Where do new words like "boobage, flamage, ownage come from?" Tracking the history of "-age" words from 1100 to 2000 in the "OED3"

Chris Smith

To cite this version:

Chris Smith. Where do new words like "boobage, flamage, ownage come from?" Tracking the history of "-age" words from 1100 to 2000 in the "OED3". Lexis. Journal in English Lexicology, 2017, Lexical and Semantic Neology in English, 12, [29 p.]. 10.4000/lexis.2167 . hal-02144382

HAL Id: hal-02144382
https://hal.science/hal-02144382
Submitted on 30 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Where do new words like boobage, flamage, ownage come from? Tracking the history of -age words from 1100 to 2000 in the OED3

Chris A. Smith
Where do new words like *boobage*, *flamage*, *ownage* come from?
Tracking the history of *-age* words from 1100 to 2000 in the *OED3*

Chris A. Smith

Introduction

This study aims to trace the evolution of nominal *-age* formation in the *OED3*, from its origins as a product of borrowing from Latin or French from 1100, to its status as an innovative internal derivation process. The suffix *-age* continues today to coin newly- or non-lexicalized forms such as *ownage*, *boobage*, *brushage*, suggestive of the continued productivity of a long-standing century-old suffix. This remarkable success appears to distinguish *-age* from similar Latinate suffixes such as *-ment* and *-ity* (see Gadde [1910]) and raises the question of the reasons behind this adaptability. The suffix *-age* has motivated a monograph study from Fleishman [1977], which takes a lexicographic socio-cultural or “integrated” approach to word formation and also later a corpus study by Palmer [2009]. The question of how suffixes found in loanwords become productive remains of consequence despite existing research into the question, such as: how and when does a suffix become productive? What criteria can be used to establish this (nonexistent base word in English, versus base word, i.e. transparency or analyzability of the derivative, phonological accommodation i.e. remodelling of the loan word)? More recently, corpus studies have flourished with the increased availability of diachronic corpora, evolving methods of diachronic analysis which embrace micro and macro evidence (Nevalainen & Traugott [2012]). In terms of methodology and statistical analysis, the issues of periodization and reliability of data have been shown to be instrumental in providing consistent verifiable data and analysis (Gries & Hilpert [2008], Allan [2012]).
Tracing the influence of Romance languages (Miller [1997]) and other contact languages on the English language has produced much research, both historically (in the 1900s) and today still (see Trudgill [2010] for a sociohistorical perspective, Durkin [2014] for a lexicographic overview). This continued interest shows that issues of language contact such as the process of external influences on English (Miller [2012a, 2012b]) is far from firmly understood and straightforward. This can be partly explained by shifting theoretical viewpoints on language change, and its processes, which shed new light on interpretations of historical data. The history of English has been studied from multiple perspectives: socio-cultural, bilingualism and code-switching, language contact (Nevalainen & Traugott [2012]). Recent methodologies tend to favour data-driven, corpus-based approaches, rather than anecdotal histories, although the focus of language change research tends to lie with macrostructures and word orders, centering on grammaticalization and subjectification, rather than morphology. Morphology and historical morphology (Gadde [1910], Dalton-Puffer [1994], Miller [1997], Anderson [2000], Lloyd [2005], Palmer [2009]) has also been challenged by the availability of wide-ranging data, thus paving the way for new methods and thinking influenced by language processing research, and probabilistic and statistical approaches, such as Baayen’s [2003] probabilistic morphology.

This data-driven lexicographic study focuses on collecting diachronic evidence regarding nominal -age derivatives in the OED3 to test prior results obtained by Fleishman [1977] and Palmer [2009]. The question at the heart of this paper is whether the continued trajectory of -age derivation evolves predictably based on historical patterns, in other words whether patterns of behavior are regular, or irregular.

To this purpose, this paper is organized as follows. Section 1 describes the corpus, and method for labelling and analyzing -age forms. It details the objective of the OED3 corpus analysis, and the factors used for assessing the morphosemantic structure of the 921 -age nouns generated. The data provide historical growth patterns, ratio of loans to non-loans, and comparison of N-age versus V-age productivity between 1150 and 2000. Mental productivity is arguably partially represented in the error forms and alterations which provide some clues regarding historical language processing. Section 2 provides a semantic classification of -age output based on key words relying on the analyzability of the senses provided by the OED3. The resulting semantic behaviour of -age forms displays considerable regularity and general absence of polysemy, coinciding with the knowledge that historical -age forms tend to be context-specific. This absence of semantic change is discussed, in contrast with two exceptions to this behaviour: polysemous ossified -age words, and outliers resulting from reanalysis. It is argued that exceptions provide anecdotal, yet crucial, insight into language processing and natural analogical change, otherwise absent in a historical corpus. Section 3 then focuses on the small output of lexicalized -age words in the 20th century, and then proceeds to discuss the productivity of non-lexicalized -age forms and their semantic behaviour using a contemporary corpus of English, English Web 2013, which is part of the TenTen family of large text corpora.
1. Data and method: the diachronic morpho-semantic analysis of -age nominal derivatives

1.1. Morphological productivity: innovation and analyzability

How does a loan suffix separate from the foreign non-transparent bases of loan words to become a productive suffix? In other words, how does a productive process emerge from the borrowed forms entering a language? The accepted answer (from language change, and language acquisitions studies) is frequency of usage, which brings about ability to recognize a pattern, i.e. cohesion within a paradigm (Miller [2010]) and extend or regularize that pattern to non-Latinate, or native bases. This form of innovation which relies on the analyzability of the loan words allows for reanalysis to be made. That means that the suffix -age emerges as a morpheme with a predictable effect on the base: it allows for extension of the model to bases of different kinds, particularly to native English, Norse and Germanic bases. For Palmer [2009, 2014], analyzability correlates directly with productivity. It is viewed as a pre-condition for productivity to occur (spontaneous new innovations).

However, to test this assumption, access to the mental lexicon of speakers in the past would be required, which is of course not realistic, especially as data are essentially limited to written data, and specific registers and genres. The unavailability of spontaneous oral data is a well-documented limitation of diachronic studies (see Allan [2012]).

Early signs of this innovation (see Bauer [2001]) are twofold. On the one hand, the existence of hybrids (Miller [1997]) constitutes a symptom of the increasing transparency, and therefore a precursor to the emergence of the morpheme. Secondly, the existence of error-forms can be interpreted as attempts to equate form with meaning in a context and so to reanalyze for increased iconicity or correspondence between form and meaning. Miller [1997] underlines that the application of Latinate suffixes to non-Latinate roots started earlier than previously thought: he counts at least 100 hybrid derivatives before 1450 in a literary corpus. Miller argues that colloquial ME allowed for a considerable amount of hybridity, which then made its way into literary texts by means of convergence.

1.2. Generating the corpus: exclusions and difficulties

This paper sets out to explore how a loan formation such as -age nouns becomes an internal pattern of neology by tracking their evolution over time. Using the OED3 to generate a corpus of 1233 -age words, a manual sorting of the forms carrying the -age nominal suffix filters over 200 homonyms and false positives. I excluded -age words which are verifiably transmission errors; smellage 1846 given as an alteration of smallage 1300 <small-ache>; lovage 1300 n1 given as being from the French lovesh, graffage 1798 deformation of graff-hedge. Also excluded are compounds of the noun age, words carrying the combining form -phage, as well as many remodelled forms such as the noun besage 1526 referring to a pair of saddle bags (from French besace, from bis-+sac), the noun cabbage 1391 from French caboche. In addition, transparent derivatives of pre-existing -age nouns as in marriage > re-marriage > mismarriage, etc. are also removed. On the other
hand, homonyms of existing -age words were kept if verifiably of a different base word. From the remaining 958 -age nominal derivatives, this empirical corpus-based study sets out to determine the diachronic behaviour of the target forms through a morpho-semantic lexicographic study.

A second filtering of 958 -age derivatives occurred during the classification phase: 921 -age output forms remain, excluding variants, errors and alterations, which will nonetheless be considered in the discussion of productivity and language change.

1.3. Classifying -age words in the OED3

The following subsection deals with the morpho-semantic labelling and classification of the filtered -age forms retrieved from the OED3.

1.3.1. The specificity of a data-driven lexicographic approach

As mentioned earlier, research into the history of loan suffixes from Latin and French is not novel (Gadde [1910] studies the history of -ery, -age and -ment, Fleishman [1977] focuses on the history of the suffix -age, Dalton-Puffer [1994] and Miller [1997] study the influence of French on ME morphology). Most recently, Durkin [2014] traces the history of loanwords. Most of the recent studies use diachronic corpus data such as the Helsinki corpus to study loans over certain key periods (Middle English, Early Modern English). While this approach has many benefits as underlined in Palmer [2009, 2014] including the availability of contextual elements, and the authenticity of usage, there are also some drawbacks, namely the lack of genre- or register-specific data.

