

Corpus-based Study of Word Formation: Anti-agentive Nominals in English and Japanese

Junya Morita
Kinjo Gakuin University
Japan
morita@kinjo-u.ac.jp

Abstract

This study reveals the characteristics of anti-agentive nouns in English and Japanese (cf. *instructee*, *hi-shidoo-sha*) and then proposes the underlying word formation mechanism. Related derivatives extracted from two large-scale corpora are examined in detail from grammatical and functional perspectives. The result exposes four generalizations: (i) both languages have core devices for their formation that involve high productivity values (*P*), anaphoric conceptualization, and the compositional core meaning ‘one who is V-ed,’ (ii) there exist unproductive subsidiary types that deviate in part from the core meaning (see *attendee* without a passive reading and *enhancee* without a personal reading), (iii) Japanese has a one-to-one correspondence between semantic units and morphemes, but English does not (cf. *hi-shidoo-sha* [passive-instruct-person] and *instructee* [instruct-passive.person]), and (iv) Japanese has two types of anti-agentive nouns, *hi-shidoo-sha* and *taiho-sha* ‘arrestee.’ We then show that an “anti-lexical” theory reconciles the common core aspects with the morphological difference, each deviation in the subsidiary types is related to “grammaticalization,” and pragmatic strategies are at work in the selection of competing anti-agentive forms.

Keywords: anti-agentive nouns, corpus research, generalizations, word-formation systems, English, Japanese

1 Introduction

This study investigates in detail the nominal expressions seen in (1).

- (1) Unlike many, Martin was able to get through her prepared remarks, whereas more often than not, *acceptees*’ thank-yous were interrupted mid-phrase by “the orchestra

from Hell,” which functioned as a sort of musical hook if *awardees* went one second past their allotted thirty seconds per speech. (*Plays and Players*, July 1993, p. 45)

The derived nouns *acceptee* and *awardee* share the meaning of ‘one who receives an action’ and involve a covert passive voice. They are therefore called “anti-agentive nouns.” Lexically, they are very low frequency words and are hence involved in a creative word formation that generates words on the spot.

The aim of this study is to collect extensive data on anti-agentive nominals in English and Japanese using large-scale corpora, to clarify their similarities and differences based on this data, and then to theoretically explain their origins. Based on extensive data collected from large corpora, we disclose the morphosyntactic, semantic, and functional properties of *-ee* derivatives (§2), other type of anti-agentive nouns (§3), and their Japanese counterparts (§4). Section 5 compares anti-agentive nominals in English and Japanese, presents several empirical generalizations, and then show their theoretical implications.

2 The Properties of *-ee* Nominals

2.1 Lexical properties

This section deals with two lexical properties of *-ee* derivatives. First, they exhibit lexical attributes in that they conceptualize “person receiving an action.” A good example of this is provided in (2).

- (2) He’s Paul Klemchalk, “*laid off* at 43,” one of the thousands felled by AT&T’s downsizing ax. ... Klemchalk’s timely resuscitation may be part of a big AT&T PR campaign to soften its image as a heartless job-slasher—an effort that included full-page newspaper ads last week touting its job bank for *lay-off-ees*. (*Newsweek*, March 25, 1996, p. 42)

The context of discourse (2) is that a person was laid off and then a certain measure was taken against those who were laid off. In response to the antecedent verb phrase, a situation arises in which “the people felled by AT&T” are conceptualized. A word, rather than a phrase, is then used to conceptualize these people. Consequently, the anti-agentive noun *lay-off-ee* is innovated on the spot. This word is a neologism because it is not found in the British National Corpus (BNC), a 100-million-word corpus, or large dictionaries like *The Oxford English Dictionary* (2nd ed.) (*OED*). Similarly, in example (1) observed above, the neologism *acceptee* is coined to label a contextually established concept.

The second lexical property concerns productivity. *Acceptees* and *awardees* in (1) are BNC hapaxes—token frequency 1¹. A complex word of very low frequency is quite likely to be created on demand rather than stored in the lexicon (Hay, 2003). Accordingly, the degree to which words of this kind can be produced is an indicator of the productivity of a word formation device (Baayen and Renouf, 1996; Plag, 1999). A productivity measure has been proposed: *Productivity* (P)= n_1/V , where n_1 is the number of hapaxes and V is the total number of word types. The productivity value (P) can then indicate the probability that a word type in *-ee* is a neologism when it is formed. Our BNC survey has detected 22 hapaxes and 94 word types of *-ee* derivatives, giving its productivity value of 0.234. The productivity of *-ee* means that one out of four word types in *-ee* used is a new word, thus indicating the high potential of *-ee* suffixation to generate new words.

In summary, *-ee* derivation is highly creative: it can coin a relevant word whenever the recipient of a particular action needs to be conceptualized in an appropriate context.

