate."

When a PA predicate has a Genitive subject, we can assign to it a structure like (12b2). On the other hand, when a PDA predicate has a Genitive subject, we need to have a structure like (10b). Diesing (1992) argues convincingly that while the subject of a stage-level predicate, as in (17a), is base-generated within VP/AP, the subject of an individual-level predicate, as in (17b), is base-generated in [Spec, T], controlling an empty subject PRO. The two types of adjectival predicates are structured as in (18a,b), and this proposal, coupled with the Mapping Hypothesis

that she proposes, as in (19), can explain why the subject in (17a) is ambiguous between existential and generic readings, while the subject in (18b) can only receive a generic reading:

(17) a. <u>Firemen</u> are available.   (generic / existential)

    b. <u>Firemen</u> are altruistic.   (generic /*existential)

(18) a. [$_{TP}$ [$_{T'}$ T [$_{VP}$ DP [$_{AP}$ A (available)]]]]   (stage-level)

    b. [$_{TP}$ DP$_i$ [$_{T'}$ T [$_{VP}$ PRO$_i$ [$_{AP}$ A (altruistic)]]]]   (individual-level)

(19) Mapping Hypothesis (Diesing 1992:10; IP being innocuously replaced by TP):

    a. Material from VP is mapped into the nuclear scope.

    b. Material from TP is mapped into a restrictive clause.

Without going into the details of the issue of the syntax-semantics interface, let us simply adopt her proposals and their consequences. Then, the genitive subject in (16b) should occur in [Spec, T] controlling an empty PRO in VP/AP, because the PDA is a kind of individual-level predicate. Hence, the syntactic size of such an adnominal clause should be at least as large as TP, as in (10b).[4] On the other hand, the Genitive subject in (16a) can occur in AP, forming a complex predicate, and hence the syntactic size of such an adnominal clause can be as small as AP, as in (12b2).

    We then predict that the native speakers of Dialect B, which assigns to NGC the structure in (10b), should accept (16b), while the native speakers of Dialect C or D, for whom the structure of a NGC is smaller than TP, should reject (16b). If the transition from Dialect B to C/D is as much an intergenerational variation as the one from Dialect A to B, as Harada (1971) argues, then we also predict that the younger generation will be more severely resistant to accepting such

---

[4] Support for this view comes from Ritter and Rosen's (1993) observation that a small clause headed by an individual-level predicate can occur as the complement of the causative *make*, whose complement they argue is TP, but not the causative *have*, whose complement they argue is VP, as in (i), and that the causative *have* can take a stage-level adjectival predicate as its complement, as in (ii):
(i)   a. *John had Bill like French cooking.
       b. John made Bill like French cooking.
(ii)   a.*The mayor has more firemen altruistic during the summer.
       b. The mayor has more firemen available during the summer.

a PDA sentence with a Genitive subject. By contrast, PA sentences with a Genitive subject such as (16a) will be acceptable for all the speakers of Dialects A to D.

In fact, given the binary distinction for the adjectival predicates within the adnominal clause, Ogawa's corpus study shows that there is a stark contrast between the two types of adjectival predicates in terms of their occurrence with a Genitive subject: the frequency of a Genitive subject co-occurring with a PA, which only requires an AP structure, has the highest frequency among all the predicate types (from 22% to 69% of all the Genitive subjects during the last 110 years), whereas the frequency of a Genitive subject co-occurring with a PDA, which requires a TP structure, is the second lowest among all the predicate types (from 3.65% to 0% of all the Genitive subjects during the last 110 years). Incidentally, the frequency of a Genitive subject co-occurring with a copulative nominal predicate, as in (4), is lower than that of the PDA (from 1.42% to 0% of all the Genitive subjects during the last 110 years). We simply assume that the nominal copula sentence needs a CP structure (see Niikuni, Wada, and Ogawa 2017 for discussion). The distinction between the PA, PDA, and COP shows that the larger structure a Genitive subject clause needs, the lower the frequency with which it occurs and that those Genitive subject clauses that need a TP or CP have disappeared by the year 2000.

