[REVIEW]

*Heterogeneity in Word-Formation Patterns:
A Corpus-Based Analysis of Suffixation with -ee
and Its Productivity in English*

By Susanne Mühleisen, John Benjamins Publishing Company, Amsterdam,
2010, xiii+245pp.

CHIAKI KOMATSU
*Ochanomizu University**

Keywords: *productivity, prototype theory, diachronic change, neologism*

1. Introduction

In her book entitled *Heterogeneity in Word-Formation Patterns*, Susanne Mühleisen gives an intensive and extensive analysis of a specific word formation process in English, the formation of nouns by suffixation with *-ee* such as *employee* and *trainee*. The *-ee* derivation as well as the related agentive noun derivation with *-er* has been much discussed in the morphological literature. The *-ee* derivation especially exhibits diverse syntactic and semantic properties and the diversity or heterogeneity provides great insight into the nature of word formation rules and the notion of productivity. In contrast to previous works which provide an analysis based on a single syntactic or semantic framework such as Bauer (1983, 1987, 1994), Barker (1998) and Lieber (2004), Mühleisen gives a thorough investigation of the properties and usage of *-ee* derivation from wide range and multiple perspectives: not only from syntactic and semantic perspectives but also from diachronic, cognitive and socio-linguistic perspectives.

The major claims of this book are threefold. (I) Prototype theory should be applied to the description of *-ee* derivation; there is a continuum between more prototypical derivation and less prototypical derivation. (II) In order to account for the heterogeneous pattern of *-ee* word formation, a synchronic as well as a diachronic point of view is necessary; diachronic changes that *-ee* derivation has undergone explain the reason for the present heterogeneity

* I would like to thank two anonymous reviewers for their valuable comments and suggestions. Needless to say, all remaining errors are my own.

of *-ee* word formation. (III) The methods of data collection and analysis by the use of linguistic corpora, especially the World Wide Web, provide a reliable database to measure the productivity of recent *-ee* word formations.

This book consists of six chapters plus a conclusion. Chapter 1 introduces some of the heterogeneous properties of *-ee* derivation and sets the aims and scope of the book. Chapter 2 examines phonological, syntactic, semantic and other various properties of *-ee* derivation and proposes that those properties are adequately described from a prototype-theoretic view which allows a range of variation from more prototypical derivations to less prototypical derivations. Chapter 3 explores the diachronic development of *-ee* derivation from its beginnings in the 14th century to the 19th century on the basis of a close analysis of Oxford English Dictionary (OED) data. Chapter 4 examines the trends and features of *-ee* neologisms in the 20th century on the basis of OED data supplemented with some data cited from Bauer (1983, 1987, 1994), Barker (1998), and other sources, after careful consideration of some of the important theoretical issues concerning morphology, including the definition and measurement of productivity. Chapter 5 analyzes more recent *-ee* neologisms in the late 1990s and early 2000s on the basis of data obtained by testing 1,000 potential new *-ee* words on the World Wide Web. Chapter 6 discusses the distribution of particular *-ee* words in different varieties of English by analyzing websites of American, British, Australian, New Zealand and Irish origin, and shows the regional heterogeneity as well as the continuing development of the word formation process under the influence of language contact in this globalized communication era.

In this review paper, I focus on the three major claims stated above and review the prototype-theoretic view on *-ee* derivation (in Section 2), the diachronic development of *-ee* derivation (in Section 3) and the corpus-based analysis of *-ee* neologisms (in Section 4). Section 5 concludes this paper.

2. A Prototype-Theoretic Approach

Mühleisen gives a detailed analysis of *-ee* derivation on the basis of quantitatively and qualitatively sufficient data which successfully reveals the heterogeneous pattern of *-ee* derivation. She advocates a comprehensive model of syntactic and semantic description which does not disregard deviant cases (discarded in earlier studies as exceptions), but rather sees them as less prototypical or marginal examples.