The use of lexicographic material to trace word histories has also greatly improved, especially the use of the OED for tracking historical evolution. Despite the lack of available data for certain periods, Allan [2012] underlines the invaluable opportunity represented in the OED3 in terms of access to semantic, morphological, etymological data, and attestation dates (also see Smith [2016]). Allan [2012: 37] points out that the OED3 structures the senses of lexemes in order of attestation, although this cannot be taken to represent the development of senses, and that proper historical development must be reconstituted by the reader or researcher. In terms of methodology of tracking the histories of words, Allan [2012: 37] advises correlating data from derivationally related lexemes and etymologically related lexemes. Durkin [2012: 103-104] also strongly recommends paying close attention to detail and considering how individual word histories interact, reminding us “of the complex interplay of factors involved in any instance of lexical change”. Given the scale of the corpus, the methodology used here will be first a quantity-based analysis, and then will include a quality-based analysis of several individual word histories. Durkin [2014] extensively tracks the history of loanwords in English with data-driven analyses. Challenging corpus studies with lexicographic data now makes a lot of sense, as attestation dates (albeit approximative) do give some insight into macrolevel diachronic evolution.

The dates of attestation are those of the initial emergence of -age words. There are naturally limitations to the dates provided in the OED3, which are approximate and cannot be construed as definitive (see Allan [2012] on the question of attestation dates and chronology), although they provide acceptable data on a large time-frame scale. The OED3 also provides some indication of restriction of usage with terms such as rare, archaic,
obsolete except historical, feudal law, which provide some contextual clues to the usage of the term. Unfortunately, the frequency band provided only serves as a current frequency indicator rather than a diachronic indicator, and so cannot provide patterns of usage over time1. However, the indications of register are not systematic, or systematized. For instance, how is archaic different from obsolete, or from obsolete except historical? As Durkin [2012] himself underlines in his experience and capacity as an OED lexicographer, there remain many issues of methodology in terms of labelling, selection and classification of senses, as well as selection of context.

A fair proportion of more recent -age forms from the 20th century labelled in the OED3 as loans from French is specific to certain usages and contexts (such as massage and winery). This coincides with general knowledge regarding the motivation of recent loans, such as described in Ayto’s [1999] dictionary of new words of the 20th century for instance.

1.3.2. The structure and origin of -age words

Labelling the structure of the -age words requires analyzing or parsing the structure of the derivative by identifying the nature of the base word, even when the structure is a loan from Latin or French. Identifying the structure of native words is far easier when the derivative is transparent, however determining the initial structure within the donor language may help provide information on language processing at the time. The data in the OED3 allowed me to provide the following classification based on morphological decomposition of the OED3 entry words.

1) LOAN: this is the term used in the OED3, although not systematically. I have used this whenever the -age form is directly retrievable from a donor language (usually Latin or French). For example: coinage 1380 < coignage French>, carriage 1386, pontage 1325 “payment for use of a bridge” are historical loans from Old French or Anglo-Norman; rapportage 1903, sondage 1914, meritage 1989 are more recent loans from French.

2) LOAN / BLEND: this is a term I have added to label hybrid forms that are neither loans nor derivatives but associate both (usually a sign of analogical formation, and the beginning of independent formation). For example: clientelage 1660, verdage 1782, floriage 1775.

3) V-age: this is used when the -age form is verifiably deverbal based on the analysis of the attestation sense. For example: warpage 1863 < warp v. + -age>, and graftage 1895 < graft v. + -age>.

4) N-age: this is used when the -age form is verifiably denominal. For example: floodage 1862 in the sense “flooded state, inundation”, hulkage 1869 in the sense “hulks collectively”.

5) V-age or N-age: this is used when the word analysis allows for both denominal or deverbal derivation. For instance; taskage 1830 given as a nonce-word, i.e. a hapax, either has the sense collective tasks or tasking; buoyage 1855 can also be used in the sense “collective N” or “act of V”, or tankage 1866 can also have the sense “collective N” or “process of V”.

6) BLEND: this is used generally for later -age forms when they are analyzed as a combination of a prior -age word with another form. For example; scavenage 1878 < scavenge + -age> in the sense action of “scavenging”, or later haylage 1960 < hay + silage>.

7) Unknown: this is used when the OED3 offers no plausible origin hypothesis. These are usually rare or obsolete words with no context and uncertain meaning, such as average n3 1537 in the sense “breaking of corn fields”, breneage 1535 in the proposed sense
“burning”, and flobbage 1535 a rare Scottish usage for “phlegm”.

(8) Obscure: this follows OED3 labelling, and tends to be used when the OED3 offers several etymological pathways, with no clear preference. Examples are village 1386, garbage 1430, soutage 1532.

The following analysis will attempt to measure the proportion of V-age to N-age over time to detect if a change in patterns occurs in favour of one or other form. The question of semantic behaviour will also be considered. Is it possible to propose a relevant semantic classification of -age output forms, and what of the semantic specialization often cited in relation to -age forms?

1.3.3. Semantic categories using key features

Semantic typologies of -age words have been provided at length (Fleishman [1977]), including in the OED3 entry for the suffix -age, which provides historical information of its beginnings. A key word search in the entry definitions for each -age word provided key features, which were then classified according to larger conceptual categories, correlating with the nature of the base word of the derivative.

(1) ACT OF V, ACTION OF V, PROCESS OF V. This sense can be in competition with -ing nouns, as found in seasoning 1400 vs obsolete seasonage 1716. Note that metonymical extension can lead to the sense PLACE OF WORK, as in standage <V+age> 1777 “a space for standing”. This sense also appears in denominal -age forms like cooperage <N+age> 1714 “the place where a cooper’s trade is carried out”.

(2) COLLECTIVE N, AMOUNT OF N, QUANTITY OF, MEASURE OF (stealage 1769 “losses due to stealing”, wastage 1756 “loss by leaking, etc.”, roofage 1829 “roof-like covering, or the area covered”, wordage 1829 “words collectively”, plotlage 1910 “land in the form of plots; the area or value of such land”).

(3) STATUS OF N, CONDITION OF N, BEHAVIOUR OF N. This usage of -age can be in direct competition with -ery and -hood derivatives: the OED3 often gives these as synonyms, for instance coltage / colthood, adultage / adulthood, etc. Note that some rare forms such as havage 1799 in the sense “lineage, ancestry” do not have equivalent -hood derivatives.

(4) PAYMENT, TOLL, FEE, TAX, DUTY, RIGHTS. This historical usage has no derivational alternative but appears extremely frequent historically: riverage 1701 has the sense “toll for travelling on a river”, greenage 1763 has the sense “fee for using a bowling green”.

The methodological choice of using a feature analysis will be discussed in the section on the distribution of senses of -age forms.

1.4. Results and morphological observations

Using the classifiers listed above (i.e. decomposition and semantic features) in addition to the dates of attestation, this subsection aims to provide a quantitative analysis of -age words to detect patterns, as well as to verify pre-existing knowledge of a well-documented suffix. The overriding questions are: when does -age derivation become an internal process, and can preferential patterns be identified over periods of time?
1.4.1. Diachronic patterns of growth of -age output

The data confirm the patterns provided by previous research regarding historical lexical growth in the English language. The rate of -age output in terms of historical periods coincides with general patterns of growth (Durkin [2014], Nevalainen & Traugott [2012]), suggesting there is nothing atypical about the growth patterns. Figure 1 shows the raw output number per century: period 1 represents 1100-1200, period 2 represents 1200-1300, period 3 represents 1300-1400, and so on until period 8 representing 1900-2000.

The graph shows highs and lows of raw output per century, which means that these increases and decreases must be offset by the available data for each period. Nonetheless, the significance of the highs and lows can be compared with overall knowledge surrounding word production at different times in the history of English. There is a steady increase overall, with a dip in -age output for period 6, which corresponds to 1700-1800. This can be accounted for by the prescriptivist movement calling for regulation and reform in late Middle English (Swift [1712] aiming to correct and improve the chaotic English language).