Given the three-way distinction among predicates, the syntactic hypothesis of diachronic clause shrinking, and Ogawa's (2016) corpus data, we predict that if we make an experiment with an acceptability rating task in the same way as Niikuni, Wada, and Ogawa (2017) did for the PA, PDA, and COP with a Genitive subject, then the younger generation will show a lower acceptability of a Genitive subject co-occurring with the PDA and COP, while there will be no such intergenerational distinction observed for a Genitive subject co-occurring with the PA. As for a Genitive subject co-occurring with the COP sentence, which is only compatible with Dialect A speakers' clause size, it will be rejected even by the subjects of the oldest age group in our experiment, whose ages are between 65 and 74, since they were in the twenties when Harada (1971) conducted his experiment some forty-six years ago and can be identified with Dialect B speakers.

3. EXPERIMENT: PA VS. PDA VS. COP AND INTERGENERATIONAL DIFFERENCES

3.1. PARTICIPANTS   Five hundred and forty native speakers of Japanese were recruited through a Web-based survey administered in the first half of 2017. Each participant belonged to one of the following three age groups: (i) 65–74, (ii) 45–54, and (iii) 25–34 years old. All the

participants met the following criteria: (i) born in the Tokyo metropolitan area (i.e., born in Tokyo, Saitama, Chiba, or Kanagawa Prefecture), (ii) raised in this area until the age of 15, (iii) now living in this area. The data were collected so that samples could be evenly divided by age group and by gender in each age group.

3.2. MATERIALS AND PROCEDURE   We created 12 PA sentences, 12 PDA sentences, and 12 COP sentences as follows:

(20) a.   PA sentence:
         Hoppeta-ga/no   akai hito-wa,   okugai-ni ita-yooda.
         cheek-Nom/Gen   red person-Top,   outdoor-at   stayed-seem
         'People/A man whose cheek is red seem(s) to have been outdoors.'
     b.   PDA sentence:
         Yuuyake-ga/no   akai koto-wa,   kangaetemiru to   fusigi dearu.
         sunset-Nom/Gen   red fact-Top,   on.reflection   if   strange is
         'On reflection, the fact that the evening glow is red is strange.'
     c.   COP sentence (cf. (5f)):
         Keiba-ga/no   syumi-dearu oji-wa   kyuujitu-ni-wa ie-ni   inai.
         horse.race-Nom/Gem hobby-is   uncle-Top holiday-in   home-at is-not
         'My uncle whose hobby is betting on horse races is out on vacation.'

The Nominative-subject part (e.g., *Hoppeta-ga akai*… 'Cheek-Nom red…') and the Genitive-subject part (e.g., *Hoppeta-no akai*… 'Cheek-Gen red…') of a sentence were always presented in pairs, and participants were asked to rate each item on a 5-point Likert scale ranging from 0 (very unnatural) to 4 (very natural). For half the participants of each age group, a Nominative-subject part of the sentence was presented above a Genitive subject one in each pair, while for the other half, a Genitive-subject part was presented above a Nominative subject one in each pair.

In addition to the 36 pairs of the target sentences above, we prepared 12 pairs of filler sentences. Each pair of these fillers contained an ungrammatical sentence. If a participant's average rating score for these 12 ungrammatical sentences exceeded 2, the participant's data were excluded from the analysis. A total of 48 pairs of sentences were presented on a page on the

web browser in individually randomized order. Participants were allowed to change their answer unless they proceeded to the next page. In the questionnaire, two pairs of dummy items were mixed into the experimental items. Participants were instructed to make the specified answer (rating "0" or "4") for these items. If a participant made a different answer from what was specified at least once for the dummy items, we excluded the participant's data from the analysis.

The experiment was approved by the human ethics committee of the Graduate School of Information Sciences, Tohoku University.

3.3. RESULTS   We finally analyzed the data from 471 participants: 152 from the 65–74 age group, 160 from the 45–54 age group, and 159 from the 25–34 age group. The mean rating scores (0 to 4: smaller values mean that the sentence is harder to accept) for the target sentences in the acceptability-rating task were entered into a three-way mixed ANOVA. The ANOVA included a between-participant factor (Age group: 65–74/45–54/25–34) and two within-participant factors (Predicate type: PA/PDA/COP; Case particle: *No/Ga*). Greenhouse-Geisser corrections were used where applicable.

