What is the State of Evaluative Affixes in Contemporary English?

Joshua Albair

University of New Hampshire, Durham, New Hampshire, USA

joshalbair@hotmail.com

Joshua Albair is a dual-degree undergraduate student studying Linguistics and Computer Information Systems at the University of New Hampshire, USA. He has a strong interest in syntax and morphology as well as working with non-native English speakers. After graduate school, he plans to pursue a doctorate in Linguistics and go on to teach either in that field or in English as a Foreign Language.

Abstract

In this paper I argue that English uses evaluative morphology and that the current measures of productivity provide counter-intuitive results. Using the Corpus of Contemporary American English I collected data containing instances of evaluative affixes. I explored evaluative affixes in terms of word formation rules and assigned each affix a rating of productivity using the Equation of P (Baayen 1989). I examine these ratings in terms of perceived productivity versus actual productivity according to native-English intuition and the Equation of P, respectively. I discuss the faults of the current understanding and measurement of productivity, as well as the obstacles to accurately collecting data using current corpus technology. Evaluative affixation allows a language to inexpensively add a richness of meaning. Although many linguists mention examples of evaluative morphology in English, there exists no current and comprehensive reference work specifically studying the life and productivity of evaluative morphology. I intend to discuss evaluative affixation as it exists in contemporary English, including: an understanding of what constitutes evaluative morphology, the word formation rules for the thirteen English evaluative affixes, and a study of the productivity of each affix.

Introduction

Before I begin, a brief introduction to terminology is in order. An affix is a type of morpheme, the smallest unit of a language that contains meaning. In English, there are two types of affixes: *prefixes* and *suffixes*. Prefixes attach to the beginning of a word, as seen in recharge, while suffixes attach to the end of a word, as in chargeable. Both *recharge* and *charge* share the same base but the prefix *re-* adds the meaning "to charge again". *Charge* and *chargeable* are similar in meaning but differ in category, with *charge* being a verb and *chargeable* being an adjective. When affixes attach to a base, they may change either the category, the meaning, or both.

Evaluative affixation is a special subset of derivational morphology in which affixes are attached to bases to form new derivatives that convey a meaning of either size or emotion, in the form of diminutives or augmentatives. Diminutives convey "ideas of smallness, pleasantness, familiarity, affection, and […] 'nicknames'" (Dossena 1998, p. 24).[1] Some examples of diminutives are: *booklet* (smallness), "isn't he a *sweetie*?" (pleasantness), *Johnny* (familiarity), *honey-bunny* (affection), *Lefty* (nickname). Augmentatives are the opposite of diminutives. They typically convey a sense of largeness and may contain a pejorative sense, meaning there is a negative connotation. An example of an augmentative in English is found with words using *ultra-*, as in *ultra-bright*, or *ultra-jerk*. *Ultra-* gives both a sense of size or

[1] These categories of diminutives may overlap.

intensity in the former example and intensifies the emotional sense of disdain in the latter example.

My research focuses on the thirteen evaluative affixes currently found in English: *hyper-, mega-, micro-, mini-, nano-, pico-, tera-, uber-, ultra-, -ette, -let, -ling,* and *-y.* Harnessing the advancement of modern technology, I have used an online corpus, the Corpus Of Contemporary American English (COCA), to collect my data. The COCA contains 400+ million words collected from written and spoken texts from popular and contemporary sources, from 1990 to 2009. The COCA provides for wildcard searches and has enabled me to collect affix data. It also allows for the user to view the context in which words are used. With this new technology, linguists and researchers are able to gain a more "real-time" look at the rules of English word formation as they exist and evolve in the present, as well as ensuring the data is the most up-to-date available.

Word formation rules

This section explores the Word Formation Rules (WFRs) for the evaluative affixes. The following WFRs show the rule for attaching an affix to the base and the derivative, i.e. N -> N means the affix attaches to a noun (N) to form a noun (N). My WFRs also give a classification to distinguish which affixes are diminutives and which are augmentatives, as well as providing for additional meanings encoded in each affix. The origin of the affix is included, and it should be noted that only the affixes *-y* and *-ling* are native to English. Finally, there follows a table including some examples, frequencies, and definitions of words used in context.

Suffixes

Table 1 shows the word formation rules for evaluative suffixes. Evaluative suffixes tend to only attach to nouns, to form diminutives, and tend not to change the category of the base, i.e. nouns remain nouns. There are four evaluative suffixes: *-ette, -let, -ling,* and *-y.* Table 2 shows examples of the evaluative suffixes in use with their definitions, along with examples featuring both a low and a high frequency.[2]

Table 1: Evaluative suffixes

Suffixes	*Rule*	*Classification*	*Additional Meaning*	*Origin*
ette	N -> N	Diminutive – "little X"	-feminine noun -feminine group noun	French
let			-	French
ling			-	Native
y			-familiar name -nick-name -rhymes	Native

[2] Although some of the data have highly lexicalised meanings, there is no current way of measuring which words are lexicalised to the degree at which native-speakers do not recognise the affixes as being evaluative. As such, I have kept all instances of "etymological" evaluatives in the data for the sake of consistency. If a measure of the degree of lexicalisation is created, some of this data may be eliminated and the results may change. However, at the present time, there is no way of operationalising this.

