SYNTACTIC REANALYSIS OF PRONOUNS AS DEMONSTRATIVES:
A CASE OF DEGRAMMATICALIZATION

YUKO YANAGIDA
University of Tsukuba

This paper argues that Old Japanese (8th century) had a distinction between strong and deficient pronouns, comparable to the distinction found in Romance languages as described in detail by Cardinaletti and Starke (1999). I propose that the loss of the deficient pronouns, which occurred in Early Middle Japanese, is best analyzed as loss of the functional category AgrP within the extended nominal projection. Due to the loss of AgrP, strong pronouns, in particular first person *ware*, underwent a shift from first person to second person and acquired the derogatory sense found in Late Middle Japanese. I argue that this change can be formally analyzed as a categorial reanalysis of personal pronouns as demonstratives, and that it involves what Norde (2006, 2009) defines as “degrammaticalization.”

*Keywords: degrammaticalization, strong pronoun, demonstrative, clitic, Old Japanese*

1. Introduction

Grammaticalization is accompanied by phonetic and phonological reduction and loss of syntactic independence. Kuryłowicz (1975: 52) defines grammaticalization as “the increase of the range of a morpheme advancing from a lexical to a grammatical or from a grammatical to a more grammatical formant.” This is described in terms of a “cline of grammaticalization”
by Hopper and Traugott (1993: 7) as follows:

(1) content item > grammatical word > clitic > inflectional affix

In so-called grammaticalization theory, this process has been taken to be strictly a unidirectional process; counter-directional developments are assumed to be either non-existent or statistically insignificant. Much recent research has, however, focused on what is known as degrammaticalization: changes that run counter to the general direction of (1) (cf. Janda (2001), Norde (2006, 2009)). Unlike grammaticalization, however, degrammaticalization has been treated as comprising a single change in isolation, such as from affix to clitic, or from function word to lexical item. Thus, Willis (2007) discusses degrammaticalization of Welsh *eiddo* from possessive pronoun ‘his’ to the noun ‘property.’ Norde (2006, 2009) discusses development of the word-marking genitive -s out of the case suffix in Swedish. Based on Kuryłowicz’ (1975) classical definition of grammaticalization, Norde (2006) analyzes degrammaticalization as a process that involves an increase in semantic complexity, pragmatic significance, and syntactic freedom:

(2) **Degrammaticalization**

Degrammaticalization is a change from a grammatical to a lexical formant, or from a grammatical to a less grammatical formant.

(Norde (2006: 203))

This paper presents a formal account of the development of Japanese strong pronouns into demonstratives. This change is exactly the reverse of the reanalysis of the Latin demonstrative *ille* as a strong pronoun in Romance, as proposed by Giusti (1998, 2001). I suggest that this change in the pronominal systems involves degrammaticalization defined as loss of AgrP within the extended nominal projection and that the newly degrammaticalized, referentially dependent pronoun gains semantic content not present in the original pronoun. Section 2 discusses the general characterization of three classes of pronouns in Romance languages proposed by Cardinaletti and Starke (1999). In section 3, it is shown that Old Japanese (8th century; OJ) has a tripartite pronominal system comparable to the Romance systems. Section 4 discusses diachronic changes undergone by these pronouns and section 5 provides a formal analysis of the shift of pronoun to demonstrative in the history of Japanese.¹

2. The Tripartite Taxonomy of Pronominal Systems

Romance languages possess three classes of pronouns, labeled by Cardinaletti and Starke (1999) as strong pronouns, weak pronouns and clitic pronouns. Cardinaletti and Starke provide a number of pieces of evidence that these three classes of pronouns have distinct syntactic behavior. The tripartite taxonomy of pronouns is represented in (3):

(3) Pronouns
   \[\text{Strong} \quad \text{Deficient}\]
   \[\text{Weak} \quad \text{Clitics}\]

Cardinaletti and Starke (1999: 152) argue that deficient pronouns (both weak and clitic) are differentiated from strong pronouns in that:

(4) A deficient, but not a strong personal pronoun:
   a. must occur in a special derived position (associated with a case feature).
   b. is incompatible with coordination and c-modification (i.e. modification by an adverb that modifies the entire NP).

The French examples in (5)–(8) illustrate the difference between strong and deficient pronouns in terms of position, coordination and c-modification:

**Position and Coordination**

(5) Strong pronouns
   a. Jean trouve elles belles.
      John finds them pretty
   b. Jean trouve elles et celles d’à côté belles.
      John finds them and those besides pretty

(6) Deficient pronouns
   a. Jean les trouve belles.
      John them finds pretty
   b. *Jean les et celles d’à côté trouve belles.
      John them and those besides finds pretty

**C-Modification**

(7) Strong Pronouns
   a. seulement lui
      only him
   b. *beau lui
      beautiful him
Deficient Pronouns
a. *seulement il
b. *beau il

In English, the personal pronouns *he and *she behave like strong pronouns in terms of coordination and c-modification, while *it behaves like a deficient pronoun. Cardinaletti and Starke show that this is a common pattern, where inanimate pronouns are deficient:

Coordination
a. I saw the book and the magazine.
b. I saw Bill and him.
c. *I saw the book and it.