Mühleisen assumes that the following properties constitute the prototype

of *-ee* words: (i) it is based on a monotransitive verb with its direct object referring to the derived noun; (ii) it denotes a human being who is passively or involuntarily involved in the activity or event denoted by the base verb; (iii) it can be used in general contexts, i.e. not limited to a specific semantic field; (iv) it has a correlative agentive *-er* noun, which can be accounted for by the so-called “paradigm pressure” (cf. Bauer (2001)). In other words, the coinage of new *-ee* words is characteristically triggered by the presence of the corresponding *-er* word. One instance of prototypical *-ee* words is *interviewee*. This word is derived from a monotransitive verb *interview* and interpreted as ‘someone who is interviewed.’ The use of the word is not limited to a specific context, and it has a correlative *-er* word *interviewer*. While these properties are observed in most *-ee* words, Mühleisen argues that there are quite a few *-ee* words which lack one or more of these properties and assumes that they constitute less prototypical examples.

First, the syntactic category of the base to which *-ee* suffix is attached is not limited to a verb. Mühleisen points out that denominal derivation is also possible as in *festschriftee* ‘someone to whom a *festschrift* is dedicated’ and *missione* ‘someone who is susceptible to the arguments of a missionary.’ She also points out a number of examples which have neither a verb nor a noun base, such as *biographee* and *absentee*. The alleged base of *biographee* is **biograph* but there is no such word, and the base of *absentee* is apparently *absent* which is an adjective.

Second, the direct object interpretation is not the only interpretation that can be obtained from *-ee* nouns. Mühleisen refers to Bauer (1983) which classifies *-ee* nouns into the following four types: Direct Object (DO) type, Indirect Object (IO) type, Object of Preposition (OP) type and Subject (S) type. The DO type is the type that we have seen above and is the most common type. The IO type is based on a ditransitive verb with its indirect object referring to the derived noun, as in *promisee* derived from the ditransitive verb *promise* and interpreted as ‘someone to whom a promise is made.’ The OP type is based on a verb which takes a prepositional phrase with the object of the preposition referring to the derived noun, as in *experimentee* derived from the verb *experiment* taking the prepositional phrase headed by *on* and interpreted as ‘someone who is experimented on.’ The S type, which is rather marginal, consists of two subcategories: one type refers to the subject of an intransitive verb as in *evacuee* ‘someone who evacuates’ and the other type refers to the subject of a transitive verb as in *attendee* ‘someone who attends a meeting, conference, etc.’

Third, while most *-ee* nouns refer to human beings, *-ee* nouns with non-

human reference are also attested. For instance, Mühleisen gives examples of recent *-ee* words referring to animals or pets like *milkee* ‘one who is milked (farm animal)’ and *brushed* ‘one which is brushed (pet),’ and an example which refers to inanimate things as in (1).¹

- (1) However, **contaminees** such as soils and fly ashes from waste incinerators often contain a considerable amount of other chlorides, which may act as a main source of chlorine in the formation of PCDD/Fs via thermal processes. (p. 46)

Fourth, “lack of volition,” which is another common semantic feature of *-ee* words, does not apply to all *-ee* nouns. The lack of volition is originally proposed by Barker (1998) as a semantic constraint on *-ee* nouns and is meant to account for not only prototypical DO type (patient meaning) *-ee* nouns but also S type (agent meaning) *-ee* nouns like *standee*. A sense of affectedness or involuntariness is implied in this agentive noun; *standee* is not just ‘someone who stands’ but ‘someone who is compelled to stand (on a bus, etc).’ Mühleisen, however, demonstrates that this constraint does not apply to all S type *-ee* nouns, giving the following example found in her Web corpus:

- (2) “LEGAL DISCLAIMER: The stories, views and opinions expressed on this site are solely those of the **expressee**.” (p. 39)

Here in (2) above, *expressee* roughly means ‘one who expresses an opinion,’ where volitional control over expressing is implied.²

Lastly, *-ee* nouns characteristically have correlative *-er* nouns in an antonymic relation, as with an *-ee* noun for the patient and an *-er* noun for the agent of the activity denoted by each given verb in *employee/employer* and *advisee/adviser*, but this does not apply to all cases of *-ee* formations: *-ee* nouns which do not have correlative *-er* nouns exist as in *retiree*/**retirer* and *divorcee*/**divorcer*.

The prototype-theoretic analysis might adequately describe the heterogeneous pattern of *-ee* derivation at a synchronic level, but Mühleisen admits that the model itself does not account for why such heterogeneity is observed in this particular derivation. To explain the reason for the heterogeneity, she further explores the diachronic aspect of *-ee* derivation, which we

¹ Mühleisen notes that inanimate *-ee* nouns are often observed “in the technical and instrumental sphere” (p. 46).