Fig. 1 shows the mean rating scores by Age groups for each experimental condition. The ANOVA found significant main effects of Age group ($F(2,468) = 6.74$, $p = .001$), Predicate type ($F(1.87,877.12) = 1232.81$, $p < .001$), and Case particle ($F(1,468) = 2195.34$, $p < .001$). The ANOVA also found significant interactions between Age group and Predicate type ($F(3.75,877.12) = 22.24$, $p < .001$), between Age group and Case particle ($F(2,468) = 21.64$, $p < .001$), between Predicate type and Case particle ($F(1.95,913.41) = 1693.70$, $p < .001$), and between Age group, Predicate type, and Case particle ($F(3.90,913.41) = 22.23$, $p < .001$).

Since the higher-order Age group × Predicate type × Case particle interaction was significant, we conducted follow-up analyses to test simple-main effects and a simple interaction of Age group and Case particle for each predicate type condition.

In the PA predicate type condition, we found a significant main effect of Case particle ($F(1,468) = 81.33$, $p < .001$) and a significant interaction between Age group and Case particle ($F(2,468) = 5.66$, $p = .004$). No main effect of Age group was significant ($F(2,468) = 0.09$, $p = .911$). A follow-up analysis for the Age group × Case particle simple interaction found significant simple-simple-main effects of Case particle for each age group (65–74 group: $F(1,151) = 8.52$, $p = .004$; 45–54 group: $F(1,159) = 44.24$, $p < .001$; 25–34 group: $F(1,158) =$

34.87, $p < .001$), while that of Age group was significant in neither the Ga ($F(2,468) = 2.15$, $p = .118$) nor No ($F(2,468) = 1.02$, $p = .362$) case particle condition.

In the PDA predicate type condition, we found significant simple-main effects of Age group ($F(2,468) = 26.97$, $p < .001$) and Case particle ($F(1,468) = 1112.47$, $p < .001$) and a significant simple interaction between Age group and Case particle ($F(2,468) = 45.91$, $p < .001$). A follow-up analysis for the Age group × Case particle simple interaction found a significant simple-simple-main effect of Age group in the No case particle condition ($F(2,468) = 52.36$, $p < .001$) but not in the Ga case particle condition ($F(2,468) = 2.01$, $p = .135$). A multiple comparison (Holm's method) for the effect of Age group in the No case particle condition revealed that all the residual score differences between age groups were significant ($ps < .001$). The post analysis also found significant simple-simple-main effects of Case particle in each age group (65–74 group: $F(1,151) = 214.44$, $p < .001$; 45–54 group: $F(1,159) = 414.19$, $p < .001$; 25–34 group: $F(1,158) = 497.98$, $p < .001$).