C-modification
a. I thought about only her.
b. *beautiful her

According to Cardinaletti and Starke, the highest layer CL hosts both case and referential features. SL hosts polarity features, and IL hosts agreement features. Strong pronouns (12a) are considered to possess three structural layers. Weak pronouns (12b) lack the highest layer and thus lack case and referential features. Clitics (12c) lack the highest layer SL of weak pronouns. This has the following result (Cardinaletti and Starke (1999: 192)):
SYNTACTIC REANALYSIS OF PRONOUNS AS DEMONSTRATIVES     403

(13) a. Deficient elements are morphologically “lighter” (i.e. have less heads to realize).
    b. Deficient elements are necessarily in a case-assigning position at S structure (to recover case).
    c. Deficient elements cannot refer, they must be referentially dependent on an antecedent prominent in the discourse.

Another difference between strong and deficient pronouns is stated in (14) (Cardinaletti and Starke (1999: 162)):

(14) a. Deficient pronouns are never semantically/prosodically focused.
    b. Strong pronouns are always semantically/prosodically focused.

What follows from the property in (14) is that while strong pronouns cannot be expletive or impersonal, weak pronouns appear in subject position in expletive and impersonal constructions.² Clitics do not appear in impersonal subject position since clitics are X⁰ categories.

3. The OJ Pronominal System

3.1. A Tripartite System for OJ

Modern Japanese (ModJ) pronouns are distinctive compared to their Indo-European counterparts in that there is a large inventory of pronominal forms which originate as nouns (e.g. watakusi ‘I’ < ‘private,’ boku ‘I’ < ‘servant,’ kimi ‘you’ < ‘lord’).³ These pronouns meet none of the criteria of pronouns suggested by Cardinaletti and Starke. While neither strong nor deficient pronouns allow adjectival modification in European languages (see (7b) and (8b)), ModJ pronouns may be freely coordinated with a noun as in (15), and modified by an adjective as in (16):

² An impersonal construction is a clausal construction in which no referential subject is realized. Thus in Italian (i) on ‘they’ is a deficient pronoun used only in a non-referential context (Cardinaletti and Starke (1999: 155)):

(i) On t’ a vendu un livre pas cher.
    they you have sold a book not expensive
    ‘They have sold you a book not expensive.’

³ First person watakusi, which originally meant ‘private’ or ‘private property’ in OJ, was used as a first person designator in Middle Japanese (MJ). Thus, Amakusaban Heike, the Jesuit romanized text of Heike, published in 1592, contains 17 tokens of watakusi used as a first person pronoun, while it contains 85 tokens of ware, which is descended from the OJ strong first person pronoun. The statistical data for Amakusaban Heike is taken from Eguchi (1986).
(15) watasi wa [utukusii kimi] ga sukida
   I Top beautiful you Nom like
   ‘*I like beautiful you.’
(16) [John to watasi] ga ikimasu.
   John and I Nom go
   ‘John and I will go.’

In contrast with the ModJ facts shown above, Whitman and Yanagida (2009)
suggest that in OJ, pronouns are divided into three classes, equivalent to the
tripartite system found in Romance:\footnote{4}

(17) Tripartite Pronominal Systems in OJ (8th century)

<table>
<thead>
<tr>
<th></th>
<th>clitics</th>
<th>weak pronouns</th>
<th>strong pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person</td>
<td>(w)a</td>
<td>(w)a-re</td>
<td></td>
</tr>
<tr>
<td>2nd person</td>
<td>na</td>
<td>na-re</td>
<td></td>
</tr>
<tr>
<td>demonstrative</td>
<td>si</td>
<td>so/ko</td>
<td>so-re/ko-re</td>
</tr>
</tbody>
</table>

These three classes of pronouns differ with respect to the cooccurrence re-
striction with case/focus particles:

(18) Pronominal Systems in OJ

<table>
<thead>
<tr>
<th>Cooccurrence with the particle</th>
<th>deficient pronouns</th>
<th>strong pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>clitic pronouns</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(w)a, na, si</td>
<td>+</td>
</tr>
<tr>
<td>case</td>
<td>ga (Gen)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>wo (Obj)\footnote{5}</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>no (Gen)</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>ni (Dat/Loc)</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>yu, yori ‘from’</td>
<td>−</td>
</tr>
<tr>
<td>kakari</td>
<td>koso, zo, namu, ya, ka (Foc)</td>
<td>−</td>
</tr>
<tr>
<td>kakari</td>
<td>pa (Top)</td>
<td>+</td>
</tr>
</tbody>
</table>

(cf. Whitman and Yanagida (2009))

\footnote{4}{The current study of pronouns is based mainly on two electronic text corpora: \textit{Nihon Koten Bungaku Honbun} Data Base from the National Institute of Japanese Literature and Yoshimura’s \textit{Man’yōshū Kensaku}, as well as the editions by Nakanishi (1978–1983), Kojima et al. (1995) and Ogihara and Konosu (1979).

\footnote{5}{OJ wo, the ancestor of the ModJ accusative case \textit{o}, marks a wider range of objects than ModJ \textit{o}, including a locative PP selected by an intransitive verb. It is thus glossed as Obj (object marker).}
3.2. Strong Pronouns

The full forms with -re such as (w)a-re ‘I’ and na-re ‘you’ are strong pronouns. The strong forms appear with kakari-focus particles such as koso and zo and all types of case particles except for genitive ga and no. Examples (19) and (20) are taken from Man’yôshû (MY):\(^6\)

(19) a. are (安礼) pimo toku (MY 3361)
   we/I sash untie
   ‘We untie our sashes.’

   b. nare (奈礼) wa ga te purenana … oti kamo
   you I Gen hand touch-not fall Q
   ‘Will you fall … though my hand never touches you?’