2.2 Semantic properties

Previous studies on *-ee* nouns have tended to simply list their various meanings (Barker, 1998; Marchand, 1969; Huddleston and Pullum, 2002). Below, by scrutinizing related words collected from a large corpus, we will perform an “elegant”

semantic classification of *-ee* derivatives and demonstrate that their meanings are divided into a single basic meaning (Patterns 1-3) and meanings that partially deviate from it (Pattern 4). The semantic features of each pattern are exemplified, while also mentioning the productivity of each pattern.

2.2.1 $V_{tr} + -ee$ (Pattern 1)

Pattern 1 consists of a transitive verb and the suffix *-ee*, which can be divided into three groups: A, B, and C. Type A is one in which the base verb participates in the passive voice and its direct object represents the referent of *-ee* derived nouns, as in *interviewee* and *examinee*. This type of *-ee* derivatives can be paraphrased as ‘one who is V-ed,’ and Marchand (1969: 268) calls this type “direct passive nouns.” Our BNC research has found 62 word types of this sort out of a total of 94 word types in *-ee*; Type A accounts for 66.0 % of the total. Moreover, 19 hapaxes have been detected among the 62 word types of Type A *-ee* nouns, giving its productivity value of 0.306. This type then is judged to be the prototype of *-ee* derivatives.

In Type B *-ee* derivatives, the underlying dative verb is passively involved and its indirect object is the referent of *-ee* nouns. This type then is interpreted as ‘one who is V-ed something.’ In our BNC survey, five word types of this sort have been discerned: *allotee*, *assignee*₂, *legatee*, *presentee*, *providee*².

In addition, Type C is related to “prepositional passive,” in which the object of the preposition represents the referent of *-ee* nouns. Our BNC survey has detected only one word type of this kind, *conferee*, which is paraphrased by ‘one who is conferred with.’

A word needs to be said here about the productivity of Type B and Type C. In addition to the productivity measure presented in the last section (P), we assume “the global productivity P^* ”: P^* of affixation is defined in terms of its coordinates in the P - V interaction region, with productivity (P) on the horizontal axis and word types (V) on the vertical axis; productive affixation occupies a central position in the region (Baayen and Lieber, 1991: 818-819). According to this definition, cases with a very small number of word

¹ I am indebted to the research engine of www.english-corpora.org/bnc/. By repeatedly using the “wild card” function of the research engine, words ending in *-ee* are extracted from BNC.

² If the underlying verb is interpreted in two ways, it is separately counted: *assignee* means ‘one who is assigned’ and ‘one to whom something is assigned,’ and hence it belongs to both Type A (*assignee*₁) and Type B (*assignee*₂).

types, such as Type B and Type C *-ee* derivations, fall outside the domain of productive processes.

2.2.2 N + *-ee* (Pattern 2)

It has been pointed out that a noun can be used as the base of an *-ee* derivative other than a verb (Barker, 1998: 705; Huddleston and Pullum, 2002: 1697). Contrary to the previous analyses, we show below that the base of this type of *-ee* derivative is not actually a noun. It is well known that one of the distinctive features of English word formation is that conversion from noun to verb produces an abundance of verbs. Quirk et al. (1985: 1561) classify the conversion from N to V into seven and point out that most of these verbs are transitive. One of the seven semantic classes is ‘ornative’: the verb *coat* is interpreted as ‘to give a coat (of paint) to,’ the base noun being converted referring to an entity given to something or somebody. This process is quite productive, giving rise to various verbs such as *poison*, *lipstick*, *bribe*, *lecture*, and *money*³.

We claim that the *-ee* nouns at issue are derived via this type of conversion: the noun *patent* is converted into a verb (‘to give a patent’) and then the suffix *-ee* is added to it to derive the noun *patentee*. Thus, the suffix *-ee* that appears to be attached to a noun is actually connected to a verb. There are six related word types found in BNC, and according to “the global productivity *P**” (cf. §2.2.1), Pattern 2 is not productive. Interestingly, of these six word types, *off-licencee*, *therapee*, and *awardee* are hapaxes, and furthermore, *off-licence* and *therapy* are not listed as verbs in major dictionaries. This fact indicates that these *-ee* nouns are not created directly from verbs, but rather through the process described above⁴.

2.2.3 Stem + *-ee* (Pattern 3)

Let us go on to a case in which *-ee* is superficially added to a verbal stem. Four relevant examples have been detected in BNC: *evacuee*, *nominee*, *amputee*₁, and *rehabilitatee*. It can be argued that this type of *-ee* derivative is produced by dropping the internal suffix *-ate* when *-ee* is added to an *-ate* final verb. First, there is a semantic reason for

this view. The readings of *evacuee* and *nominee* are ‘one who is evacuated’ and ‘one who is nominated,’ respectively, and if the suffix *-ate* is underlyingly involved, the reading of each derivative is easily and naturally defined. The truncation at issue is also justified by a context-dependent neologism, as exemplified in (3).