² Mühleisen also points out that the word *expressee* is ambiguous and can be used with DO meaning.

will turn to in the next section.

3. Diachronic Development of *-ee* Derivation

In order to explore the diachronic development of *-ee* derivation, Mühleisen takes a close look at the lexical entries of the OED with their first attested use and examines the syntactic, semantic and other various features of *-ee* formations. For more recent *-ee* formations in the 20th century, she supplements OED data with data cited from Bauer (1983, 1987, 1994), Barker (1998) and various other sources. Providing and comparing the lists of new *-ee* word entries for each century, she discusses the diachronic development of *-ee* derivation from the following points of view: (a) the word class of base of derivation; (b) the type of passive noun; (c) the passive character; (d) the semantic field of occurrence; and (e) human reference.

First, Mühleisen points out some changes in the possible bases of *-ee* derivation. Although the contemporary *-ee* words are dominantly verb-based, she demonstrates that *-ee* formation originally started as noun-noun derivation: *-ee* words were primarily created to form a correlative patient noun to the agent *-or* noun. The earliest *-ee* words are French loan words, originally substantivized French past participles (*-ée* for feminine and *-é* for masculine) like *ordinee* (1330) 'ordained clergyman or minister' which was spelled as *ordiné* at first.³ Some of the earliest *-ee* words rest on the existence of correlative word pairs in Old French, as in *apelour* vs. *apelé* which were slightly changed to *apellor* vs. *appellee* (1387) in English. Soon after the appearance of such correlative noun-noun derivation, the formation of *-ee* words based on English verbs became possible as in *assignee* (1419) and *presentee* (1498), and the verb-based pattern hereafter became the dominant pattern of *-ee* formations. The productive use of the verb-based pattern, however, has not resulted in the complete abandonment of noun-noun derivation. Mühleisen gives relatively recent examples of noun-noun derivation as in *astrologee* (vs. *astrologer*) and *professee* (vs. *professor*). A further notable diversification occurred in the 20th century, where the number of new *-ee* entries dramatically increased: compound words and phrasal verbs entered the possible bases of *-ee* derivation. Mühleisen gives examples like *blindatee* and *moneylendee* for *-ee* words based on compounds, and *cracku-*

³ Hereafter, a four-digit number in brackets stands for the first year of the attestation with its meaning.

pee and *tryoutee* for those based on phrasal verbs.

Second, Mühleisen discusses the diversification of the possible types of *-ee* words. The earliest *-ee* words have a passive character and denote the direct object relation to the base verb, since they are modeled after French substantivized past participles. However, a few decades after the first appearance of such French-modeled DO type words, *-ee* words having the indirect or prepositional object relation to the base verb such as *vowee* (1380) 'one to whom a vow has been made' are added to the inventory of *-ee* words, and the number of such IO/OP type *-ee* words gradually increase during the subsequent centuries. In fact, IO/OP type formations outnumber DO type formations in the 15th and 16th centuries, indicating a diversion from the French borrowing tradition. DO type formations, however, gradually increase in the 17th and 18th centuries, and dramatically increase in the 19th century: DO type entries account for 19.1% of the total new *-ee* entries in the 17th century, 30% in the 18th century, 61.8% in the 19th century. As one of the possible factors for the increase of DO type formations, Mühleisen points out the influence of continued borrowings from French such as *protégé* (1778), *émigré* (1792) and *revolté* (1890).

Third, Mühleisen points out a change in the passive character of *-ee* type words: some *-ee* nouns early on had an agentive meaning rather than a passive meaning. In addition to DO and IO/OP type *-ee* words, S type *-ee* words such as *absentee* (1537) 'one who is absent' and *refugee* (1685) 'one who, owing to religious persecution or political troubles, seeks refuge in a foreign country' appeared and came into use. Mühleisen points out that these unusual formations are French borrowings (based on the verbs *s'absenter* and *se réfugier*), and suggests that those few deviant examples might have served as exemplars to yield a new or modified rule of word formation. In fact, the 20th century sees the increase of S type *-ee* words such as *pledgee* (1937), *retiree* (1945) and *arivee* (1989).