(20) a. are wo (安礼乎) tanomete asamasi mono wo
   I Obj dependent offensive thing Excl
   ‘Although he made me dependent on him, he found me offensive.’

   b. ware yori (和礼欲利) mo madusiki pito (MY 892)
   I than even poor person
   ‘A person poorer than myself.’

   c. pikoposi mo ware ni (和礼尔) masarite (MY 3657)
   Hikoboshi even I Dat surpassing
   ‘Even the Hikoboshi surpassing me.’

Strong pronouns often appear morphologically unmarked (19), but can freely cooccur with case particles; wo in (20a), yori in (20b) and ni in (20c). The combination of strong pronoun and genitive ga does not occur in the OJ corpus, that is, *(w)a/na-re ga*. As explained in section 4, I propose that strong pronouns and ga are in complementary distribution because they appear in the same functional head. The strong pronouns are compatible with modification and coordination. As shown in (21), we find examples where ware coordinates with a common noun (21a) in parallel with (5b) and (9b), and is modified by a c-modifier makoto ‘true’ (21b), in parallel with (7a) and (10a):

---

\(^6\) Following standard practice in Japanese historical linguistics, I have indicated where the pronoun and particle are written phonogrammatically to ensure the reliability of the data.
3.3. Deficient Pronouns

Unlike strong pronouns, deficient pronouns cannot be left morphologically unmarked. This may be attributed to the generalization (13b) proposed by Cardinaletti and Starke. The case feature of deficient pronouns must be overtly realized by a case particle. \((W)a\) and \(na\), which are the deficient counterparts of \(ware\) and \(nare\), are accompanied by the topic marker \(pa\) (22a), and the case particles \(ga\) (22b–d) and \(wo\) (22e) (and in some cases, \(ni\)), but they do not cooccur with other case particles.\(^7\) Following Kayne’s (1994) antisymmetry hypothesis, I suggest that \(pa\), \(ga\) and \(wo\) are functional heads and that the clitics \((w)a/na\) are directly adjoined to these heads (see section 3.5):

\[(22)\]  
\[\text{a. wa pa (和波) soto mo pa-ji (MY 3451)}\]
\(\text{I Top outside Foc chase-Not}\)
\(\text{‘I don’t even chase outside.’}\)
\[\text{b. nageki so a ga (安我) suru (MY 3524)}\]
\(\text{sigh Foc I Gen do}\)
\(\text{‘I heave a heavy sigh.’}\)

\(^7\) In the \(Man’yôshû\), there are three examples (3167, 3468, 3478) in which \(ni\) appears with clitic pronouns, but in all cases, they are used with the verb \(yosu\) ‘attract’ selecting the \(ni\) marked NP as in (i):

\[(i)\]  
\(\text{na ni yotsori-keme (MY 3468)}\)
\(\text{you Dat attract-Aux}\)
\(\text{‘I might have been attracted to you.’}\)

Yanagida and Whitman (2009) argue that \(ni\) in this use originates as a head of a small clause. The locative use of \(ni\) never appears with the clitics.
SYNTACTIC REANALYSIS OF PRONOUNS AS DEMONSTRATIVES

407

c. yumapi wa ga (和我) suru (MY 4382)
  disease I Gen do
  ‘I have a disease.’

d. na ga (奈我) nake ba (MY 3785)
  you Gen cry when
  ‘when you cry …’

e. na wo (那乎) do kamo si-mu (MY 3556)
  you Obj what Q do-Aux
  ‘What shall I do with you?’

While a clitic only appears as a core argument of a clause, marked by the case particle ga or wo, weak pronouns cooccur with all types of case particles, showing that the latter need not be the argument of a clause. (For further discussion, see section 3.5.)

According to Cardinaletti and Starke, weak pronouns have no referential feature; thus as stated in (13c) they can be interpreted as referential only if they are associated with a (non-deficient) antecedent, through coreference. OJ third person pronouns si/so are weak pronouns in that, as Hashimoto (1966) points out, they only appear in contexts where a specific antecedent, such as pana ‘flower’ in (23a), ayu ‘sweet fish’ in (23b), is present in the preceding discourse:

(23) a. wa ga yado ni pana so saki-taru so
  I Gen garden Loc flower Foc bloom-Past it
  wo (乎) mire do kokoro mo yuka-zu (MY 466)
  Obj see but mind Foc go-not
  ‘In my garden, flowers are in bloom. But as I see them, I am never comforted.’

b. tora-samu ayu no si ga (之我) pata pa ware ni
catch-Aux sweetfish Gen it Gen fin Top I Dat
  kaki-muke (MY 4191)
toward
  ‘capture the sweetfishes, put them on dishes with all their fins toward me’

The demonstrative ko, the deficient counterpart of kore, may be used deictically, but as Hashimoto points out, ko tends to refer to the previously mentioned element in a given discourse. In (24) ko refers to the moment or situation when the speaker is sitting alone, thinking:
(24) pitori wite mono opmopu yopi ni pototogisu ko alone sit thing think evening Loc cuckoo this yu (此間) naki wataru (MY 1476) from cry pass

‘In the evening as I think of things sitting all alone, a cuckoo passes by (through this scene) crying.’

Finally, while full (strong) pronouns with -re appear with kakari-focus particles, deficient pronouns do not. This follows from Cardinaletti and Starke’s generalization (14) which states that deficient pronouns are never focused.