- (3) In particular, the ERCs fail in some cases to *rehabilitate* the men and women who pass through them ... Their routine is quite demanding and in some ways it expects a level of determination on the part of *rehabilitatee* that would probably be high enough to get him back to work without a course of formal *rehabilitation*. (BNC FS6: 761)

What is of interest here is that the noun *rehabilitatee* is a hapax, thus demonstrating that a new *-ee* derivative for ‘one who is rehabilitated’ is composed from the prior verb *rehabilitate*. It is shown that when an *-ee* noun is created, the internal suffix *-ate* is generally removed, thus lending support to *-ate* truncation. This process is unique to *-ee* suffixation, since it does not occur in the related “process nominalization,” as evidenced in *rehabilitation*⁵.

2.2.4 Others (Pattern 4)

There are a set of *-ee* nouns that lack one of the definitional features of *-ee* nouns. They are divided into four classes depending on what feature is missing. The first class is one in which the underlying verb has to do with active voice and its subject represents the referent of *-ee* noun. Our BNC research has detected four related *-ee* word types: *absentee*, *escapee*, *returnee*, and *standee*. Interestingly, Barker (1998: 707, 718) points out that the referent of this kind of *-ee* noun performs an active action and the action is forced or constrained by an external force. Accordingly, *standee* is better interpreted as ‘one who is forced to stand’ (cf. *OED*). Nouns such as *escapee* and *returnee* can be analyzed similarly; the senses of *escapee* and *returnee* are, respectively, something like ‘one who is forced to escape (from an

³ Semantic dependence can be pointed out as a criterion for the direction of conversion from N to V: if defining one (*to net* ‘put into a net’) is difficult without using the other (*net*), then the latter is the source of conversion (Quirk et al., 1985: 1558-1559).

⁴ For more on how noun-to-verb conversion can generate a wealth of unlisted verbs, see Clark and Clark (1979).

⁵ There is a case where no base verb can be found that combines with *-ee* (cf. *appellee*, *donee*, *lessee*). This type of *-ee* nouns were borrowed entirely from French and are not considered to be formed in English soil.

oppressive or irksome condition)’ and ‘one who is ordered to return (from abroad to one’s native land).’ Consequently, the anomaly of this class is that the base verb is not in the passive voice.

The second one is devoid of passive sense. Three related word types are taken from BNC: *attendee*, *cohabitee*, *signee*. According to *OED*, *signee* is defined as ‘one who has signed a contract or register.’ In contrast to the corresponding agentive noun *signer* ‘one who signs,’ this *-ee* derivative emphasizes that a person is in a resulting state of an action. In addition, while *attender* ‘one who regularly goes to an event’ implies an intentional act, *attendee* foregrounds one’s state like ‘one who is at an event.’ *Cohabitee* can be treated in the same way as *attendee*.

The third class of atypical *-ee* noun is one whose referent is non-human. Five related word types are found in BNC: *adoptee*₂, *enhancee*, *investee*, *selectee*₂, *toucher*. *Adoptee* means not only ‘person who is adopted’ but also ‘thing which is adopted.’ The last class is one whose referent is a part of a body. The only relevant *-ee* noun, *amputee*₂, is discerned in BNC, which has the reading ‘one who has had a limb amputated’: what receives the action is a body part. The four classes of *-ee* derivations are unproductive because of the small number of related word types (cf. §2.2.1). It comes from the fact that each of them deviates from the prototypical *-ee* derivation and lacks one of its properties.

To summarize Section 2, based on objective data from a large corpus, we have presented evidence supporting conversion from noun to verb (§2.2.2) and the removal of *-ate* (§2.2.3), thus demonstrating that *-ee* is added only to verbs. As a result, we have confirmed that the meanings of *-ee* derivatives are divided into two categories: the meaning of ‘one who is V-ed’ and the meanings that partially deviate from that meaning. Furthermore, by quantifying the productivity of *-ee* derivatives (§2.1) and their subclasses (§2.2), we have confirmed that the former is the primary meaning, while the latter is secondary. It should be emphasized that this study is “corpus-based”: rather than simply listing various meanings of *-ee* derivatives collected from a large corpus, it

“elegantly” classifies them. In doing so, it focuses on “*-ee* hapaxes” and uses them importantly as a measure of productivity, for detecting context-dependent creative word formation, and for arguments in favor of conversion and the removal of an intraword suffix. Thus, this study exemplifies the significance of “corpus-based research” in morphological studies.