In the COP predicate type condition, we found significant simple-main effects of Age group ($F(2,468) = 6.51$, $p = .002$) and Case particle ($F(1,468) = 3989.19$, $p < .001$) and a significant simple interaction between Age group and Case particle ($F(2,468) = 3.18$, $p = .043$). A follow-up analysis for the Age group × Case particle simple interaction found a significant simple-simple-main effect of Age group in the No case particle condition ($F(2,468) = 8.06$, $p < .001$) but not in the Ga case particle condition ($F(2,468) = 0.08$, $p = .927$). A multiple comparison for the effect of Age group in the No case particle condition revealed that the mean rating score of the 65–74 age group was significantly higher than that of the 45–54 group ($p < .001$) or the 25–34 group ($p = .007$), but there was no significant difference between the 45–54 group and the 25–34 group ($p = .36$). The post analysis also found significant simple-simple-main effects of Case particle in each age group (65–74 group: $F(1,151) = 1217.72$, $p < .001$; 45–54 group: $F(1,159) = 1684.06$, $p < .001$; 25–34 group: $F(1,158) = 1175.39$, $p < .001$).
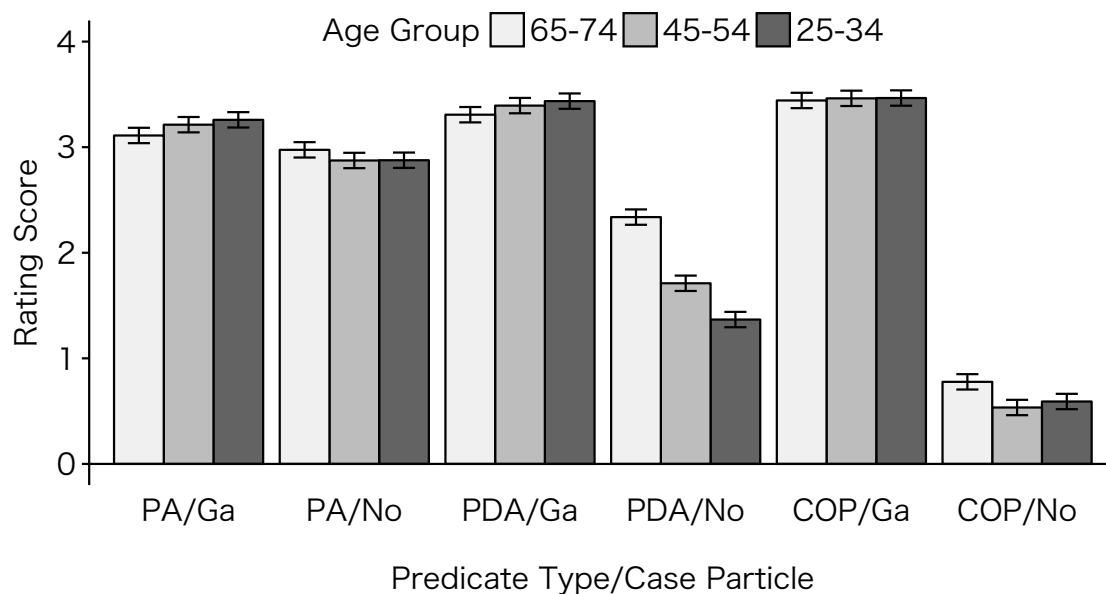
Fig. 1 Mean rating scores of the acceptability rating task by Age groups for each experimental condition. Error bars denote 95% confidence intervals. Higher scores indicate higher levels of acceptance.

In summary, the following results were obtained from the statistical analysis:

(i)   In each predicate type condition, the Genitive-subject sentences were less likely acceptable than the Nominative-subject ones for each age group.

(ii) In each predicate type condition, there was no significant difference in the acceptability of the Nominative-subject sentences among the three age groups.

(iii) There was no significant difference in acceptability of the Genitive-subject-PA sentences among the three age groups.

(iv) The younger age groups were significantly less likely to accept Genitive-subject-PDA sentences than the older age group(s).

(v)   The two younger age groups (45–54 age group and 25–34 age group) were significantly less likely to accept Genitive-subject-COP sentences than the 65–74 age group; there was no significant difference in acceptability of the Genitive-subject-COP sentences between the 45–54 age group and the 25–34 age group.

3.4. DISCUSSION   Among the five results obtained, (ii) and (iii) are exactly what we predict from the proposed hypothesis, and nothing additional needs to be said about them. As for (iv), however, given what we have said so far, we may predict that Dialect C and D speakers equally reject a Genitive-subject-PDA sentence, since neither of them has TP whose Spec could be occupied by a Genitive subject. This implies that intergenerational variation between Dialects C and D as to the acceptability rating score of a Genitive-subject-PDA sentence remains unexplained. A similar consideration also applies to (v), in which the significant difference between Dialect B and Dialect C/D speakers remains unexplained, since neither of them has a CP that could host a nominal copula sentence. Hence, we introduce an additional condition to explain the three-way intergenerational distinctions, as in (21):

(21)   The Markedness Condition on Acceptability Judgment:
        The larger the degree of deviance is from the unmarked structure that can be generated
        by the speaker, the less acceptable the generated structure is judged to be.[5]

(21) is a hypothesis about how (adult) native speakers tend to rate the acceptability of a sentence they are asked to judge with reference to their own grammar, which was fixed in their childhood: more concretely, it says that a sentence is judged less acceptable if the degree of deviance from the unmarked structure their grammar generates is larger. Recall our hypothesis, which states that the structure of an adnominal clause with a Genitive subject has been shrinking from (10a) to (10b), from (10b) to (12a), and from (12a) to (12b) in the last 110 years. Now, we refer to (10a), (10b), (12a), and (12b) as the unmarked structure for speakers of each of the Dialects A to D, respectively. Given the four different dialects and (21), we predict that each of the four different dialect speakers will judge the PDA sentence (which requires at least TP) as follows: Dialect B speakers, who have TP as the unmarked structure for NGC, need no deviance from their unmarked structures, while it is deviant from the unmarked structure of a Genitive subject clauses for Dialect C speakers, (= vP) by one projection and from the unmarked structure of a