3.4. The Position of Functional Heads

As we have seen, weak pronouns are differentiated from clitics in that weak pronouns can freely appear with a case particle, while clitics are restricted to cooccurring with the topic particle pa or the case particles ga and wo. Yanagida (2005: 121–124) suggests that the cooccurrence restriction between clitics and case particles is accounted for by assuming that certain particles are clausal heads in the light of Kayne’s (1994) antisymmetry hypothesis. Kayne (1994: 143) proposes that (Modern) Japanese nominative ga and accusative o are clausal heads (T and v respectively) that select a complement on their right, as represented in (25) (see also Whitman (2001)):8

\[(25)\]

\[
\begin{array}{c}
TP \\
\text{NP}_{\text{SUBJ}} \quad T'
\end{array}
\]

\[
\begin{array}{c}
\text{ga} \\
\text{NP}_{\text{OBJ}} \\
\text{o}
\end{array}
\]

\[
\begin{array}{c}
vP \\
v'
\end{array}
\]

\[
\begin{array}{c}
\text{VP}
\end{array}
\]

Given that clitics occur in a derived (functional) position, as stated in (4a) by Cardinaletti and Starke (also Kayne (1991)), we can provide a straightforward account for why the clitics appear with ga and wo, but not with other case particles. Ga and wo (the ancestor of accusative o) are functional heads and the clitics are head-adjoined to these particles. Clitics do

8 Under Kayne’s (1994) head initial hypothesis, the object NP and the accusative o do not form a constituent. The scrambling of an object NP may be analyzed as movement of the remnant vP containing tVP.
not cooccur with postpositional case particles because movement of a clitic to a lexical head is illicit. Weak pronouns, on the other hand, project to a phrasal level according to Cardinaletti and Starke, and thus can appear with a postposition, as in ko yu ‘from this’ ((24) above), with the structure \([\text{PP} \ [\text{XP} \ ko] \ yu]\).

Yanagida (2006) points out that OJ has a peculiar word order restriction; when the subject and object are both case-marked, the object necessarily precedes the subject. The clitic adjoined to genitive \(ga\) must appear immediately before the nominalized (\(rentai\)) verb (head adjunction is represented by ‘\(=\)’ as in \(wa=ga\)):\(^9\)

\[
\begin{align*}
\text{(26)} & \quad \text{pito dumakoro wo iki ni wa=ga suru} \\
& \text{(MY 14/3539)} \\
& \text{person wife Obj long for I Gen do} \\
& \text{‘I long for another person’s wife.’}
\end{align*}
\]

Under Yanagida’s (2005: 102) analysis, genitive \(ga\) is the head of AgrP and the clitic \(wa\) is head-adjoined to \(ga\). The clitics \((w)a/na\) are moved from Spec, NP/VP, the position where the external theta role is assigned, and left-adjoined to Agr:

\[
\begin{align*}
\text{(27)} & \quad \text{AgrP} \\
& \quad (w)a_i=ga \ NP/VP \\
& \quad \quad t_i \ NP/VP
\end{align*}
\]

The idea that OJ genitive \(ga\) is the head of AgrP captures the word order restriction as shown in (26) and also the parallelism between nouns and nominalized clauses in OJ.\(^{10}\) The structure is also similar to the one proposed by Cardinaletti and Starke (1999) in (12c); clitics appear in the functional category specified for agreement, which appears immediately above a lexical category LP (either nominal or verbal).

Finally, OJ has a second genitive particle, \(no\), the ancestor of the modern standard Japanese genitive marker. Note that clitics never appear with \(no\), but weak pronouns can. A genitive \(no\) phrase in OJ, as in ModJ, behaves

---

\(^9\) Approximately 120 occurrences of subject \((w)a=ga\) are found in the \(Man’yōshū\) (based on the Yoshimura’s electronic text); all are immediately adjacent to the verb. Data cited here include personal pronouns written with phonographs but not the freestanding ideograph 吾, which can be read with or without a case particle, depending on the metrical context.

\(^{10}\) Assuming with Chomsky (1995) that AgrP is eliminated and replaced by \(vP\), Yanagida (2006) proposes that the OJ subject with \(ga\) appears in Spec, \(vP\), and that the object obligatorily moves to Spec, CP.
like an adjective in that it can iterate within a nominal projection. A genitive *ga* phrase, on the other hand, cannot iterate. Thus, in (28a) the two genitive phrases modify the head noun, whereas in (28b), each genitive modifies the NP that immediately follows it:

(28) a. watatumi no izure no kami
   watatumi Gen which Gen god
   ‘which god of Watatumi’
   b. wa=ga seko ga onpune
   I=Gen lover Gen ship
   ‘the ship of my lovers’

The structures of (28a) and (28b) are represented in (29a) and (29b) respectively:

(29) a. [NP [PP watatumi no] [N izure no] [N kami]]
   b. [AgrP [AgrP wa=ga [NP seko]] ga [NP onpune]]

The genitive *no* is the head of PP, whereas the genitive *ga* is the Agr-head that takes the NP complement on its right. The fact that weak pronouns, but not clitics, cooccur with genitive *no* supports the view that *no* is simply a postposition, just like other case particles such as *yu*, *yori* ‘from.’