3 The Characteristics of Personal Substantivized Adjectives

The preceding section has been concerned with the characterization of the anti-agentive *-ee* nouns. There is another type of personal nominals involving a passive element, as shown in “cohabitees have more liberal views towards marriage and divorce than *marrieds*” (BNC CB8: 1868) and “the differences between the conquerors and *the conquered* began to disappear” (BNC B2V: 13). The italicized expressions are complex nominals signifying ‘people who receive an action,’ and neither of them actualizes a semantic unit representing people. This type of nominals always has a generic meaning and has traditionally been treated as a subtype of “conversion” from adjective to noun (Jespersen, 1949: 231-234). Poutsma (1914: 365-368) calls *the* + *Adj* formation “partial conversion,” to distinguish it from “total conversion” of *V-eds* formation⁶. While nouns like *marrieds* can rarely be used in the singular, expressions like *the conquered* require a definite article, cannot take the plural *-s*, and always refer to a class of referents (Quirk et al., 1985: 42). The latter are considered a type of anti-agentive nominals because they exhibit lexical properties: (i) coordinating with nouns (cf. *the interest of women, minorities and the disabled*) and (ii) conceptualizing people’s socioeconomic or physical conditions (cf. *the oppressed/the wounded*)⁷.

⁶ Regarding *the* + *Adj* nominals, there is also a syntactic analysis that assumes an empty noun after the participial adjective, rather than partial conversion (Kester, 1996).

⁷ Since only six related word types are discerned in BNC, *V-eds* formation is not productive (§2.2.1). The reason is that functional categories such as inflections and prepositions

generally do not appear inside words (cf. **[bookstore]_N*, **[in-city traffic]_N*), but *V-eds* nouns do not meet this restriction (cf. *[[[married]_N] s]_N*). As for *the* + *Adj* formation, it is currently difficult to comprehensively extract relevant expressions from BNC, making it difficult to quantify its productivity.

4 Japanese Anti-agentive Nominals

4.1 *hi-VN-sha*

We turn now to the characterization of anti-agentive nouns in Japanese⁸. A detailed observations of Japanese anti-agentive nouns using the method presented in Section 2 has revealed that they can be divided into four types. The characteristics of each type are described below. The prototypical type is *hi-VN-sha* noun, which consists of the passive prefix *hi-*, verbal noun (VN), and the person-denoting suffix *-sha*⁹. Accordingly, the noun *hi-koyoo-sha* has the same meaning as the *-ee* noun *employee*: ‘one who is employed.’ The internal structure is $[[hi-VN]_A -sha]_N$, where the prefix *hi-* is added to a verbal noun (VN) to form an adjective and then the suffix *-sha* is added to it to make a noun. VN is defined as a stem that has a predicate function and combines with the light verb *suru* ‘do’ to yield a complex verb, although the light verb is omitted within a word. Evidence that the *hi-VN* part is an adjective is that it allows the addition of the suffix *-sei* ‘-ity,’ which combines with adjectives.

Additionally, we can point out the constraints on the historical origin of the base. Japanese words are divided into three kinds: (i) native (Japanese) words, (ii) Sino-Japanese words, and (iii) words of foreign origin. The prefix *hi-* only attaches to Sino-Japanese VNs, and hence we can say *hi-koyoo-sha* ‘employee,’ but not **hi-yatoi-nin*, where the prefix is added to the native VN *yatoi*.

Turning to the communicative function of *hi-VN-sha* derivatives, they may serve as anaphoric substitutes, as is exemplified in (4).

- (4) *nyuushosha* *karano* *hiyoo-chooshuu*
resident from fee-collection-
joogengaku-o *hikiageru* *totomoni* ...
upper.limit-Acc raise and ... *hi-*
sochi-sha-chooshuu-gendo no ... pas-treat-
person-collection-limit of

‘in addition to raising the upper limit of fees that are collected from residents ... in the collection limit for **those who are so treated** ...’ (BCCWJ¹⁰)

The content of the VN *sochi* ‘treat’ in the italicized derivative is given in the first sentence. This propositional content is then incorporated to create the personal passive noun *hi-sochi-sha*. Importantly, the internal element of the derivative (*sochi*) can only be properly interpreted in relation to the preceding propositional expression¹¹.

Let us move on to examining the creativity of *hi-VN-sha* derivation. The productivity value is calculated, following the measure presented in section 2.1. Our BCCWJ survey has detected 91 word types of the nouns including 35 hapaxes, and so its productivity value is 0.385. The productivity shows the potentiality of creating 35 kinds of new words when 91 kinds of *hi-VN-sha* derivatives are used, and hence *hi-VN-sha* formation is very productive. Like *-ee* derivation in English, *hi-VN-sha* derivation is central to the formation of personal passive nouns in Japanese.

4.2 *VN-sha*

The second type of personal passive nouns consists of Sino-Japanese VN and *-sha*. Personal nouns in *VN-sha* can be divided into agentive (*shiyoo-sha*₁ ‘user’) and passive (*taiho-sha*₂ ‘arrestee’) nouns. The reason for this is that a Sino-Japanese VN combines with *suru* ‘do’ to make a verb, and when it is followed by *-sha*, *suru* is omitted, and further *-sha* itself is a personal suffix rather than an agentive suffix. Consequently, the personal passive nouns concerned are formally indistinguishable from agentive nouns.