Fourth, Mühleisen points out the widening of semantic fields where *-ee* words can be used. As the earliest *-ee* nouns were borrowed from French legal terms, subsequent *-ee* nouns were restricted at first to the legal domain until the derivation was gradually incorporated into English as productive word-formation for wider semantic usages. The words listed for the 16th century were still largely limited to the legal or religious domain, but the semantic fields of *-ee* words widened as its productivity increased; the 19th century sees the dramatic increase of new entries, which include "new words that have a playful, humorous or ironic character—ironic in the sense that they mockingly transfer the legalistic character of *-ee* from its earlier uses

to the sphere of interpersonal relationships" (p. 81) as in *kissee*, *laughee* and *gazee*, as well as many non-humorous *-ee* words such as *addressee*, *advisee* and *employee* which are firmly established as common words in contemporary English.

Lastly, Mühleisen points out that *-ee* words started with human reference and expanded their meanings to include non-human reference. We can see this from the semantic development of the word *patentee*; the first attested meaning is 'person to whom a patent has been granted' (1442), followed by 'the inventor and proprietor of something' (1616), and then 'the invention of something' (1691). The 18th and 19th centuries produce several *-ee* words referring to animals or pets like *trainee* 'animal undergoing training' and *vivisectee* 'animal undergoing vivisection,' and the 20th century sees some inanimate *-ee* nouns like *benefactee* and *possessee* used as linguistic terms.

So far, we have seen the major diachronic changes that *-ee* derivation has undergone. Mühleisen argues that the accumulation of those diachronic changes has led to the present heterogeneity. This is a reasonable way of thinking. It is unquestionable that the vocabulary is the product of its history and incorporates various diachronic changes. However, I would like to suggest here that her diachronic findings can be analyzed from a slightly different point of view: the chronological order of appearance in different syntactic and semantic properties of the derivation accords with the extension of the properties of contemporary *-ee* derivation from the core to the periphery. The properties of *-ee* derivation observed at the earliest stage of development, such as DO meaning, human reference and correlation with *-er* words, correspond to the prototypical characteristics of contemporary derivation. In contrast to prototypical characteristics, less prototypical characteristics such as S meaning and compound base are quite recent phenomena. Is there any substantial relationship between the synchronic and diachronic developments of meaning and usage of *-ee* words? In relation to this, Mühleisen carefully distinguishes two types of productivity, 'diachronic productivity' and 'synchronic productivity,' and states that "synchronic productivity of a morphological operation cannot be regarded as completely separate from its diachronic version, as the production of new words at a given point in time is also based on the quality and quantity of previous formations" (p. 108). If there is any connection between synchronic and diachronic developments, the next question is how this can be explained in prototype theory or other general theories of linguistics. I believe that Mühleisen's fine-grained analysis of diachronic development of *-ee* words will open up new stimulating possibilities of further theoretical discussion if

it is properly combined with her synchronic study.

I also would like to remark that the concepts and ideas that Mühleisen discusses to account for language changes are very intriguing and worth exploring. She assumes that even a few of the creative neologisms introduced by analogical coining can lead to the establishment of a new productive word formation rule (cf. Botha (1968), Lehrer (1996) and Fischer (1998)). She states that “it is those actual words which deviate from the established patterns that are the most influential in changing the morphological patterns over time, since they are the basis of new words formed by analogy which might then result in rules” (p. 118). This accounts for the rise and development of IO/OP and S type *-ee* words. It should be noted here that the very issue of a dichotomy of analogy and rule has been controversial in the morphological literature (cf. Chomsky (1970), Jackendoff (1975), Bauer (2001)) as well as in the psycho- and neurolinguistic literature (cf. Rumelhart and McClelland (1986), Pinker and Prince (1994)). Specifically, it has been discussed whether the distinction between the two different modes of word creation can be clearly drawn, and whether the distinction, if any, cuts across inflectional morphology and derivational morphology. Although it needs more evidence and discussion to be verified, Mühleisen’s view on the interaction between analogy and rule provides an important perspective for future research.