4. Degrammaticalization of Pronouns

The tripartite pronominal system of OJ almost completely disappeared in Early Middle Japanese (EMJ), after 800. Deficient pronouns were lost and replaced by their strong counterparts *ware*/*sore*. The clitic *wa* came to be used only in the conventionalized form *waga*. As discussed above, the OJ clitics *w(a)=ga/na=ga* appear strictly adjacent to adnominal (*rentai*) verbs. In EMJ, however, *waga/naga* acquires greater positional freedom (“deflexion”). Thus, in *Konkōmyō Saishō Ōkyō* (The Sutra of Golden Light), there are quite a number of examples in which *waga/naga* appear in clause initial position, as illustrated in (30a, b):\(^{11}\)

---

\(^{11}\) *Konkōmyō Saishō Ōkyō* (The Sutra of Golden Light) appeared in the late 8th century. It was originally written in India and was translated into Chinese in 703. This Chinese text was read in Japanese through a system called *haku-ten* ‘white glosses’ which appeared on the original Chinese text, and were used as a way of glossing the Chinese to be read in Japanese. What is crucial is that since these markings were added to the original text by Buddhist monks in the early *Heian* Period, we are able to reconstruct the language of that period. (Readings for the text are provided by Kasuga (1969).)
SYNTACTIC REANALYSIS OF PRONOUNS AS DEMONSTRATIVES

(30) a. waga katari wo idasa-simu toki … (Ch. 8: 144)
I.Gen talk Obj make-Aux when
‘the time when I made a word to be spoken’
b. naga yoku kono Myaugyau-wau-wo
you.Gen successfully this Myaugyau-wau-Obj
rufu-si … (Ch. 8: 146)
propagate-Past
‘You propagated this Myoukyouou successfully.’

These examples show that the morphologically complex \((w)a=ga\) came to be an independent pronoun and appear in the subject position, possibly Spec, TP. A deflexion process of this kind is a well-known type of de-grammaticalization as cited in the literature (see Allen (2003)).

The second important development is what Whitman (1999) calls “intrapersonal pronoun shift” whereby first person pronoun \(\text{ware/waga}\) came to refer to second person (i.e. the hearer). (31) is from \(Uji-shūi\ monogatari\) (1218) (cf. Whitman (1999)):

(31) ware pa miyako no pito ka. Iduko pe opasuru zo
you Top capital Gen person Q where to go Foc
(Uji-shūi)

‘Are you from the capital? Where are you going?’

In OJ \(\text{ware}\) simply appears as a neutral first person pronoun used without any regard to gender or social status. But once \(\text{ware}\) appears as a hearer-designator, it comes to convey various stylistic and sociolinguistic implicatures. In Late Middle Japanese (LMJ) \(\text{ware}\) (second person) is used towards people ranking lower on the social scale. Ultimately it came to carry a derogatory implicature (cf. Nakamura et al. (1982)). This is illustrated in (32a, b):

(32) a. ware ni kaka-se tatemu-ramu to omofu ni
you Dat write-Cau Hon-Aux that think Loc
(Ōkagami, Saneyori den)
‘(I) think I would let you write.’
b. itu ware ga ore ni sake wo kure-ta zo
when you Nom me Dat sake Acc give-Past Foc
(Kyōgen, Kofu muko)
‘When did you give me sake?’

(33) is an example of \(\text{waga}\) being used for second person with a derogatory implicature:
According to Whitman (1999), intrapersonal pronoun shift is widely observed in East Asian languages that lack agreement, and in the case of shift from first to second person, it is always mediated by a diachronic stage where the pronoun has a reflexive function SELF. Note that while strong pronouns with the plural suffix -ra are not found in the OJ corpus, we find some examples of *ware+ra* in MJ texts.\(^{12}\) (34a, b) are from *Utsubo monogatari*:

\[
\begin{align*}
\text{(34) a. } & \text{Tennyo no ifaku, saraba ware-ra ga omofu} \\
& \text{heaven.maiden Gen say, then I-Num Gen think tokoro … (Utsubo)} \\
& \text{‘A heavenly maiden said, then I think that …’} \\
\text{b. [ware-ra ga ko] pa oya ni masaru nasi} \\
& \text{I-Num Gen child Top ancestor Dat surpass not (Utsubo)} \\
& \text{‘My children do not surpass their ancestors.’}
\end{align*}
\]

Nakada et al. (1994) indicate that in MJ *ware+ra* can be interpreted as either singular or plural. In (34a, b) *ware-ra* is unambiguously interpreted as first person singular. But it is natural to assume that this singular use of *ware+ra* has a reflexive function; thus (34a) has the meaning *zibun ga omou tokoro* ‘self thinks that …’ comparable to ModJ. Note that while in Indo-European languages, reflexives, such as *myself/yourselves*, have person features, in East Asian languages, reflexives bear no person specification (i.e. *zibun* ‘self’ in Japanese, *ziji* ‘self’ in Chinese). Following Whitman (1999), I propose that intrapersonal shift is a diachronic process in which a pronoun loses its agreement features (i.e. is depersonalized) and acquires a reflexive meaning. This change is formally analyzed as the loss of AgrP within the extended nominal projection (see section 5.3).

Lexical to functional reanalysis leads to what is known as “semantic bleaching,” that is, loss of lexical semantic meaning. Roberts and Roussou (2003), however, argue that semantic bleaching does not mean that a category loses all lexical content; some part of the lexical meaning may remain

\(^{12}\) I counted the occurrences of *ware+ra* in the OJ and MJ texts available in *Nihon Koten Bungaku Honbun* Data Base from the National Institute of Japanese Literature.
in the reanalyzed functional element. The semantic change indicated in (31)–(34), however, is counter to grammaticalization. *Ware/waga* started out as first person pronouns which lacked any descriptive or sociolinguistic content. They were used regardless of social status, in-group/out-group status, etc., unlike the so-called personal pronouns of ModJ. Over time, however, they came to acquire lexical properties, specifically, the derogatory sense found in LMJ.