In what circumstances are nouns in *-sha* used as personal passive nouns? There are two pragmatic determinants. First, *VN-sha*₂ nouns are acceptable when the agents of the verbal nouns are self-evident. For example, in the case of *taiho-sha*₂ ‘arrestee,’ the party carrying out arrests (the

⁸ While fragmentary observations have been made regarding anti-agentive nouns in Japanese (Kageyama, 2011; Morita et al., 1989), no systematic analysis has been conducted.

⁹ There are two endings for personal passive nouns based on Sino-Japanese stems: *-sha* and *-nin*. The former is more widely used, so *-sha* is used as the representative.

¹⁰ For collecting anti-agentive nominals in Japanese, I am indebted to the research engine of *Chunagon* (BCCWJ): <https://chunagon.ninjal.ac.jp/auth/login?service=https%3A>

[%2Fspring_cas_security_c heck](https://chunagon.ninjal.ac.jp/%2Fspring_cas_security_check).

¹¹ An anonymous reviewer asks if there is an English equivalent for *hi-sochi-sha* in (4). The English equivalent would be *treatee*, which is listed as “rare” in *OED* and not found in BNC. The verb *treat* is well known to require an adverbial (cf. *he treated John *(well/badly)*). Therefore, *treatee* ‘one who is treated’ is generally not acceptable, although it may be acceptable if some related adverbial is given or inferred from the context, as in (4)

police) is clear, so that the informational focus is on the party being detained. Under this circumstance, *taiho-sha*₂ is only given the anti-agentive sense. The same applies to, *kiso-sha*₂ ‘one who is accused.’ The second case is when common sense and world knowledge allow us to anticipate the “doing” side, while the informational focus is on the “done” side. In the case of *kyuujo-sha*₂ ‘rescued person,’ the rescuers are anticipated and emphasis in communication is placed on those who are rescued, as seen in news reports. Consequently, *kyuujo-sha*₂ is most likely to be interpreted as an anti-passive noun¹².

4.3 *yoo-VN-sha*

The third type of personal passive nouns is *yoo-VN-sha*, which is formally the second type with the prefix *yoo-* ‘required’ added. Thus, *yoo-kyuujo-sha* represents ‘one who is required to be rescued.’ The internal structure is $[[yoo-VN]_A -sha]_N$, of which the *yoo-VN(-no)* part is an adjectival passive, similar to *hi-VN(-no)*. This is illustrated by the BCCWJ compound *yoo-shien-hi-hoken-sha* (required-support-pas-insure-person, ‘those who are required to be supported and are insured’). In this compound, the adjectives *yoo-shien(-no)* and *hi-hoken(-no)* are coordinated. In principle, coordinates are syntactically of the same kind, which provides evidence that not only the adjective *hi-shien(-no)* but also the adjective *yoo-shien(-no)* implies the passive.

In terms of productivity, 14 word types of *yoo-VN-sha* nouns including 4 hapaxes are discerned in BCCWJ, giving its productivity value of 0.286. This indicates a considerable productivity of this derivation: one out of four word types in *yoo-VN-sha* used is a new word. The reason for the lower number of word types compared to *hi-VN-sha* derivation is likely due to the additional element *yoo-* ‘required,’ which limits their use.

4.4 *V-passive.suffix-nin*

Japanese also has personal passive nouns based on native verbs: *tora-wa-re-nin* (arrest-verb.suf-pas-person) and *kira-wa-re-mono* (hate-verb.suf-pas-person). The key point is that they contain

inflectional suffixes. For instance, in *tora-wa-re-nin*, the native verb stem *tora-* is followed by the verb inflection *-wa* and the passive inflection *-re*, and then the person-denoting suffix *-nin*. (*-Mono* is another suffix denoting person.) This contrasts sharply with Sino-Japanese-based personal passive nouns, which have no internal inflectional suffixes (cf. *taiho-sha* ‘arrestee’). *V-pas.suf-nin* nouns have connotations of scorn or mockery and therefore take the place of the compared complex nouns in such situations. Thus, *wara-wa-re-mono* (ridicule-verb.suf-pas-person ‘ridiculed person’) and *kira-wa-re-mono* (hate-verb.suf-pas-person ‘one who is hated’) are preferable to ??*hi-chooshoo-sha* (pas-ridicule-person) and ??*hi-keno-sha* (pas-hate-person)¹³.