Figure 1: -age output per century 1150-2000

Graph 1 also shows the proportion of non-loans for each period. The number of non-loans per period increases slightly more than the total output, from 1300 onward, with the gap between the two lines increasing between 1300 and 1500, reaching a maximum around 1350. This suggests that internal formation began to outperform loans in that period. The lines then tend to run parallel, indicating a levelling of the output form. Then, a reduction in the gap between -age output and loan output occurs for the period 1600-1700, suggesting a renewed tendency to import loans from French or Latin. The data for the most recent period are less accurate as the total output falls considerably for the period 1900-2000, rendering any interpretation of the gap between loan output and non-loan output questionable. In addition, further observation suggests that -age formation does in fact continue today, contrary to the appearances based on the number of...
lexicalized -age words in the OED3 for 1900-2000. This matter will be discussed in section 3.

1.4.2. Ratio of loan words to non-loans: early signs of morphological productivity

The proportion of loans to the non-loan words overall regardless of period shows that the number of derivatives by far outweighs the total number of loans. Out of 921 -age nouns, the total number of loans is 331. In addition, the total number of -age componential forms (i.e. derivatives) is 585. The total number of what I have termed blend words is 13: these blend words can be analyzed as the combination and clipping of two retrievable base words or, occasionally, a hybrid of a base word and the -age suffix. Historical forms such as the obsolete verdage 1775 referring to “green herbage” [verd- (as in verdure) + -age suffix; introduced by Marshall], and obsolete floriage 1782 [< Latin flōr-, flōs flower, after foliage] are analyzed as hybrid analogical forms in so far as they blend a Latinate loan with a native base word. Sense 2 of rummage referring to a loud noise 1575 is analyzed as a combination of the [French loan arrimage + Scottish rummish]. On the other hand, blends such as haylage 1960 [hay+ilage], ecotage [ecological-sabotage], backage 1887 [back+frontage], septage 1977 [septic-sewage] combine two base words. Finally, the total number of unknown or obscure words is 6. The OED3 gives soutage, village, garbage as obscure, in that no indisputable etymological path can be determined. This typology provides evidence that -age words display the ability to form transparent derivatives from a base word rather than merely resulting from borrowed word forms, in other words, they display morphological productivity.

Figure 2: Proportion of LOANS to -age derivatives overall

The proportion of loans to -age forms over the period 1150-2000 favours transparent structure. However, the period 1150-1400 has quasi 100% loan output, except for potential transparent structures such as hidage 1195 <hide+age> in the sense PAYMENT FOR N, baronage 1300 <baron+age> in the sense STATUS OF N, hermitage 1290 <hermit+age> in the sense PLACE OF N. The first supposed blend loan (or analogical form) occurs around 1300; costage 1325 <costage Anglo-Norman> COLLECTIVE N, rummage 1406, clientelage 1660 <loan or blend clientele / clientela+age>. By 1500, the proportion of transparent derivatives starts to climb, which will be quantified in the following section devoted to the trajectories of N-age and V-age output.
This evidence coincides with findings concerning the historical evolution of English, as the production of internal derivatives is thought to be dated from 1400 (Palmer [2009], Durkin [2014]). What remains to be seen is whether the structure of the forms is stable, or if there is a change in morphosemantic pattern over time. This will be tackled in the following subsection.

1.4.3. Proportion of V-age to N-age over time

Out of the 585 -age forms, 295 are labelled as denominal N-age; 189 are deverbal V-age. In addition, no fewer than 101 are either deverbal or denominal, or a combination of both, depending on the sense. The proportion of non-N or non-V bases is very small, although not insignificant, mostly composed of lexicalized -age words such as Adj-age shortage 1868, savage 1250, and Adv-age outage 1851. Figure 3 shows the disproportionate percentage of N-age, which represent 50% of total output.

Figure 3: Proportion of V, N or V / N / Adj-age derivatives

The progression of V-age vs N-age forms is shown below in Figure 4; the first potential perceived derivative hidage is attested in 1195. The graph represents the accrued number of N-age and V-age output over time, using 50-year time slices. The progression lines are very similar, with N-age outperforming V-age consistently, as visible in the parallel growth trajectories. This suggests that the productivity of V-age and N-age remains relatively constant, with N-age outperforming V-age at the same rate. Although V-age is attested earliest, N-age has a constantly larger output than V-age. Two peaks evidencing periods of increased growth appear roughly around the same period for both V-age and N-age. First around 1500 and then again 1600 for N-age, whereas V-age forms peak slightly earlier 1450 but do not reach the levels of output of N-age which increases at a steadier growth. The highest period of increase for N-age is 1650-1750, which accounts for the greater discrepancy in the following period between N-age output and V-age output. The graph shows no sudden or distinctive change of pattern of output, which can be interpreted as an indicator of morphological stability.
The decision to label -age output as V-age or N-age can be a complex task, as in some cases both analyses are possible. However, there may be a predictable correlation between type of base and semantic behaviour. In the V-age category, verbs of action are often used in a specific sense to form technical -age nouns. Take *runnage* 1742 ACT OF V which has the specific sense of “smuggling”. The noun *havage* 1799 is limited to the sense of ancestry: it is a synonym of *lineage*, *ancestry*. As for *breakage* 1775, the noun refers to the ACT / PROCESS initially, then the sense shifts in 1848 to the PRODUCT / RESULT OF V, in other words the effected object of the verb. With this sense PRODUCT / RESULT, *breakage* can in fact be used in different registers such as music or nautical contexts. This suggests -age output is dependent on the selected sense of the base word, rather than semantic change of the -age noun itself.

Let us here note that another form exists: the irregular *brockage* 1879 which is given by the *OED3* as the clipped form or past participle broken + -age. *Brockage* 1879 refers to a specific subtype of *breakage*, that of damage to a coin, suggesting reference to an industrial coinage process. An additional, more general sense appears some 10 years later (i.e. concomitant for all intents and purposes) in a Scots usage to refer to anything broken. This sense dating from 1888 makes it a near-dialectal synonym of *breakage* 1842.

In the N-age category, a similar observation can be made. For instance, *riverage* 1701 (obsolete sense) refers to the toll for using a river, synonym of prior *rivage* 1598 perceived as a Latin loan (*rivagium*), or 1749 a collective sense for any stretch or river where mooring is possible; here a synonym of prior French loan *rivage* 1330.

In the following section, I aim to provide evidence of the semantic behaviour -age forms:

1. To what extent do -age derivatives follow the pattern established by loan word senses?
2. To what extent do the senses follow from the sense of the base?
3. To what extent do -age words show potential for semantic change?

Let us now turn our attention to hybrid forms which are evidence of remodelling in English.
1.4.4. Hybrids: language internal remodelling within English

As indicated above out of 921 -age nouns 331 are loans, 585 derivatives, 13 blend or hybrid forms and 6 of unknown origin (i.e. unclear or debatable) namely: village, garbage, soutage 1532 in the sense “coarse cloth”, cabbage n3, average n2, and flottage 1535 in the sense “phegm”. I should note at this point that I have been conservative with the list of unknowns; when the OED3 offers an interpretation, the form goes into the reanalysis category. For example, lobage is classified as N-age although the OED3 indicates a probable misinterpretation of louange. I have also not included shewage given as a reanalysis alteration of scavage from the base shew. However, distinguishing blends from derivatives can be uncertain, and there is some potential overlap. Five -age forms combine a Latinate base with the -age suffix, without proof of an existing Latin or French base, making them hybrid savant blends (for want of a better word). The following hybrids of a Latinate base and the -age suffix all refer to collective nouns: fructage 1650 <fruct- + -age>, lactage 1753 <lact- + -age>, sylvage 1773 <sylv- + -age>, verdage 1775 <verd- + -age>, florage 1782 <flor- + -age >. A potential 6th, postliminage 1661 a rare synonym of postliminy, is either a blend of Latin postliminium + age or postliminy + age, and refers to a legal right in Roman law. A further 10 -age words appear to involve hybrid influences of native forms, leading to semantic transfer: rummage 1486, scavenage 1878, spoilage 1806 <spoliate + age> ACT OF V, clientelage 1660 <cliente + age>, brockage 1887 <broke + age> RESULT OF V, backage 1887 <back + frontage>, proprietage 1827 <proprietor + age> COLLECTIVE N, haylage 1960 <hay + silage>, septage 1977 <septic + sewage>, ecotage 1971 <eco / ecological + sabotage>. Of this list, the recent blend ecotage is no longer recognizable as an -age form, which suggests that it can be disregarded as a direct -age form. As most recent blends, it reads transparently as a blend of the French loan sabotage 1910 <saboter + age> with the adjective ecological. The phonological structure of the ending sets it apart from the regular phonological assimilation of -age ending from a calque of the French /tːʒ/ to weak /tɒʒ/. Evidence of phonological confusion is visible in the semantic change of rummage. The noun rummage 1486 is a loan from the French arrimage as exhibited in sense 1: “The arranging or rearranging of cargo (esp. casks, etc.), in the hold of a ship”. However, the influence of the verb rummish linked to Scots (north-east) reemish, remish, reemage “to search thoroughly, to poke around and stir things up in searching” transfers to the sense 2 of the noun rummage attested in 1575 in the sense “loud noise, commotion”. Interestingly, the verb rummage 1544 is viewed as a conversion from the noun rummage 1486, thereby coming full circle from the influence of the verbal Scots form reemish. Some -age words of unknown origin are no less intriguing. Garbage is a common -age noun still in use, and yet its origins appear opaque (as is the primary sense attributed to garbage as “offal”). No unequivocal base word is retrievable either from a native base or a Latinate and French base. However, the OED3 suggests the word is adopted from Anglo-Norman, and related to French garb “sheaf”. The sense of garbage has evolved from the initial sense of “offal / entrails” (1482) to the more general sense of “refuse” (1582). This semantic extension could suggest some confusion with the noun garble (1503) indicated by the OED3 also in the sense “refuse, extraneous matter”. Of course, this semantic evolution is not incompatible with standard semantic change by extension since refuse is what is not desired, in the same way as offal represents what is not desired. Garbage in the sense first attested in 1592 is a figurative sense of non-material things which have no value such as...
“literary garbage”, and garbage attested in 1526 has the obsolete sense of gerbée, a dialect French term for horse food.