4. A Corpus-Based Analysis of *-ee* Neologisms

As Mühleisen puts it, the OED is considered to be as “the best possible source for a diachronic reconstruction of vocabulary, especially for the earlier formations” (p. 65). The OED and other dictionaries of neologisms, however, do not provide enough data to investigate the potentiality of more recent word-formations because they lag behind current usage and do not necessarily reflect actual usages due to the selecting and editing processes these publications employ. The same is true of language corpora such as the British National Corpus which is generally used for linguistic research. Thus, for the investigation of more recent formations, the author examines the data from the Web, which she dubs as “the one large body of text collections which is unbeatable with regard to size, diversity of Englishes, varieties of users (i.e., here producers of text), text types and recentness” (p. 123). The use of the Web is not new (cf. Hundt et al. (2007), Kilgariff and Grefenstette (2003)), but the methods and procedure for the Web search that Mühleisen devises are very intriguing.

As the first step for the Web search, she creates a set of 1,000 potential search words by checking the OED and selecting those bases (mostly verbal bases) for which no *-ee* derivation has been found in the OED or any other common dictionaries and scholarly articles. The selection of base verbs is not conditioned by any syntactic, semantic or phonological features so as not to exclude any possibilities of new forms. The search words are alphabetically ordered and given in Appendix 2 of her book, which starts with *abandoned* and ends with *yellee*. The creation of search words is indispensable because unlike in coded corpora, a search for “*ee#” is not possible and a simple search for *ee* would produce an incredibly enormous amount of data due to the abundance of irrelevant words such as *leek*, *peer*, and so on.

Mühleisen checks each one of those potential words for its actual realization in any of the English language pages by using the common search engine, the Google query. Her research was conducted between February 2005 and June 2005 and this temporal limitation avoids one of the problems with the use of the Web: the Web is not a finite body of texts and differs from standard corpora in that sense. She also downloads a website for each search word and creates a reproducible corpus for further quantitative and qualitative analysis. She categorizes the search words according to the frequency of occurrence as i) Not found (no meaningful token found); ii) Hapax Legomenon (1–2 tokens); iii) Rare (3–20 tokens); iv) Established (21–100 tokens); v) Frequent (more than 100); and vi) Not quantifiable (many tokens found but not quantifiable because they are personal or brand names, French lexemes, and so on).

Some of her major findings can be summarized as follows. First, *-ee* derivation is highly productive or much more productive in contemporary English than previously assumed. Google produced 748 lexemes out of the 1,000 search words tested. Assuming that the productivity of a given derivation can be assessed by the ratio of potential words (her 1,000 previously unattested search words) and actual words (the words actually produced by the Web search), Mühleisen argues that such a high frequency of new *-ee* neologisms (almost 75%) indicates its high productivity.

Second, the co-occurrence of an *-er* word plays an important role in the formation of new *-ee* words. For the 748 new *-ee* words found, Mühleisen examines the possibility of co-occurrence with the correlative *-er* word in the same environment and finds that 552 items (73.8%) co-occurred with the corresponding *-er* word within one text on a website. Such a high frequency of co-occurrence with *-er* words is notably seen in higher frequent categories (Frequent: 94%; Established: 90%) in contrast to lower frequent

categories (Rare 79%; Hapax Legomena: 47%).

Third, almost all of the *-ee* neologisms in the Frequent category comply with the prototypical features of *-ee* words (i.e. verb-based, DO meaning, reference to animate and mostly human entity, denotation of a non-volitional participant in an event, and the use in general contexts).⁴ In contrast to the words in the Frequent category, those in the Hapax Legomena category exhibit syntactic and semantic diversity. Among the 246 neologisms in the Hapax Legomena (21.4% of the test words), she finds some *-ee* nouns which do not have possible verb bases like *astrologee* and *professee* and quite a few *-ee* nouns which are formed as part of a compound word as in *lap dancee* and *party ruinee*. She also finds S type nouns like *disrobee* ‘someone who disrobes,’ and *-ee* words with non-human reference as in (3).

- (3) The continuous removal of a solid ‘**erodee**’ is effected by gradual consumption of a solid ‘eroder’. In the process under study, the **erodee** melts and the (liquid) melt diffuses to the eroder, where an exothermic, indefinitely rapid, diffusion-controlled surface reaction occurs. (p. 160)

From these facts, Mühleisen argues that neologisms that have prototypical characteristics of *-ee* words are used more often and hence, become more easily established than those with marginal characteristics which tend to end up with nonce words or hapax legomena. This curious correspondence of frequency and prototypicalness remains unexplained.