5 Comparison of English and Japanese Anti-agentive Nominals

5.1 Empirical generalizations

Based on the observations made so far, this section identifies two major similarities and two key differences between English and Japanese anti-agentive nominals (§5.1) and then indicates their theoretical implications (§5.2). The first commonality concerns semantics. As laid out in section 2, only *-ee* derivation is productive in the formation of English personal passive nouns. Furthermore, among the four types of *-ee* derivation, Pattern 1 (e.g. *employee*) is the most productive core device and therefore the meaning ‘one who is V-ed’ is the compositional core meaning of *-ee* derivation. Patterns 2 and 3 (*patentee*, *evacuee*) basically have the same structure as Pattern 1. In comparison, each of the four subtypes of Pattern 4 conveys meaning that lacks part of the primary meaning (*standee*, *signee*, *enhancee*, *amputee*₂) (cf. §2.2.4). Because of their low productivity, each meaning can be considered a secondary meaning of *-ee* derivation. It is important to clearly distinguish between these two kinds of meanings. As discussed in section 4.1, the Japanese counterpart of Pattern 1 is *hi-VN-sha* derivation: it is highly productive and has the prototypical meaning mentioned above. In

¹² Note that, given an appropriate context, the *VN-sha*₂ noun in the second case (*kyuujo-sha*₂) can be used as an agentive noun. As for the productivity of *VN-sha*₂ derivation, it is difficult to apply the above productivity measure, so it is unclear whether the derivation is productive or not. One reason for the difficulty is that it is not always clear from the

context whether a given *VN-sha*₂ noun meets the above formation conditions.

¹³ Six word types of *V-pas.suf-nin* derivatives are detected in BCCWJ, so this derivation is not productive due to the restriction mentioned in Note 7.

Japanese, the other three patterns also express the central meaning, although the *yoo-VN-sha* pattern has the added meaning of ‘be required’ (e.g. *yoo-kyuujo-sha* ‘one who **is required** to be rescued’). Thus, in English and Japanese, the most productive anti-agentive nominalization provides a consistent core and compositional meaning.

Second, the main class of *-ee* derivation and its Japanese counterpart are highly creative (cf. §2.2.1, §4.1). The productivity values for each derivation are 0.306 and 0.385, respectively, implying that one out of three or four word types in *V-ee* or *hi-VN-sha* is a new word. The anaphoric and dynamic facets of the word composition are particularly noteworthy, as in examples (1) and (4), where an appropriate label can only be given to an entity by referring to a larger unit in discourse.

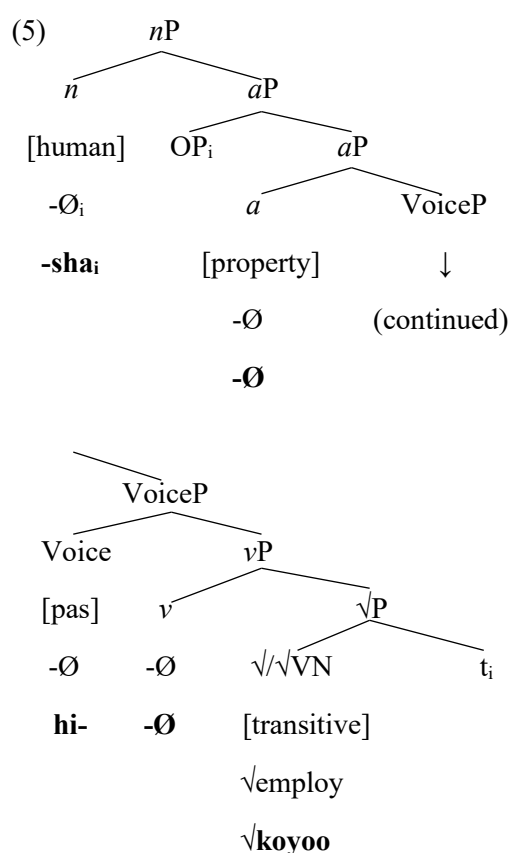
Let us move on the differences between English and Japanese personal passive derivatives. The observations in sections 2.2 and 4.1 identify a morphological difference between the main classes of personal passive nouns. Specifically, in the Japanese derivative *hi-koyoo-sha* ‘employee,’ the passive import is given to the prefix *hi-* and the person import is given to the suffix *-sha*. In its English equivalent, however, the suffix *-ee* represents ‘one who is V-ed,’ which involves the passive and personal elements. In addition, unlike English, Japanese has the isomorph *VN-sha* for both agentive (*tsuiseki-sha*₁ ‘pursuer’) and anti-agentive nouns (*taiho-sha*₂ ‘arrestee’), so that it has two kinds of anti-agentive nouns, *hi-koyoo-sha* and *taiho-sha*₂ (cf. §4.2).

5.2 Theoretical implications

The above generalizations have three theoretical implications. The first is what kind of theory is needed to grasp the core part of personal passive nominalization. *-Ee* and *hi-VN-sha* derivations can generate numerous nouns expressing a compositional meaning. Their generative and compositional nature is characteristic of sentence construction. It is therefore quite natural to capture this aspect in syntax. As for the morphological difference—unlike English, Japanese has a one-to-one correspondence between semantic units and morphemes—it is precisely what morphology should address: “fusion” is applied in English but not in Japanese, realizing the two semantic units of passive and person with a single morpheme, *-ee*.