The following section investigates the semantic evolution of -age forms and the impact of such change or lack of change on the status of -age output and productivity. I seek to answer the questions posed in the section on historical output: (1) To what extent do the -age derivatives follow the pattern established by loan word senses? (2) To what extent do the senses follow from the sense of the base? (3) To what extent do -age words show potential for change?

2. Morpho-semantic patterns of -age nouns

2.1. The evolution of the semantic features of -age nouns

The semantic typology of -age words is provided in 1.3.3. and corresponds to well-known and well-documented studies of -age words, including the OED3’s own entry for the -age suffix which summarizes the history of the form.

(1) ACT OF V, ACTION OF V, PROCESS OF V; (2) COLLECTIVE N, AMOUNT OF N, QUANTITY OF, MEASURE OF N; (3) STATUS OF N, CONDITION OF N, BEHAVIOUR OF N; (4) PAYMENT, TOLL, FEE, TAX, DUTY, RIGHTS FOR USING N.

The semantic distribution of vage words is difficult to pin down using key words in the OED3 definitions, although not impossible. Despite the necessary approximations associated with using the key terms listed, an overview of the distribution of senses attributed to the 998 -age forms is proposed here. As can be seen, they do coincide to some extent with the categories listed above.

Figure 5: Semantic feature analysis of -age words

Figure 5 shows that reducing -age to several categories using repeated key words is more difficult than predicted. Only around a third of -age words fit into these categories, while another third (346 words) do not fit into the main categories listed above. There are three plausible explanations for this; first the categories chosen are too restrictive to allow for adequate classification into large types. If this is true, regrouping the outliers (subtypes) into the larger types should solve the issue. Secondly, OED3 entries can be inconsistent,
since there is no uniformity on a large scale lexicographic corpus that is still in the process of being updated. The third explanation is that -age words lack regularity, transparency, and that the remaining third of -age words are lexicalized and ossified so that they cannot be decomposed easily.

At this stage, some questions can be answered, albeit partially as further study would be required to systematize.

1. Is there semantic regularity of -age output? The multiple categories suggest that semantic regularity is difficult to pin down using a feature analysis.

2. Is there a correlation between denominal and deverbal structures and a semantic category? Yes, to a certain extent, as long as the base word is not multicategorical, in which case dual interpretation remains plausible.

3. Is there a change in semantic behaviour over time of persisting -age forms? Yes, the disappearance in the 20th century of the PAYMENT / TAX / TOLL sense is the one major distinction. This disappearance may be a remotivation or reanalysis based on socio-cultural factors. The last -age word classified as having a PAYMENT sense is millage 1891 “The rate of taxation in mills per dollar to which a given area, group, etc., is liable” in US usage.

What other factors may account for semantic change or semantic permanence? In order to investigate further, the next subsection focuses on patterns of semantic change in -age words.

2.2. Semantic change in -age words

2.2.1. Lexical stability and specialization of -age words

There is little diachronic semantic change observable across the 921 -age forms in the corpus. Most words appear remarkably constant with one or two senses and they exhibit little polysemey. This may however be a sign of lexicalization as argued in Lipka [2002], also see Lipka et al. [2004]. The absence of multiple lexicalized meanings does not necessarily equate with monosemey. It does suggest that usage tends to be restricted to a context (which has been established elsewhere) and does not favour normal semantic change (Koch [2012]) such as metonymy and metaphor. Only a handful of -age forms display subsenses: rattage is first attested 1807 and then 1878 (see below); rummage is first attested 1486 in a now obsolete sense, and then 1575 in the sense “loud commotion” in Scottish usage, and again 1598 in the sense “rubbish, junk”; garbage is first attested in 1430 and later 1582; ownage is attested in 1576 in a rare usage in the sense “ownership”, reappearing again in 1998 in the sense “act of defeating”; damage, first attested 1300, has the most subsenses including the PAYMENT sense attested in 1430 as a legal term (in the plural) and then in 1755 as a slang term in the sense “cost / expense”; lastly courage 1300 also has multiple subsenses. For most of these, the semantic change that is observable seems to stem entirely from the context of usage.

Take the term rattage, which is given two distinct senses, both directly linked to a difference in base word rather than natural change. Rattage is first attested in 1807 as a deverbal V-age “nonce-word” in the sense “percentage of people who rat”; the following sense rattage attested 1878 is analyzed as N-age in the sense “loss, damage caused by rats” in US usage.
Amongst the -age nouns with more than one sense, a small proportion have become longstanding conventionalized -age words (although most are labelled as obscure origin words by the OED3), which are discussed in the following section (courage, damage, baggage).

2.2.2. Tracking some conventionalized -age words

In this section I consider the semantic ossification of some centuries-old -age forms still in use today. These appear to be the very antithesis of neologisms. The words remain frequent today and are used in standard discourse, compared to other -age forms which tend to be limited to historical practices (feudal law) and technical usage (nautical, industrial, and more recently specific terminology related to massage or wine).

It can be argued that such early -age forms have become conventionalized. Some early forms such as courage 1300, village 1386, damage 1300, garbage 1430 are hardly perceivable as -age derivatives at all. Note that garbage and village are given as being of obscure origin in the OED3, so their non-analyzability stems from their very emergence in English. Given their longstanding existence and continued usage, tracking their semantic history could give some insight into historical practices and usage.

Take the noun village attested in 1386. Despite multiple senses identified in the OED3, all appear stable, with no unpredictable changes of meaning; the definition ranges from historical, to adapt to more modern habitations (Greenwich village, New York), and then extends metonymically to the population. Notably one of the subsenses listed is a derogatory Cambridge University slang term 1864 for “a disreputable suburb of Barnwell”.

Courage 1300, from Anglo-Norman from French <cor + -age>, refers to N (French noun <coeur> as in “heart”) collectively, specifically the heart as the “seat of emotions”, although the transparency of this analysis is debatable. The subsenses are quite numerous, showing semantic change based on the interpretation of emotion (from anger, lust to bravery 1382), but also possibly involving other etymological pathways and natural semantic change such as metonymy and metaphor (see the history of anger by Geeraerts et al. [2012]). It is notable that one of the senses attested 1545 is metonymically derived with the collective sense of “brave individuals”, triggering the sense COLLECTIVE attributed to the -age suffix.

Damage 1300, a loan from Old French damage (1100), shows a more straightforward trajectory of change; from the initial sense “loss” (1300), to “trouble / inconvenience” (1398), to “misfortune / regret” (1385), to the legal sense (1430) which has remained in use today. A slang use of the term corresponding to the sense “expense, loss”, in other words monetary loss, i.e. specialization of sense in terms of cost, actually dates back to 1755. This new sense could be motivated by the triggering of the PAYMENT sense often attached to -age words, or possibly an alteration of the pre-existing legal sense “loss, damages” attested in 1430. The triggering of this sense is therefore in line with both a context and the potential for meaning of -age suffix (as originating in the Latin usage of the suffix).