I argue that the changes illustrated above, summarized in (35a–c) are direct consequences of the loss of AgrP:

(35) a. deflexion of clitic *waga*
   b. intrapersonal shift (first person > reflexive > second person)
   c. rise of new lexical content

In section 5.3, it is argued that the loss of AgrP leads to the categorial reanalysis of pronouns as demonstratives, involving what Norde (2006, 2009) defines as degrammaticalization ((2) above).

5. A Formal Analysis of (De)grammaticalization-type Categorial Shift

5.1. From Demonstrative to Definite Article/Complementizer

In recent studies, it is proposed that demonstratives are not unanalyzable categories in the grammar but decompose into the three morphological elements: definite, deictic and noun, as illustrated in (36) (cf. Kayne and Pollock (2009), Giusti (1998, 2001) among others):

(36) a. that book: *th-* (definite) -at (deictic) book (noun)
   b. this book: *th-* (definite) -is (deictic) book (noun)

Kayne and Pollock (2009) suggest that the English demonstratives *that* and *this* without an overt noun decompose into *th* (definite), *at/is* (deictic), and the unpronounced noun labeled THING. Giusti (1998, 2001) presents a similar analysis for the demonstratives of Romance and Germanic languages. She argues that demonstratives are associated with [+definite] and [+deictic] features, and project up to DemP. Under Giusti, definiteness is specified in D, and DemP originates in AgrP and moves to Spec, DP in the extended nominal projection, as represented in (37):

(37) \[\text{DP \ DemP} \_\text{D +def [AgrP} \_\text{t} \_\text{Agr [NP \ldots]]}]]\]

Following Renzi (1997), Giusti argues that the development of the Romance definite article from the Latin demonstrative *ille* can be viewed as a grammaticalization process whereby DemP in Spec, DP is reanalyzed as the head of DP (Giusti (2001: 167)): 
Roberts and Roussou (2003) discuss a similar reanalysis for the development of Germanic complementizer *that* from a demonstrative pronoun. According to Roberts and Roussou, demonstrative and complementizer are variants of the same basic abstract category. They differ in that demonstratives have a deictic feature realized as the morphological contrast *this*/that and a number feature realized e.g. as *this*(Sg)/*these*(Pl). Number structurally corresponds to a NumP in the extended nominal projection and the demonstrative *that* moves from Num to D, as shown in (39) (cf. Ritter (1991) among others):

\[
(39) \quad \begin{array}{c}
\text{DP that/those, } \begin{array}{c}
\text{[DemP D'}}
\Rightarrow
\text{Spec D'}
\end{array}
\end{array}
\]

Relative clauses in the older Germanic languages are characterized as having both a demonstrative and a complementizer. The demonstrative synchronically moves from inside IP to Spec, CP (Roberts and Roussou (2003: 119)):  

\[
(40) \quad \begin{array}{c}
\text{[CP that, } \begin{array}{c}
\text{[CPrt] IP } t_i \ldots]
\Rightarrow
\text{[CP [that (+Prt)]]}\end{array}
\end{array}
\]

Diachronically, when language learners encounter no direct evidence that the demonstrative *that* moves from inside IP to Spec, CP, *that* comes to be base-generated under Comp. Loss of movement, that is, results in the categorial reanalysis of *that* as a complementizer. The semantic change is a direct consequence of the category change. Roberts and Roussou provide ample evidence that loss of movement leads to grammaticalization, and they claim that grammaticalization always involves structural simplification, as illustrated in (38) and (40).

5.2. Japanese Demonstratives

As discussed above, in Romance and Germanic languages, definiteness is specified in the functional head D, and demonstratives, which project up to DemP, move to Spec, DP to check their definite feature. Japanese demonstratives, however, need not move to Spec, DP. Consider the contrast given in (41) and (42) (cf. Fukui (1995)):

\[
(41) \quad \begin{array}{c}
a. \quad \begin{array}{c}
\text{[akai kono/sono fuku] o karita red this/that clothes Acc}
\end{array}
\end{array}
\rightarrow
\begin{array}{c}
\text{I borrowed red this/that clothes.'}
\end{array}
\]
b. [John no kono/sono fuku] o karita
   John Gen this/that clothes Acc borrow
   ‘*I borrowed John’s this/that clothes.’

(42) a. [akai kore/sore] o katta
   red this/that Acc bought
   ‘*I bought red this/that.’

b. [John no kore/sore] o katta
   John Gen this/that Acc bought
   ‘*I bought John’s this/that.’

Unlike their English counterparts, Japanese demonstratives can appear lower than the adjective or the genitive phrase. Furthermore, Whitman (personal communication) points out that Japanese demonstratives allow an indefinite reading, while their English counterparts do not. This is shown in (43):

(43) a. kono mise ni wa sono hon ga 5 satu aru
   this store Loc Top that book Nom 5 CL be
   ‘*There are five of that book in this store.’

b. Tukuba si ni wa kono kuruma ga nandai ka aru
   Tsukuba city Loc Top this car Nom some Q be
   ‘*There are some of this car in Tsukuba.’