Lastly, I would like to stress that Mühleisen’s study is notable in that the Web is effectively used to search for morphological objects, especially potential words or neologisms. The creation of 1,000 potential words and the statistic analysis of the data from the Web search, which must have been tremendously hard work, make it possible to successfully reveal the still continuing development of *-ee* formation. I believe that this method of data collection can be applied to the investigations of other derivational processes as well, and will be a new basis for future morphological researches.

5. Concluding Remarks

To sum up this review, I would like to stress again how detailed Mühleisen’s analyses are and how much data from various sources she ex-

⁴ Mühleisen admits that this may be due to a feature preconditioned by the selection of bases for the test words, which are mostly verbal.

amines to justify her claims. Although the scope of her research is limited to one particular English derivation, her synchronic and diachronic analyses of the derivation are very insightful and make us fully aware that morphology interfaces with syntax, semantics and the lexicon which is the product of history and can be influenced by social factors. While discussing the problems of the definition and measurement of productivity, Mühleisen also redirects our attention to the notions of analogical coining and neologisms which have often been neglected as marginal in the morphological theory. Also, use of the Web as a corpus has received increased attention in linguistic research over the last few years, and Mühleisen's carefully thought-out way of collecting and analyzing data sheds new light on the usefulness and potentiality of the Web. Due to the wide ranging and multiple perspectives that it offers, I believe that the book under review should interest many students and researchers working on morphology, lexicology, cognitive linguistics, historical linguistics and corpus linguistics.

REFERENCES

- Barker, Chris (1998) "Episodic *-ee* in English: A Thematic Role Constraint on New Word Formation," *Language* 74, 695–727.
- Bauer, Laurie (1983) *English Word Formation*, Cambridge University Press, Cambridge.
- Bauer, Laurie (1987) "*-ee* by Gum!" *American Speech* 68, 315–319.
- Bauer, Laurie (1994) *Watching English Change: An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century*, Longman, London.
- Bauer, Laurie (2001) *Morphological Productivity*, Cambridge University Press, Cambridge.
- Botha, Rudolph P. (1968) *The Function of the Lexicon in Transformational Generative Grammar*, Mouton, The Hague.
- Chomsky, Noam (1970) "Remarks on Nominalization," *Readings in English Transformational Grammar*, ed. by Roderick A. Jacobs and Peter S. Rosenbaum, 184–221, Ginn and Company, Waltham, MA.
- Fischer, Roswitha (1998) *Lexical Change in Present-Day English: A Corpus-Based Study of the Motivation, Institutionalization and Productivity of Creative Neologisms*, Gunter Narr Verlag, Tübingen.
- Hundt, Marianne, Caroline Biewer and Nadja Nesselhauf, eds. (2007) *Corpus-Linguistics and the Web*, Rodopi, Amsterdam.
- Jackendoff, Ray (1975) "Morphological and Semantic Regularities in the Lexicon," *Language* 51, 639–671.

- Kilgariff, Adam and Gregory Grefenstette (2003) "Introduction to the Special Issue on the Web as Corpus," *Computational Linguistics* 29, 333–347.
- Lehrer, Adrienne (1996) "Why Neologisms Are Important to Study," *Lexicology* 2, 63–73.
- Lieber, Rochelle (2004) *Morphology and Lexical Semantics*, Cambridge University Press, Cambridge.
- Pinker, Steven and Alan Prince (1994) "Regular and Irregular Morphology and the Psychological Status of Rules of Grammar," *The Reality of Linguistic Rules*, ed. by Susan D. Lima, Roberta L. Corrigan and Gregory K. Iverson, 321–351, John Benjamins, Amsterdam.
- Rumelhart, David E., and James L. McClelland (1986) "On Learning the Past Tenses of English Verbs," *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*, ed. by James L. McClelland, David E. Rumelhart and the PDP Research Group, 216–271, MIT Press, Cambridge, MA.

[received July 26 2011, revised and accepted January 7 2012]

1-34-4 Zoshigaya, Toshima-ku
Tokyo 171-0032
e-mail: bxr02340@nifty.com