Nonetheless, the change evidenced through individual institutionalized forms appears to be dependent on sociocultural norms and changes.
2.3. Outliers: error words, alterations as evidence of cognitive processing

From the 1233 generated -age final syllable words in the *OED3*, 958 were found to be carriers of the -age morpheme (either through borrowing or derivation). However, a second filtering based on etymological information identified spelling variants, so called alterations, deformations, or even errors. Many of these word alterations appear to stem from phonological or phonotactic reanalysis, which is significant for the study of productivity. Reanalysis may be considered a symptom of language processing, allowing for reinterpretations based on perceived iconicity. This suggests, as mentioned earlier, that errors can be perceived as clues of the mental processes of historical speakers, therefore giving some insight into the mental lexicon of past time. In so doing this signals a form of productivity, the ability to form spontaneous -age words based on analogy.

*Billage* is attested in 1627 as originating from the noun *bilge* 1522, possibly via phonological assimilation. *Billage* refers to “the bottom of a ship’s hull, or that part on either side of the keel which has more a horizontal than a perpendicular direction, and upon which the ship would rest if aground; also, the lowest internal part of the hull.”

*Graffage* is given as dialect word attested in 1798 as a blend of *graff* N “grave” and *hedge* N “Graffage... a wooden frame somewhat like a Stile, placed in a bank, where there is a water-course”.

*Careage* attested in 1704 is given as originating from *caruage* 1610 “ploughing” from <Old Northern French *caruage*, in modern French *charruage*).

*Stintage* is labelled in *OED3* as a “spurious” misreading of *stinting* 1642 “a portion of the common meadow set apart for the use of one person” and later 1827 *stint+age* “allotment of stints”.

*Cabbage* n2 in the sense “lair, den” is given as probably an analogical alteration based on analogy with -age words of the noun *cabin* in the obsolete sense “a natural cave or grotto; the den or hole of a wild beast”.

*Curtilage* 1330 is a legal or formal term, given as originating from Anglo-Norman *curtilage*, <Old French *cor-*, *courtilage* (medieval Latin *cor-*, *curtilagium*) interpreted as a reanalysis *court+lodge* according to the *OED3* via popular etymology in the 17th century, i.e. through remodelling. *Curtilage* refers to “a small court, yard, garth, or piece of ground attached to a dwelling-house, and forming one enclosure with it, or so regarded by the law; the area attached to and containing a dwelling-house and its out-buildings”.

*Eatage* 1842 is labelled as a deverbal -age form, of northern dialect usage, from *V* *<eat* + -age> but appears comparable with the senses of *eddish* OE, with which *eatage* may have been confused according to the *OED3*. The first attestation of eatage has the sense “grass available only for grazing; esp. the aftermath, or growth after the hay is cut. Also, with some defining word, as after-, spring, winter”. A further sense attested in 1843 appears to be modelled on Latin -age words referring to payment and rights, in this case the right to use such pasture. This signals a reinterpretation based on analogy with other regular -age forms. *Eddish* OE is identified as being of obscure etymology in the *OED3*. “Usually identified with Old English *edisc* park or enclosed pasture (glossed vivarium)”. The *OED3* also comments on the unpredictable change of meaning from the original sense attested around 700 to the modern sense of *eddish* 1468 “grass which grows again, an aftergrowth after mowing”. In fact, *eatage* appears to share its meaning with the latter sense of *eddish* (a stubble field), signalling a remodelling by analogy.
The entry for lovage is labelled as a loan from French luvash itself stemming from classical Latin levisticum (4th cent.; also as levisticus, libestica, libisticum, libusticum, livisticum, lubestica, luvesticum, from the 11th cent. in British and German sources), “probably an alteration of classical Latin ligusticum, apparently denoting the same plant” according to the OED3. The OED3 also offers the following explanation for the potential remodelling of nach into -age. “The Anglo-Norman forms in -ache, -asche probably show folk-etymological alteration after ache, asche, ache n2 referring to plant names. There seems no compelling reason for assuming (as is often done) that the English word itself shows folk-etymological alteration after love n1 and ache n2, although the ending of the word may show remodelling after words in -age suffix”.

Another example in the corpus of interaction with Anglo-Norman and Old French ache referring to a variety of parsley can be found in the noun smallage <small adj. + ache n2> attested in 1300 in the sense “Any of several kinds of celery or parsley; esp. wild celery, Apium graveolens, formerly used medicinally and to flavour food”.

Following Hay and Baayen [2002] and Bybee [2007] on analogical productivity, I posit that these cases of reanalysis confirm evidence of potential productivity. Historical reanalysis or remodelling via phonological and semantic assimilation can be interpreted as processing changes that occurred in the mental lexicon of speakers at the time. If this theory holds true, they offer insight into potential productivity from a diachronic perspective.

2.4. Preliminary conclusion: need to assess productivity through non-lexicalized forms

Identifying semantic behaviour based on dictionary senses is not a straightforward task, especially with a considerable number of forms to analyze. Limitations associated with the decision to enter or exclude the forms into the dictionary, and the assessment of meaning via accessible occurrences make it difficult to establish an indisputable semantic pattern. Once lexicalized, forms tend to become conventionalized, and in many instances, although impossible to quantify due to the inconsistency of such tags, they are limited in register (slang, technical) or variety (Scots, Manx); some are obsolete and rare (a large proportion of words in -age have not stood the test of time). Many -age words are restricted in their usage to the reference to certain practices, such as feudal systems historically, and more recently to wine, massage, technical, nautical uses. There is evidence that -age derivatives remain highly analyzable, therefore susceptible to reanalysis based on contextual cues. A range of potential semantic behaviour can be realized depending on the conditions of usage. This is the very definition of productive word forms, i.e. non-institutionalized ossified meaning. To summarize the conclusions of this section, here are the answers to the questions posed at the beginning:

1. To what extent do the -age derivatives follow the pattern established by loan word senses? Patterns appear to remain stable, there is little semantic change overall, although it is true that many of these -age forms are rare or obsolete, which can of course be a major factor in explaining the absence of change.

2. To what extent do the senses follow from the sense of the base? There does appear to be a predictable effect between base and derivative, i.e. a pattern that is internalized.

3. To what extent do -age words show potential for change? The evidence suggests that -age words which do remain in usage are adaptable to socio-cultural practices.
The following section considers the future predictions of productivity and semantic change in -age words, providing a counterpart to purely lexicalized forms and meanings.

3. Predictions of productivity and current -age words

3.1. Productivity measures

Before describing the predictions of productivity, let us consider a brief overview of the main principles of productivity and the means of assessing it. Bauer [2001] defines productivity as a type of innovation different from creativity in terms of its predictability. However, as underlined by Bauer and others, how productivity is defined can vary significantly, leading to some confusion in approaches. There are different types of productivity, although in this case it is morphological productivity that is intended, i.e. in the case of derivation the ability of a bound morpheme to be added to a base to create a derivative, following a pattern of constraints. However, the complexity surrounding morphological productivity goes beyond the ability to form new words based on a pattern:

1) PRINCIPLE 1: Productivity is scalar, not absolute (Bauer [2001], Baayen [2009]). It is now established that there is no binary system of productive versus non-productive. For instance, there is no such thing as zero productivity, unless it refers to the impossibility of forming a word. Overall, productivity is best measured on a scale of high to low productivity.

2) PRINCIPLE 2: Following principle 1, it is possible to measure the degree of productivity quantitatively.

3) PRINCIPLE 3: There are several types of productivity: 1) realized productivity, i.e. the number of word forms existing and following a pattern, 2) expanding productivity (the number of new words being formed, and 3) potential productivity, i.e. the prediction of creation of new forms - also referred to as category-conditioned productivity measured by calculating the proportion of hapaxes to the total number of tokens containing the affix (Baayen [2009]). This means that productivity can be measured with respect to three different factors. Baayen [1992, 1994] proposes a now well-established quantitative measure of productivity based on frequency ratios of type-token. This method has been perfected by Hay and Baayen [2002, 2003: 14] who argue that the ratio of the frequency of the base compared with the frequency of the derivative is a better indicator of morphological productivity.

4) PRINCIPLE 4: on a qualitative level, decomposability, or parsing, is directly correlated to rates of productivity, from a cognitive language processing viewpoint (Hay and Baayen [2002]).

5) PRINCIPLE 5: Productivity is fundamentally motivated in terms of frequency of repetition and pattern-formation via analogical reanalysis. Bybee [2007] explains that the existence of frequent words ending in the same suffix produces a pattern of semantic regularity that the speaker can analyze and apply to new forms. This suggests that productivity is essentially a factor of analogy in the mental lexicon, which translates to the lexicon to a smaller degree (Bybee [2007], Hay and Baayen [2002]. Also see Baayen [2014] for an overview of the benefits of a psycholinguistic experimental approach to derivation). Indeed, lexicalized forms exhibit ossification, and therefore lack analyzability or transparency, which contrasts with productivity.
No single factor is responsible for productivity. Instead productivity is a multifactorial equation producing probabilistic measures. The following sections consider what data are available to assess productivity from a diachronic perspective when information on users’ production and processing is unavailable (i.e. no access to the mental lexicon of historical language users).