A reviewer points out that in examples like (43), when the quantifier is floated, 

sono ‘that’ refers to a particular entity in a given discourse context. I argue that the “particular entity” reading in (43) comes from the feature [deictic], not the feature [definite]. The idea that sono, unlike English that, has no specification of definiteness is supported by the fact that sono can have a bound variable interpretation, as suggested by Hoji (1989). Thus, (44) is acceptable with the reading in which sono is bound by the universal quantifier dono ‘every’:

(44) dono gakusei mo, sono gakusei, no soba ni hana
ey every student Q that student Gen next Loc flower
   ga oite aru
   Nom have be
   ‘*Every student has a flower next to that student.’

The English demonstrative that is inherently specified for the features [definite] and [deictic], and hence an indefinite reading is not allowed. Japanese sono ‘that,’ on the other hand, is specified for the feature [deictic] and is thus referential, but is not necessarily definite.

Furthermore, Hoji (1989) argues that unlike so-type demonstratives, the a-type demonstratives, that is, are/ano ‘that’ and third person kare ‘he’ cannot be construed as bound variables. (45a, b) are taken from Hoji (1989):
(45) a. *daremo ga [NP [s kare ga tukut-ta] omotya] o
everyone Nom he Nom make-Past toy Acc
kowasi-ta
break-Past
‘Everyonei broke the toy that hei had made.’
b. *daremo ga ano hito no i hon o
everyone Nom that person Gen book Acc
sute-ta
throw away-Past
‘Everyonei threw away that personi’s book.’

Given that kare ‘he’ behaves like the demonstrative are/ano ‘that,’ Hoji proposes that kare is not a pronoun but a demonstrative. Note that there is a general consensus that OJ kare (which is very rare) is also used not as a pronoun, but as a third person demonstrative (Omodaka (1967)):

(46) tare so kare to ware wo na topi so (MY 2240)
who Foc that that I Obj Neg ask Foc
‘Don’t ask me who that person is.’

Li (2002: 167), however, claims that in (46) there is no need to assume that there is a third person present in the speaker’s location at the time of utterance. Kare can be interpreted as referring to the hearer. Li, in fact, claims that the demonstrative kare is not specified for person because in OJ through MJ, kare can be used to refer to either second person or third person present in the location of the speaker. In (47) cited by Li (2002: 168), kare unambiguously refers to second person (i.e. the hearer):

(47) san nin no fito toite iwaku, kare wa namu zo
three Gen person ask say you Top what Q
no fito zo (Utsubo)
Gen person Q
‘Three people asked and said, What kind of person are you?’

The fact that Japanese demonstratives can occur lower in the extended DP structure than adjectives and that so-type demonstratives have an indefinite reading supports the view that definiteness is not an inherent feature associated with Japanese demonstratives. Furthermore, in ModJ and at earlier stages, kare is specified for the [+deictic] feature, but has no specification for person.

5.3. From Pronoun to Demonstrative: Degrammaticalization

As noted in section 4, all the deficient pronouns—both weak and clitic pronouns—were lost in EMJ and replaced by their strong counter-
Based on Norde’s (2006, 2009) definition of degrammaticalization in (2), I suggest that reanalysis of the strong pronoun *ware* as a demonstrative is an instance of degrammaticalization: *ware* loses its grammatical status and acquires new semantic content. The change is exactly opposite in direction to the change of the Latin demonstrative *ille* into the Romance strong pronouns, as analyzed by Giusti (1998, 2001) in (38). The OJ strong pronouns *wa-re/na-re* have person agreement morphologically realized as *(w)a/na* generated in Agr, while the suffix *-re* has its definite feature specified in D. Strong pronouns are morphologically derived by movement. The spellout of agreement, *wa/na*, generated inside NP, moves to D to check the definite feature, as in (48):

(48) Strong Pronouns

\[
[\text{DP} [\text{wa}=\text{re}] [\text{AgrP} t [\text{NP} t \ldots]]]
\]

The difference between pronouns and demonstratives is the presence or absence of a person and a number feature. A demonstrative is specified for a number feature, while a pronoun is specified for a person feature. The reanalysis of *ware* as demonstrative is then represented as in (49):

(49) a. DP

\[
\begin{array}{c}
\text{D} \\
\text{AgrP} \\
\text{wa-re} \\
\text{Agr} \\
\text{NP}
\end{array}
\]

b. DemP

\[
\begin{array}{c}
\text{Dem} \\
\text{NumP} \\
\text{wa-re} \\
(-\text{ra}) \\
\text{NP}
\end{array}
\]

\[
\{ \text{definite} \} > \{ \text{deictic} \}
\]

\[
\{ \text{person} \} > \{ \text{number} \}
\]

The loss of the agreement feature is triggered by an increase in the occurrence of the plural form *ware+ra* in MJ (see section 4). The idea that the plural suffix *-ra* is the head of NumP is supported by the fact that *-ra* selects an appositive NP, as illustrated in ModJ (50):

(50) a. *[kare ra gakusei] no kibou

him PI student Gen hope

‘the hope of them students’

b. *[kare gakusei] no kibou

him student Gen hope

‘the hope of him student’

The contrast between (50a) and (50b) can be accounted for by assuming that the suffix *-ra* is the head of NumP and selects the NP; that is, that (50a)
has the structure \[[\text{DemP} \text{kare} [\text{NumP} \text{ra} [\text{NP} \text{gakusei}]])\]. The pronoun > demonstrative shift, as shown in (49), is similar to what Whitman (2000) calls “relabeling.” Whitman (2000) argues that categorial reanalysis of a head may affect syntactic structure, but only in the “minimal domain” (Chomsky (1995)) of its specifier and complement. Under the present analysis, the reanalysis of D as Dem results in a change in its selectional properties—specifically, in what kind of complement it selects.