---

[5] An anonymous reviewer asks whether we can ignore the effect of the context imagined by the test subjects in making the hypothesis in (21). However, as we have collected judgments by more than five hundred test subjects, there is good reason to suppose that any effects due to different contexts imagined by different people have been offset or reduced to a minimum. See also Schütze and Sprouse (2014) for an argument for the importance of using acceptability judgments in linguistics and an introduction of specific judgment tasks one of which we have adopted.

Genitive subject for Dialect D speakers (= VP/AP) by two projections. We can thus explain why the three different dialects' speakers will make the three-way distinct ratings for PDA sentences with a Genitive subject.

Next, let us consider why the Genitive-subject-COP sentences are rated worse for all three generations than the Genitive-subject-PDA sentences, and why there is no significant difference between the 45–54 group and the 25–34 group as to the acceptability of the Genitive-subject-COP sentences. For the first question, we can naturally attribute it to the fact that the COP sentence requires CP while the Genitive-subject-PDA sentences only need TP, and the former needs one more projection than the latter. Since the unmarked structures for speakers of each of Dialects B to D (for which we administered the present experiment) are all smaller than CP, the Genitive-subject-COP sentence is deviant from their unmarked structures by at least one projection and at most three projections. This is why the Genitive-subject-COP sentence is judged less acceptable than the Genitive-subject-PDA sentence for all three generations. As for the second question, we assume that the low acceptability of the Genitive-subject-COP sentences has passed below the lower bound or threshold of acceptability for the younger two groups, since the average rating scores for these groups are between 0 and 1. In other words, most of the subjects in the two younger generations consider the Genitive-subject-COP sentences gibberish. This is why no significant difference occurs between the two younger generations, hence the lack of difference between the two younger generation in terms of the rating of the Genitive-subject-COP sentences.

Finally, let us consider the result in (i). The Genitive-subject-PA sentences can be generated by unmarked structures for all the Dialect A to D speakers, and the Nominative-subject-PA sentences, which have a CP structure by assumption, can also be generated by them. Hence, we do not expect any difference in the rating scores of the two sentences in a minimal pair. But Ogawa (2016) shows (i) that the total frequency of a Nominative subject in an adnominal clause has gradually increased and doubled (from 1000 to 2000 per million characters) in the last 100 years, while the total frequency of a Genitive subject in an adnominal clause has monotonically decreased to one-fifth of the original value (from 2300 to 500 per million characters) in the same period, (ii) that the frequency of the former has been larger than that of the latter since the 1980s, and (iii) that the gap between the two constructions has been growing for the last 30 years (see Niikuni, Wada, and Ogawa (2017) for a graph showing the turnover in the latter half of the 20th century). It can then be conjectured that those who participated in our experiment last year,

regardless of their age, had a preference for using a Nominative subject sentence rather than a Genitive one, even if an adnominal clause could be syntactically compatible with either type of subject. This is probably the reason for the result in (i).

4. CONCLUSION   In this article, we have argued for the hypothesis that the syntactic size of a Genitive subject clause in Japanese has been shrinking from CP to VP/AP in the last 100 years or so, while that of a Nominative one has been stable as CP for the same period. As a result of the clause shrinking that has only taken place for the Genitive subject clause, it has been increasingly less likely to co-occur with various types of predicates and its use has become almost entirely limited to co-occurrence with a PA or a stative verb. We have shown that the clause shrinking that we claim has been taking place can be justified by both data collected from publications and an Internet-based survey of acceptability rating administered to more than 500 people. Although Nambu (2014) was careful in judging whether the language change that Harada (1971) identified some forty-five years ago has been in progress, it is safe to conclude from our results that the language change is still in progress. There remains the question of why such clause shrinkage occurs for a particular construction, but the descriptive generalization we obtained seems to be highly compatible with the simple hypothesis. Next we must ask how general the diachronic clause shrinking is and why such a change has continued in a language for more than 100 years. We leave these interesting issues for future research.