3.2. Semantic change in -age words: case studies

The data provide few recent lexicalized -age words in the OED. However, this does not suggest that productivity is decreasing, given that lexicalization is a historical process. Some interesting semantic features can be observed in 20th century -age words such as flamage 1983 (computing slang) referring to “Vitriolic argument or ranting, esp. via email or in postings to a newsgroup; the action or practice of sending inflammatory, emotive, or abusive email messages”. This definition would now correspond to the practice of trolling. The sense of flamage does not correspond to the usual types of meaning associated with V-age or N-age as in this case the meaning of the base is to be taken metaphorically, from the verb flame [1377] in its 1981 sense, as in “to produce inflammatory reactions”.

Ownage has two distinct attestation dates: 1576 in the (now rare) sense “ownership, the fact of owning”, and 1998 in the sense “act of defeating”. The time lag between the two attestation dates is rather remarkable. Both forms can be parsed as V-age although the sense of the verb own differs. This calls into question whether the later form has any relationship at all with the initial form, suggesting rather that these are two disconnected derivatives based on different attestations of the verb own. Rather than polysemous change deriving from the noun ownage 1576, this is the result of two distinct processes of -age affixation.

Similarly, the deverbal noun stickage produces two distinct senses which do not appear directly related but rather a product of separate deverbal affixation from two different senses of the verb stick. Stickage is attested in 1647 in the obsolete rare sense “hesitation, reluctance, delay” and can be directly correlated to the sense stick 1535 “stop in perplexity, be unable to progress, find an answer”. On the other hand, stickage attested in 1726 in the sense “tendency to jam” can be correlated with the sense of stick 1635 “to become lodged, mired, unable to move, jammed”.

These observations show that multiple derivatives appear to be frequent, resulting from natural semantic change in the base word rather than in the derivative. This confirms that the derivative remains stable, i.e. that the pattern remains retrievable.

3.3. Current development of -age formations

Quantifying and retrieving non-lexicalized forms depends on several factors. The main factor is the source of data, which excludes lexicographic material. A corpus of contemporary recent English will be the most helpful in identifying novel -age forms, including rare usage and nonce-words as they signal ability to access and produce (Baayen [2003, 2009], Bauer [2001]). The corpus of choice is a large internet-based corpus, which will provide a sample of natural non-literary language. This type of data is well adapted to finding neologisms (see the NeoCrawler in Kerremans et al. [2012]).
This signals a strong intuitive language internal productivity; for instance, some old -age forms, such as snappage 1602 (N-age based on the noun snap, a slang word for thief) are used in contemporary English with a different analysis. Snappage today, although not listed in the OED3 possibly because the entry has not been updated, is in use as evidenced by a corpus search. A search in EEBO (Early English Books Online) historical corpus for snappage generates 6 hits, from two different sources dating 1602 and 1608. In all six contexts, the sense of snapping conforms to OED3 definitions and dates.

Snappage occurs four times in the OEC (Oxford English Corpus) in the sense ACT OF snapping: in 3 occurrences, specifiers of the snapping are provided within close proximity, as in handlebar snappage, cassette snappage, harness snappage, and one use of metaphorical snappage in the sense of inability to deal with a situation. In English Web 2013, 53 hits are found, although some of them are false hits corresponding to the proper name SnapPage: they also show snappage used in at least three main senses: 1) instance of snapping as in a photograph; 2a) and 2b) instance of snapping (as in wire, ligaments); 3) instance of snapping as in psychological wear and tear.

(1) to my surprise Rebecca brought out her Polaroid instamatic and started to snap away. She loves Polaroid shots – the gritty, unfinished feel of them – and she assured us that she had tons of film for her retro snappage [...].
(2a) I told her, and you won't hear a peep of complaint out of me. But I am not – I repeat NOT – strapping on those slats of death and shame and ligament snappage while I'm there [...].
(2b) He slept through most of the night, that I'm aware of, but got me up at four, wanting to go outside. We have been sitting here drinking my coffee, him at my feet, ever since. I did try to manipulate his leg, with a resultant furious squirmage and snappage, and he hurt himself trying to get away from me, so that wasn't a success, in any way [...].
(3) And Mrs Loomis is easy enough to explain coming from all of that followed by her son getting killed. A little snappage after all that would be understandable [...].

The considerable difference in the results from one corpus to the next demonstrates that context and discourse type are key to accounting for usage and interpretation. This context-dependence is also indicative of a high productivity, in the sense that snappage can be used potentially in all the structures of V-age or N-age as long as the context enables processing. This usage is synonymous with snapping, which could be used in all the contexts, but which would have a different tone or connotation. I would argue that -ing nominals continue to bear a strong verbal behaviour (with the exception of highly lexicalized nouns such as drawing, painting), whereas -age nouns behave like fully-fledged members of the noun category. It appears that -age nouns behave with a higher degree of “nouniness” as Wierzbicka [1988] might say. Deverbal -age nouns take on a quality that -ing deverbals do not: they become particularly susceptible to quantification. In context 3) for example, the use of squirmage and snappage instead of squirming and snapping gives the reference an atemporal and agentless sense, referring merely to the resulting state rather than the event. The subject of squirming and snapping is not inherent as it is with -ing forms that are viewed as nominalizations of events. This creates a discrepancy between the event and its quantifiable results, possibly interpreted with comical effect. Further study would be required to investigate the effects of deverbal nominal derivation using -ing and -age morphemes.

The pragmatics of -age words would be worth investigating to determine the versatility of usage depending on context (reminiscent of the semantic behavior of
substantivized adjectives, which also adapt to contexts, creating a strong sense of pragmatic adaptability, see Smith [2005]).

3.4. Using web corpora to track innovations in non-formal discourse

As suggested by Kerresmans et al. [2012], tracking neologisms is best achieved by using a web crawler, which enables the authors to devise a tool to track neologisms, called NeoCrawler. The use of a web crawler is more suited to finding neologisms due to two main factors; first, it will be a reflection of contemporary Internet usage, and secondly, it will be large enough to reflect small changes, and new usages. This makes a web crawler a far more efficient tool than more limited corpora such as the BNC, which reflects a comparably small sample of usage from 1980s and 1990s.

A search in English Web 2013 for instance, which counts 19 billion words, is more likely to provide an accurate portrait of contemporary usage. To test this, a search for hourage in the BNC provides no results, whereas English Web 2013 offers six occurrences. Although six hits in a corpus of several billion words is a needle in a haystack, these uses can be perceived as innovations. Usage in (4) is explicitly questioned as being innovative via the use of an insert, and the usage of hourage in (5) is analogical and modelled explicitly on mileage. Neither of the occurrences is technical or semantically ambiguous.

(4) Okay, I felt good. Most mornings (at least in recent years) when I wake up it is not by choice, and I am a little short of the full hourage (is that a word? well if not, it’s 9 words now! ha!) that I require to function with reasonable amounts of grace and a positive outlook on life [...].
(5) The problem is that treadmills don’t work so well for distance targets – you picks yer speed, and off you go. So I’ll have to build up a ton of hourage, rather than mileage [...].

A query for the noun flamage, which is listed in the OED3 as computer slang, shows no fewer than 67 hits, which is considerably more than hourage, but the frequency remains minute compared to the overall size of the corpus. Perhaps it is not entirely unpredictable that there is a use for a term referring to online overreactions. A collocate search for the significant co-collocates show the specialization of flamage for this practice in (6); spam, bashing, chatter and profanity all apply to online behaviour. The results also show the occurrence of another sense, that of flames considered collectively. In (7b) the use of quotation marks indicates a specialized use for flames in firefighter speak, whereas in (7a) the use of flamage seems to be motivated by hyperbole. In (6) the sense is compatible with an analysis of flamage as V-age, whereas in (7) flamage is compatible with N-age.

(6) Although there’s powerful moderation (forum owners can withdraw messages and even eject members), it’s rarely used: because usernames are fixed and messages traceable, people are generally very responsible, and there’s no spam and very little flamage [...].
(7a) The carpenter came by and cleaned up the box for the stove and I’ve got the gas line hooked up and have flamage! I’m looking forward to cooking something hot for breakfast, it’s been too long without one [...].
(7b) Example 1. At about 9:45PM EST a 911 call was made from “Short Stop” on highway 11 (behind Seymour Johnson Air Force Base) that the Hard Ware store adjacent to them was demonstrating “flamage” [...].
ownage to breakage ratio is getting worse all the time. After months of research and wedding gift cards burning holes in our marital pockets, we purchased a French press [...].