The dual nature of the derivatives supports an “anti-lexical” model: the derivative’s core structure is constructed in syntax, which is then materialized in the post-syntactic morphology module (cf. Matushansky and Marantz, 2013).

The specific derivation process is shown below. Given that affixes are the heads of phrases, the underlying structure of the noun *employee* and the Japanese equivalent *hi-koyoo-sha* is constructed by the merging of a root ($\sqrt{\text{employ}}/\sqrt{\text{koyoo}}$) and category-defining heads like *v*, *a*, and *n*, as depicted in (5). The lexical items in bold are for the *hi-VN-sha* nominal and the non-bold lexical items are for the *-ee* nominal¹⁴.



¹⁴ Root ($\sqrt{}$) is defined as bound morpheme that becomes the core of a word and $-\emptyset$ signifies a bound zero-morpheme.

adjectival passive. Structure (5) signifies that the base verb is involved in the passive voice and its object expresses the referent of the entire derivative. Thus, the core meaning of both nominals, ‘one who is V-ed,’ is derived from a common underlying structure depicted in (5).

The syntactic output is then sent to the morphology module, where it is required to be constructed into a word form: the “boundness” nature of the inserted lexical items triggers morphological operations that combine them to form a word. A particularly important operation here is “fusion,” which is defined as the process of coalescing two adjacent terminal nodes dominated by a given higher node into a single terminal node (Calabrese 1996: 5). In English, $\text{Voice}_{[\text{pas}]}$ and $a_{[\text{prop}]}$ in (5), which are adjacent terminal nodes under the higher aP node, are coalesced into a single terminal node ($a_{[\text{prop}, \text{pas}]}$), and then the resulting a node and the n head ($n_{[\text{human}]}$) are fused into a single terminal node, $n_{[\text{human}, \text{prop}, \text{pas}]}$. Since the fused node is assigned a new set of features, a new lexical item is required. Vocabulary—a set of lexical items—contains the item $-ee$ with the features “[N][human][prop][pas],” which signify ‘ $-ee$ is a nominalizer representing a person involved in the adjectival passive.’ Accordingly, the morpheme $-ee$, retrieved from Vocabulary, is correctly inserted into the fused node. In Japanese, on the other hand, fusion is not applied to the nodes $\text{Voice}_{[\text{pas}]}$, $a_{[\text{prop}]}$, and $n_{[\text{human}]}$, so each is given an independent morpheme¹⁵.

Second, the peripheral aspect of $-ee$ nominalization may be associated with a kind of “grammaticalization,” which is generally defined as “the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status” (Heine et al., 1991a: 3). For example, the Ewe noun *meɓbé* expands from a noun referring to person’s back to a noun referring to the back of an object, and further to a spatial adverb (‘behind’) (Heine et al., 1991b: 163). Here we define it simply as extending the grammatical functions of a morpheme. Its prime motivation is to extend the use of an existing form in order to label a new concept, rather than creating a new label (Heine et

al., 1991a: 27). The formation of the expressions at issue is motivated by this. In the first peripheral $-ee$ noun, *standee*, the passive element of $-ee$ is not linked to the base verb but to a hidden causative verb. In order to conceptualize “one who is forced to stand” in contrast to “one who stands intentionally,” *standee* is formed by $-ee$ affixation, where the morpheme $-ee$ functions as a ‘forced’ marker. In the second noncentral *signee*, the ‘state’ import is extracted from the “stative passive,” while the passive element itself completely disappears. Consequently, the suffix $-ee$ comes to signify ‘one who is in a resulting state of an action.’ Thus, it functions as a ‘state’ marker in order to conceptualize “one who has signed” in contrast to “one who signs” (*signer*). In the third and fourth peripheral types, *enhancee* and *amputee*, the ‘person’ marker $-ee$ becomes grammaticalized to acquire the additional functions of an ‘object’ marker and a ‘body part’ marker. We can see that the extension of $-ee$ exhibits the nature of grammaticalization: $-ee$ extends the function by changing one of its properties while adhering to or retaining the remaining properties (“persistence,” Hopper and Traugott, 1993: 90)¹⁶.

And finally, the fact that Japanese has two kinds of anti-agentive nouns provides interesting data on an important morphological concept, “competition” in lexical insertion (cf. Rainer et al., 2019): why can only one of the synonyms, *hi-VN-sha* or *VN-sha*, be selected? Two pragmatic strategies are at work here, leading to the choice between them. The first one is the strategy of “Avoid ambiguity,” which prioritizes *hi-VN-sha* nouns over *VN-sha* nouns, since the former explicitly indicate a passive sense, whereas the latter can show both passive and active senses. Contrastively, in the two situations where *VN-sha* nouns are disambiguated (cf. §4.2), they win out over their more complex rivals, due to the “Avoid synonymy” strategy where a well-established simpler form preempts a complex semantic equivalent (Kiparsky, 1982: 13). Thus, two pragmatic strategies play a central role in the selection between two competing noun forms.