References:

Bošković, Željko. 1997. *Nonfinite complementation: An economy approach*. Cambridge: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

———. 1995. *The minimalist program*. Cambridge: MIT Press.

———. 2001. "Derivation by phase." In *Ken Hale: A life in language,* edited by Michael Kenstowicz, 1–52. Cambridge: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

———. 2006. *Restructuring and functional heads: The cartography of syntactic structures, Volume 4*. New York: Oxford University Press.

Diesing, Molly. 1992. *Indefinites*. Cambridge: MIT Press.

Felser, Claudia. 1999. *Verbal complement clauses: A minimalist study of direct perception constructions*. Amsterdam: John Benjamins Publishing Company.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalizations in language*. Oxford: Oxford University Press.

Halle, Morris. 1964. "Phonology in generative grammar." In *The structure of language: Readings in the philosophy of language*, edited by Jerry A. Fodor and Jerrold J. Katz, 334–352. Englewood Cliffs: Prentice-Hall. First published in *Word* 18 (1962): 54–72.

Harada, Shin-ichi. 1971. "*Ga-No* conversion and idiolectal variation in Japanese." *Gengo Kenkyu* 60: 25–38.

———. 1976. "*Ga-No* conversion revisited." *Gengo Kenkyu* 70: 23–38.

Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

———. 1994. *The antisymmetry of syntax*. Cambridge: MIT Press.

———. 2000. *Parameters and universals*. New York: Oxford University Press.

Kemenade, Ans van. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.

Kratzer, Angelica. 1996. "Stage-level and individual-level predicates." In *The generic book*, edited by Gregory N. Carlson and Francis J. Pelletier, 125–175. Chicago: University of Chicago Press.

Maling, Joan. 1983. "Transitive adjectives: A case of categorial reanalysis." In *Linguistic categories: Auxiliaries and related puzzles, Volume 1*, edited by Frank Heny and Barry Richards, 253–289. Dordrecht: D. Reidel Publishing Company.

Miyagawa, Shigeru. 2011. "Genitive subjects in Altaic and specification of phase." *Lingua* 121: 1265–1282.

Nambu, Satoshi. 2007. "Teiryo-teki bunseki-ni motoduku '*Ga/No* kotai saiko" (Reconsideration of *Ga/No* conversion based on a quantitative analysis). *Gengo Kenkyu* 131, 115–149.

———. 2014. "On the use of case particles in Japanese: Corpus and experimental study." Ph.D. diss., Osaka University.

Niikuni, Keiyu, Yuichi Wada, and Yoshiki Ogawa. 2017. "Youninsei-no sedaikan-sa-ga simesu gengo henka-no yousou: Syukaku-zokkaku koutai-no baai" (Nominative-Genitive conversion and language change still in progress: Evidence from acceptability judgments by different age groups). *Cognitive Studies* 24: 395–409.

Ogawa, Yoshiki. 2016. "Zokkaku syugo-o hukumu kankei-setsu, keisiki-meisi-setsu-no tsuujiteki zyootai-ka to goi-ka-nituite" (On diachronic stativization and lexicalization of relative clauses and clauses in apposition to formal nouns containing genitive subjects). Paper presented at the Third Workshop on Language Change and Language Variation: Facts of Language Change and Language Variation Unavailable from Introspective Judgment and Linguistic Theory, Tohoku University, Japan, September 2016.

Ormazabal, Javier. 1995. "The syntax of complementation: On the connection between syntactic structure and selection." PhD diss., University of Connecticut.

Ritter, Elizabeth, and Sara Thomas Rosen. 1993. "Deriving causation." *Natural Language and Linguistic Theory* 11: 519–555.

Schütze, Carlson T. and Jon Sprouse. 2014. "Judgment data." Research methods in linguistics, edited byRobert J. Podesva and Devyani Sharma, 27-50. Cambridge: Cambridge University Press.

Stowell, Timothy. 1983. "Subjects across categories." *The Linguistic Review* 2: 285–312.

Travis, Lisa. 1984. "Parameters and effects of word order variation." PhD diss., Massachusetts Institute of Technology.