(8b) a tenkeyless mechanical keyboard in a simple, elegant form factor, perfect for portability and total ownage at tournaments. Stripped of its numpad, the Razer BlackWidow Tournament layout is light and highly compact making it perfect for gamers with limited desk space [...].

Ownage finds the most hits in English Web 2013, with 244 tokens. In (8a) ownage analyzed as V-age refers to the total amount of things owned (OBJECT OF V) but in most cases ownage refers to a metaphorical sense of the verb own, i.e. to win (ACT OF V) (8b) refers specifically to the practice of competitive video game battles.

3.5. Competition with -ing nominal forms: specificity of -age derivatives

Competition between V-ing and V-age is mentioned explicitly in OED3 entries for several word forms, such as streetage / streeting used as a synonym for the layout of streets, ramblage 1883 as a more specific type of rambling (particular to Manx law), sinkage 1883 and sinking. Some alternations of V-ing and V-age exhibit potentially humorous differences in processing. Take for instance the noun screwage attested in 1865 and labelled as rare in the OED3 and compare with screwing. The OED3 gives screwage in the sense “act of applying political pressure” as N-age rather than V-age ACT OF V. The noun screw 1404 is attested before the verb, but both predate the derivative screwage 1865. This means that there is opportunity for semantic transfer to screwage, especially since the verb screw is particularly rich in metaphorical extensions of meaning, each based on a particular feature, manner or aim of the act of using a screw. The sense “compress, exert pressure” of the verb screw is attested 1617, well before the derivative screwage, suggesting that this analysis is not impossible.

A search in English Web 2013 for the noun screwage provides 60 results, providing evidence that the word is by no means obsolete in contemporary English. However, the sense of screwage is variable, and some occurrences show screwage is used as the name of a game as in (9a). This usage seems to rely on the sense of be screwed meaning “to lose badly”. In some occurrences (9b), screwage refers to the fee paid at a restaurant for bringing a bottle of wine and having to use a cork screw to open the bottle. This usage is highly metonymical in that the reference to screw is closely indexed to the situation or the FRAME.

(9a) So we can have games that involve direct conflict but are not wargames per se. Sometimes that direct conflict involves violence (as in the MMO), sometimes not (as in the economic or business game). Sometimes these are what I call “screwage games”. These games for from three to many players are usually directly competitive but do not require a lot of reasoning for success, games that involve a strong dose of chance as well as skill [...].

(9b) Typically, we will order a bottle of white or bubbly to start an evening, and ask the somm/waitperson to cork our bottle. Often, the screwage fee is waived if we buy the first bottle [...].
In comparison with screwage, the noun screwing appears to be more straightforward as referring to an ACT OF V based on some of the commonly used senses of the verb: messing up and having sex with.

The parallelism or potential synonymy between V-ing and V-age can be quite explicit as in the use of both flamage and flaming in (7c). One of the main differences in usage is the measurability of the notion described by flamage in contrast with flaming which is limited to the description of an event.

(7c) Flaming is universal, but different systems handle it in different ways. Both the technology and cultural norms matter. On Usenet, for instance, most news reader applications support a feature known as a “killfile,” which allows an individual to screen out postings by a particular user or on a particular subject. It is also sometimes referred to as “the bozo filter.” This spares the user who is sufficiently sophisticated from further flamage [...].

From the above contrasts (screwage / screwing, flamage / flaming), it appears that V-age forms and N-age forms contain a strong concrete semantic potential of adaptability to context. They appear to be heavily context-dependent, or context-limited. -age deverbal and denominal derivation is particularly intriguing as the base can often be multicategorial, like flame, leading to dual and contextual interpretation and analysis of the -age noun. The deverbal derivative in particular has strong event nominalization features, by reference to the event or via metonymical contiguity to its participants, including effected objects, i.e. “that which is Ved / The quantity of Ved”. This distinguishes -age nouns from -ing forms which appear to remain more notional (compare flaming ACTION OF V versus flamage RESULTS OR PHYSICAL, TANGIBLE EFFECTS OF V). It can be argued there is a far higher degree of embodied meaning in -age derivatives than in -ing nominals. Of course, further systematic comparative study would be required to test this theory, as indicated above in the analysis of snappage and snapping. Rather than becoming lexicalized and ossified, -age derivatives tend to remain grounded in usage and practices, making them more pliable to semantic adaptation based on the context of usage, and specifically register.

Take the noun rakeage (1851) given as obsolete rare with the sense “that which is raked together”, i.e the object of raking, or the total amount of leaves raked. The context given by the OED3 is the following single occurrence taken from H. Mayhew, London Labour II, where both scrapeage and rakeage occur in a coordinated structure as objects of the verb remove.

(10) engaged in removing the scrapeage and rakeage... from the surface of the streets.

A search in English Web 2013 for rakeage returns one false hit, however scrapeage returns two hits, of which one context is fully comprehensible. In (10b), the term refers to poor guitar playing, producing a scraping sound. The verb rake is far less widespread than the verb scrape, which can explain the discrepancy in usage of V-age derivatives.

(10b) That's good for just having started playing. I was expecting all manner of scrapeage, but I was pleasantly surprised [...].

Scrapeage, without the letter <e> from the base, can also be found in OEC in (10c), although only one result is found. In the occurrence below, taken from a 2008 personal blog,
scrapage refers to the practice of scraping a drawing using preprinted self-adhesive patterns called screentones.

(10c) I love letratone! Look at this unremarkable drawing and how a bit of letratone with clever scrapage really makes it into a nice, proper illustration [...].

These two examples show that some -age words labelled as obsolete in the OED3 may not actually be obsolete at all. Even though the frequency of usage is rare, the occurrence of scrapeage is by no means unclear or difficult to comprehend in either occurrence given the context. In fact, it is arguably quite the opposite, the sense is easily retrievable and accessible, leading to contextual interpretation. This is arguably one of the most convincing signs there is of the productivity of -age forms that allows for mental processing and adaptable usage of -age nouns (following Principle 5).

Summary and concluding comments

This data-focused study analyzed morphosemantic patterns of 921 -age nouns in the OED3 over a period ranging from 1150 to 2000. The objective was to establish a better understanding of the productivity of a loan suffix which has stood the test of time from medieval Latin to current contemporary English, and which has been studied previously essentially from a corpus perspective as in Palmer [2009]. The continuing productivity of -age formation begs the question; how does a loan suffix become an internal suffix, by what language processes (are they different to standard processes of change?), and how does the suffix behave over a large scale diachronic timeline? The questions of morphological stability or change, the question of semantic restriction, of the possibility of normal lexical change or the absence of polysemy have been the guiding objectives of this study.

I set out to answer these questions using the data collected from the OED3:

(1) Do novel -age forms correspond to a continued pattern or do they exhibit new behaviour?
(2) Does the suffix -age hold a special place amongst borrowed Latinate or French suffixes?

The methodology selected provided some answers which coincide with generally established studies of loans and derivatives, as well as diachronic periods of the English language. Loan output outperforms derivatives only until 1500, from this period onwards, V-age and N-age derivatives account for proportionally more new -age words than others (blends and loans). N-age takes precedence over V-age consistently over the period from 1500-2000. Over the course of one thousand years, there is considerable stability, shown through steady growth of -age output. Two periods of more intense growth coincide with already established periods of increased neologism: 1450-1600 and 1750-1800.

On a semantic level, the senses attributed to -age output remain stable, which sets them apart from forms which are more liable to general language change such as metonymy and metaphor. Rather than undergoing semantic change, -age forms tend to be very dependent on the senses of the base word, which in turn account for new sense of -age output. It can therefore be argued that new senses do not derive from senses of existing -age nouns but instead derive from semantic changes pertaining to the base. This in turn is evidence of the continued stability and transparency of -age output. The continued
transparency itself is a symptom of continued productivity; the ability to analyze and model is the basis for the creation of new forms. The semantic specificity of -age words does not make them good candidates for semantic change, however multiple derivatives based on different senses of the base appear common, thereby suggesting that -age forms tend to remain transparent rather than becoming conventionalized, ossified or institutionalized.