¹⁵ Marantz (1997) and Harley and Noyer (2000) are studies that support anti-lexicalism using examples of process event nominals.

¹⁶ This kind of grammaticalization does not occur in Japanese, and the reason for this seems closely related to the

English-Japanese morphological difference mentioned above: in Japanese, unlike English, the passive and personal elements basically are manifested separately, so their respective meanings are preserved.

6 Conclusion

In forming anti-agentive nominals in English and Japanese, it is significant to distinguish between core and peripheral parts. While the core patterns, Pattern 1 *-ee* derivation and *hi-VN-sha* derivation, exhibit the phrasal attributes of “free combination” and compositional meaning, there is a morphological difference between them in terms of the combination of semantic elements and morphemes. To capture this dual nature of the primary nominalization, an “anti-lexical” model would be required. Meanwhile, in English, four minor types of *-ee* derivation are recognized, each of which deviates in part from the prototypical pattern. It would be appropriate to address these deviations within a grammaticalization framework. Furthermore, in Japanese, there are two types of anti-agentive nouns, *hi-VN-sha* and *VN-sha*₂, and competition for selection occurs. Two pragmatic strategies are strongly involved in this process. We have revealed the characteristics of anti-agentive nominals in English and Japanese using large-scale corpora, and have proposed the word formation systems behind them. Further corpus-based exploration will be needed to elaborate on the proposed systems.

Acknowledgments

I would like to express my gratitude to three anonymous reviewers for their valuable comments and suggestions on an earlier draft of this paper.

References

- Harald R. Baayen and Rochell Lieber. 1991. Productivity and English derivation: a corpus-based study. *Linguistics*, 29:801-843.
- Harald R. Baayen and Antoinette Renouf. 1996. Chronicling *the Times*: productive lexical innovations in an English newspaper. *Language*, 72:69-96.
- Chris Barker. 1998. Episodic *-ee* in English: a thematic role constraint on new word formation. *Language*, 74:695-727.
- Benjamin Bruening. 2014. Word formation is syntactic: adjectival passives in English. *Natural Language & Linguistic Theory*, 32:363-422.
- Andrea Calabrese. 1996. Some remarks on the Latin case system and its development in Romance. ms., Harvard University.
- Eve V. Clark and Herbert H. Clark. 1979. When nouns surface as verbs. *Language*, 55:767-811.
- Heidi Harley and Ralf Noyer. 2000. Formal versus encyclopedic properties of vocabulary: evidence from nominalisations. In B. Peeters, editor. *The Lexicon-Encyclopedia Interface*, pages 349-374. Elsevier, Amsterdam.
- Bernd Heine, Ulrike Claudi, and Friederike Hünemeyer. 1991a. *Grammaticalization: A Conceptual Framework*. University of Chicago Press, Chicago.
- Bernd Heine, Ulrike Claudi, and Friederike Hünemeyer. 1991b. From cognition to grammar: evidence from African languages. In E. C. Traugott and B. Heine, editors. *Approach to Grammaticalization*, volume 1, pages 149-187. John Benjamins, Amsterdam.
- Paul J. Hopper and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge University Press, Cambridge.
- Rodney Huddleston and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge.
- Otto Jespersen. 1949. *A Modern English Grammar on Historical Principles*, volume 2. George Allen and Unwin, London.
- Taro Kageyama, editor. 2011. *Nichiei Taisho Meishi-no Imi-to Kobun* (Meanings and Constructions of Nouns: A Contrastive Study of Japanese and English). Taishukan, Tokyo.
- Ellen-Petra Kester. 1996. Adjectival inflection and the licensing of empty categories in DP. *Journal of Linguistics*, 32:57-78.
- Paul Kiparsky. 1982. Word-formation and the lexicon. *Proceedings of the 1982 Mid-America Linguistic Conference*, 1:26.
- Alec Marantz. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics*, 4:201-225.

- Hans Marchand. 1969. *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*, 2nd ed. C. H. Beck, München.
- Ora Matushansky and Alec Marantz, editors. 2013. *Distributed Morphology Today: Morphemes for Morris Halle*. MIT Press, Cambridge, MA.
- Yoshiyuki Morita, Shinjiro Muraki, and Masao Aizawa. 1989. *Kesu Sutadi Nihongo-no Goi* (Case Study: Japanese Vocabulary). Ofusha, Tokyo.
- Ingo Plag. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Mouton de Gruyter, Berlin.
- H. Poutsma. 1914. *A Grammar of Late Modern English*, part II, IA. P. Noordhoff, Groningen.
- Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, London.
- Franz Rainer, Francesco Gardani, Wolfgang U. Dressler, and Hans Christian Luschützky, editors. 2019. *Competition in Inflection and Word-Formation*. Springer, Cham, Switzerland.