The strong pronoun *ware* does not appear with the nominative marker *ga* until Early Modern Japanese (EModJ). The *Amakusaban Heike* (1592) contains 85 tokens of *ware*, but no tokens of *ware+ga*. There are, however, 16 tokens for the plural form *ware+ra*, and 7 of these appear with *ga*. Similarly, *Esopono Fabvlas* (1593), another romanized Jesuit text, contains no tokens of *ware* with *ga*, but 13 tokens of the plural *ware+ra* with the genitive *ga*. This indicates that the genitive *ga* first appears in the form of *ware+ra* (Pl)+*ga* in MJ. We see thus that *ware+ra* has a full DemP structure (49b). On the analysis I have proposed it appears in the specifier position of the extended nominal or verbal projection headed by *ga*:

\[
\begin{align*}
(51) & \quad \text{DP/IP} \\
& \quad \text{DemP} \quad \text{D/I'} \\
& \quad \text{ware-ra} \quad \text{ga} \quad \text{NP/VP}
\end{align*}
\]

In OJ through MJ, *ware* never appears with *ga* because *ga* in Agr would block the movement of *wa*, due to the head movement constraint (HMC). The licit derivation is shown in (48). Once *ware* is reanalyzed as DemP, it comes to be able to cooccur with *ga*.

In EModJ texts, we find some examples in which *ware* appears with *ga*. In (52a), *ware* is used as reflexive comparable to ModJ *zibun* ‘self’. In (52b) it is used as a second person designator with a derogatory

---

13 Under the DP hypothesis, Abney (1987) proposes that determiners and pronouns are the functional head D that selects NP as their complement. A pronoun followed by an appositive noun phrase can function like the definite article the, as in \[[\text{DP we/the} [\text{NP linguists}]]\].

14 The statistical data are taken from Eguchi (1986) and Otsuka and Kita (1999).

15 As discussed in section 4, the categorial reanalysis of *ware* as demonstrative is mediated by a diachronic stage where *ware* has a reflexive function SELF. This text uses a mixture of artificially archaic classical Japanese (e.g. -keri, the classical narrative past) and ModJ (e.g. -masyoo, the modern propositive). *Ware* appears to be fundamentally reflexive (for the use of *ware* in the Edo period, see Maeda (1974)).
implicature:

(52) a. ware ga kago ni noru koto wasure keru mo
   self Nom palanquin Loc ride that forget Past Foc
   funny
   ‘It’s funny that he (self) forgot to ride in the palanquin.’

b. itu ware ga ore ni sake wo kureta zo
   when you Nom me Dat sake Acc give-Perf Foc
   (Kyôgen, Kofi muko)
   ‘When did you give me sake?’

These examples in EModJ show that the strong pronoun *ware* lost its (person) agreement feature and was reanalyzed as a demonstrative.

To summarize, I have proposed that OJ *ware* has the full DP structure (49a) and that it is reanalyzed as DemP, as in (49b), due to loss of AgrP in the extended nominal/verbal projection. I have discussed the D > Dem reanalysis in terms of evidence provided by the emergence of the plural form *ware*+*ra*, intrapersonal shift mediated by a reflexive function, and the emergence of *ware* with the D head *ga*.

6. Summary

In this paper, I have presented syntactic evidence that OJ possesses three classes of pronouns comparable to the three-way division in Romance languages. When the defective pronouns were lost in EMJ, learners were no longer presented with evidence that pronouns in Japanese contained an agreement projection. The loss of AgrP results in an overall change in the pronominal system. Demonstratives were replaced with pronouns (such as *kare, sonata, anata*), the surviving strong pronoun *ware* was reanalyzed as a demonstrative, and epithets (*boku, watakusi*) were co-opted into the role of referentially dependent pronouns. Following Norde’s (2006) definition stated in (2), I have argued that this process is an instance of degrammaticalization.

Roberts and Roussou (2003) propose that grammaticalization always involves structural simplification. Thus, according to Roberts and Roussou, the development of the Latin demonstrative *ille* into a definite article involves structural simplification in that a full structure DemP is shifted to the D head. In contrast, the reanalysis of pronouns as demonstratives does not imply any structural simplification, but it may be analyzed as a categorial change of one element into another as illustrated in (49), a process similar
to “relabeling” in the sense of Whitman (2000).

REFERENCES


Renzi, Lorenzo (1997) “Fissione di Lat. *ILLE* Nelle Lingue Romanze (The Split of


Primary Texts


Ogihara, Asao and Hayao Konosu (1979) *Kojiki Jôdai Kayô* (Poetry from the *Kojiki* and Other Old Japanese Texts), Shogakukan, Tokyo.

**Corpora**

Yoshimura, Makoto (http://infux03.inf.edu.yamaguchi-u.ac.jp/~manyou/ver2_2/manyou.php)

**Dictionaries**


[received April 20 2010, revised and accepted August 9 2010]

Institute of Modern Languages and Cultures
University of Tsukuba
Tennodai, Tsukuba
Ibaraki 305–8571
e-mail: yanagida@sakura.cc.tsukuba.ac.jp