Current neologisms generated from a corpus of contemporary Internet English confirm that despite the high frequency of a small number of conventionalized -age words, the usage of rarer types continues, restricted to certain usages, registers and contexts. Affixation with -age appears to have a strong internalized capacity for production based on the practices and uses specific to certain populations at certain times (for instance hourage and minutage are used for specific circumstances where the measure of time in hours or minutes is relevant to a practice or process). This can be interpreted as high compatibility with an embodied perspective, much like the use of adjectives as nouns such as whites, greens (which can refer to a range of objects, such as respectively people, clothes, laundry, outfits, parts of an egg, and dollars, vegetables, army or nurse uniforms, etc.). It seems neologisms may be more finely indexed to usage and practices than conventional labels, and novel -age words seem to correlate with this behaviour, making them more adaptable. Contrary to -ity and -ment, the form seems to have increased hybridity / adaptation, making it less recognizable as a foreign or French import. Phonological change may have played a crucial role in the adaptation and success of -age derivation in English.

BIBLIOGRAPHY


**Dictionary and corpus references**


NOTES

1. The genres that compose diachronic data (such as the EEBO and The Penn Parse Historical Corpora) are mostly semi-oral discourse in the form of plays, and correspondence, as well as historical records and to a lesser extent fiction.

2. This -age / hedge remodelling notably brought further reinterpretation of the loan word *curtilage* 1292 originally from Anglo-Norman *curtilage*, but later reanalysed by speakers in the 17th century as *court-ledge* or *court-lodge*.

3. This is a matter for future research, using a historical corpus such as the Penn Historical Corpus. Tracking the usage of some -age words over time in a large-scale historical corpus would provide invaluable information of frequency of usage, including the usage of rare -age forms.

4. The OED3 actually uses the term badly to assess the non-standard structure.

5. There is anecdotal evidence in the data that the assimilation of -age to /ɪʤ/ can lead to a resemblance, and occasionally a confusion with other words of similar meaning and carrying a phonetically similar ending, such as -ish. This type of remodelling is well-known and could in part account for the success of the -age ending.

6. Many thanks to one of the reviewers who suggested that the question of -age forms could be tested by analysing the etymology of the combined root morpheme. Etymological pathways could explain semantic developments.

7. Possibly formed via analogy with the preexisting noun *cribbage* 1641, referring to a game of cards.

8. Compare with the preexisting *corkage* 1838 referring to the charge applied by waiters or staff to uncork a bottle not supplied by themselves.

ABSTRACTS

This diachronic lexicographic study aims to analyze the morpho-semantic behaviour of -age forms in the OED3. The objective is to provide evidence of the diachronic processes which enabled a loan form to become an independent productive pattern of derivation in English. Using the OED3 as a corpus, a list of all the words ending in -age were generated and then filtered to exclude all those that did not undoubtedly carry the -age ending. This filtering removed many false results, such as compounds carrying the noun age and the combining form -phage, as well as a multitude of derivatives of existing -age nouns, leaving a total of 921 definitive -age nominal derivatives. A classification of these forms from a morpho-semantic perspective was then carried out with a view to determining the patterns of formation. The classification is based on historical attestation date, on word origin or base word analysis (LOAN, LOAN BLEND, DEVERBAL OR DENOMINAL DERIVATIVE, BLEND, or UNKNOWN) and finally on semantic features (ACTION / RESULT, PAYMENT, QUANTITY, COLLECTIVE). The distribution of these factors over time then allowed me to attempt to answer the following questions. When do language internal derivatives begin? Are the -age forms stable over time or do they evolve, and if so in which direction? The findings of this initial morpho-semantic analysis showed that the pattern of -age derivation
has remained remarkably stable since its beginnings around 1200 until today, with a slight preference for denominal -age nouns over deverbal -age nouns. While these results tend to confirm pre-existing findings in historical lexical morphology, the following step involved assessing the productivity of -age words over time. The novel part of this study was twofold; first focusing on whether -age forms are susceptible to semantic change, and secondly focusing on error forms and transmission errors which enabled me to consider the evidence of historical productivity. In the final section these productivity concerns were extended to contemporary English via a Web Crawler corpus so as to investigate whether -age forms continue to exhibit similar behaviour or whether new patterns can be determined. Three major findings stand out.

1) So-called obsolete -age forms in the OED3 are not obsolete after all, showing the remarkable productivity and adaptability of -age forms over the centuries;

2) -age words follow a reliable semantic pattern, fitting into four main categories of ACTION / RESULT, TAX / RIGHTS / PAYMENT, STATUS / POSITION, COLLECTIVE / QUANTITY. There is a predictable relation between base word and derivative, which may explain the continued success of -age forms, in that they can be used in slang, technical usage, etc. Their remarkable adaptability allows for a sense to be directly dependent on a context, as exemplified by the usage of ownage, or rakeage in English Web 2013;

3) -age forms are not very polysemous, they remain transparent in their relation to the base word. They also appear to be less likely to be lexicalized given this transparency. This low rate of lexicalization is a sign of high productivity and may also explain the high rates of obsolescence of historical -age words in the OED3.

Cette étude diachronique lexicographique vise à étudier le comportement morphosémantique des mots en -age du lexique de l’anglais. L’objectif est de mieux comprendre les mécanismes historiques qui ont permis le passage d’une forme empruntée au latin et au français pour devenir un mécanisme de dérivation productif de l’anglais. À partir du dictionnaire OED en ligne, l’ensemble des mots se terminant en -age ont été relevés, pour ensuite trier et supprimer ceux qui ne sont pas porteurs du morphème lié -age (tels que toutes les lexies contenant le morphème libre nominal -age) ainsi que toutes les formes qui sont dérivées de formes en -age. Ces exclusions permettent de réduire le nombre de lexies à l’étude de plus de 1200 à un peu plus de 900 lexies. Plusieurs étapes motivent le travail de classement et d’observation des lexies : un classement diachronique par année d’attestation, un classement morphologique selon la forme de la lexie (emprunt, forme hybride modelée sur le français ou le latin, dérivé, ou encore amalgame), un classement sémantique des formes selon des grands types sémantiques récurrents : ACTION, RESULTAT, COLLECTIF, QUANTITÉ, STATUT.

À partir de ces données, une analyse diachronique de la répartition des lexies en -age à travers le temps devient possible. Les questions posées sont les suivantes : les lexies en -age ont-elles suivi une courbe stable au fil du temps, ou bien y a-t-il des périodes de changement ? La répartition des lexies déverbales et dénominales est-elle stable ? À quel moment la dérivation a-t-elle pris le dessus sur l’emprunt ? Quelle est la productivité actuelle des lexies en -age et que nous apprennent les nombreuses formes obsolètes du corpus ?

Cette étude permet de montrer la remarquable stabilité de production des lexies en -age au fil du temps, avec une prépondérance de lexies dénominales sur les lexies déverbales. Les courbes de croissance des lexies déverbales et dénominales s’avèrent relativement stables et parallèles, sans grand bouleversement des tendances. Plusieurs périodes fastes de création de lexies en -age sont confirmées par les données. Alors que ces résultats tendent à confirmer les études existantes sur la question des emprunts et des dérivés, la dernière partie de ce travail sur la productivité actuelle des lexies en -age de l’anglais, ainsi que la confrontation entre dérivés en -ing et dérivés en -age permet d’observer des caractéristiques fondamentales.

1) Les lexies obsolètes en -age du dictionnaire ne le sont pas ou plus véritablement, ce qui montre...
la capacité d’adaptation remarquable de ces formes (ce qui les différencie des dérivés en -ity et -ment en particulier ;

2) Les lexies en -age ont un comportement sémantique cohérent qui correspond aux quatre types sémantiques définis (ACTION / RESULT, TAXE / DROIT / PAIEMENT, STATUT / DROIT, COLLECTIF / QUANTITÉ). Le sens du dérivé reste prévisible à partir du sens de la base, ce qui explique en partie leur remarquable succès, en particulier dans des registres variés allant de l’argot aux registres techniques. On note aussi une adaptation exceptionnelle, rendant le sens motivé par son contexte socioculturel (comme le montrent les emplois des lexies ownage, ou encore rakeage dans le corpus contemporain English Web 2013) ;

3) Les dérivés en -age ne sont pas polysémiques : c’est le sens de la base qui explique le sens du dérivé. Ils restent peu lexicalisés, preuve de leur productivité continue.

INDEX

Keywords: neology, productivity, loan words, diachronic lexicography, OED3, corpus analysis, lexical cognitive semantics, Sketch Engine

Mots-clés: néologie, productivité, emprunts, OED3, analyse de corpus, diachronie, lexicographie, sémantique lexicale cognitive, Sketch Engine

AUTHOR

CHRIS A. SMITH

Université de Caen
chris.smith@unicaen.fr