Table 2: Examples of Evaluative Suffixes

Suffixes	***Examples***	***Frequency***	***Meaning***
ette	astronette	1	female astronaut
	cigarette	10,608	small cigar
let	buglet	1	small bug
	booklet	1,031	small book
ling	demonling	1	small/young demon
	fledgling	1444	young one
y	barfy	1	nickname for a cat
	larry	23,125	familiar name for Lawrence
	Humpty -Dumpty	117	rhyme

While these suffixes carry the sense of being diminutive, some carry additional meanings. The suffix -ette may carry the additional meaning of "femininity" as seen in *astronette, bachelorette, dudette,* etc. There are 32 examples of feminine forms of generally masculine or gender neutral nouns. There is also a tendency found in the data for *-ette* to form words meaning "female member for team/group X" as in ex-*rockette (1), raiderette (4),* and *redskinette* (2). In each example, the word refers to a woman in a group of women who support a team or are in a band. There were eight occurrences in the dataset indicating this sort of female group branding. This suffix may also be used to form feminine versions of masculine names, such as Georgette, *Nicolette*, and *Harriette*. The suffix *-ette* is of French origin and exists in many of our English words borrowed from French, such as baguette, barrette, marionette. Although these words may not always have a diminutive meaning in English, they have been included in the dataset if they retain a diminutive status in French.

The suffix -y is highly versatile and has several additional meanings apart from being a diminutive: familiarity, nickname formation, and rhyme formation. Familiar names are diminutive versions of proper names and can be seen in examples such as *Tommy* (for Thomas) *Bobby* (for Robert) and *Suzy* (for Suzan). Familiarisation of first names is a common occurrence in English. My dataset includes 183 examples (see Appendix A) of English names that have been made familiar by the evaluative suffix *-y*.

Nicknames are another type of familiarity formed using the -y evaluative suffix. A nickname is a diminutive name given to friends, pets, strangers, or even objects and often focuses on a particular feature or action specific to the individual. Some examples include *Barfy* (the name of a cat that throws up often), *Knitsy* (the nickname for a particular woman who often knits), *Baldy* (for a person who is bald). My dataset includes 109 nicknames formed using the *-y* suffix, found in Appendix B.

Finally, *-y* is used in the formation of rhyme schemes. In English, it is common to form a sort of nonsensical rhyme for talking to children, for creative purposes, or to be silly. Some common examples are *oopsy-daisy* (baby talk), *Humpty-Dumpty* (character in a fairy-tale for children), and other nonsense rhymes such as *palsy-walsy* and *squeaky-deaky*. My dataset includes 138 examples of rhymes made with the *-y* suffix, found in Appendix C.

Prefixes

Table 3 displays the Word Formation Rules for evaluative prefixes. As with the evaluative suffixes, the prefixes tend not to change the category of the base, i.e. nouns remain nouns. Although some of the prefixes may attach to nouns, adjectives and verbs, examples of evaluative prefixation of verbs are rare in my dataset. There are five evaluative augmentative prefixes: *hyper-, mega-, uber-, ultra-, tera-*. Additionally, there are four evaluative diminutive prefixes: *micro-, mini-, nano-, and pico-*. Table 4 shows examples of the evaluative prefixes in use with their definitions, along with examples featuring both a low and a high frequency.

Table 3: Evaluative prefixes[3]

Prefixes	*Rule*	*Classification*	*Additional meaning*	*Origin*
hyper	• N -> N • A -> A • V -> V	Augmentative – "big X"	over, beyond, above	Greek
mega			metric : 1 million	Greek
uber				German
ultra				Latin
tera			metric : 1 trillion	Greek
micro		Diminutive – "little X"	metric : 1 millionth	Greek
mini				Latin
nano			metric : 1 billionth	Latin
pico			metric : 1 trillionth	Latin

Table 4: Examples of evaluative prefixes

Prefixes	*Examples*	*Category*	*Frequency*	*Meaning*
hyper	hyper-abrupt	A	1	really abrupt
	hyperachievers	N	1	super achievers
	hyper-corrects	V	1	action of over-correcting
	hypertension	N	1467	high level of tension
mega	mega-adjustable	A	1	really adjustable
	mega-dork	N	1	super dork
	megawatts	N	523	1 million watts
uber	uber-active	A	1	really active
	uberchallenge	N	1	really difficult challenge
	ubermensch	N	27	German for *superman*
ultra	ultrabad	A	1	really bad
	ultrabots	N	1	super robots
	ultraviolet	A	1649	type of super wavelength
tera	terayacht	N	1	huge yacht
	terabytes	N	50	1 trillion bytes
micro	microcoded	A	1	type of very tiny coding compression
	microbattery	N	1	very tiny battery
	microscope	N	1721	scope to see at the microscopic level

[3] The base of these compounds provides the base meaning, but it is the evaluative affix that intensifies the positive or negative connotations.

Prefixes	*Examples*	*Category*	*Frequency*	*Meaning*
mini	mini-mall-size	A	1	size of a small mall
	mini-bagels	N	1	small bagels
	mini-series	N	766	small series
nano	nano-enhanced	A	1	enhanced on a very small scale
	nanocircuits	N	1	really small circuits
	nanotechnology	N	664	technology on a very small scale
pico	picoarchitectural	A	2	architectural design done very small
	picobrain	N	1	insult - very small brained individual
	picoseconds	N	16	1 trillionth of a second

The suffixes *mega-*, *micro-*, *nano-*, *pico-*, and *tera-* all carry an additional scientific meaning of metric measurement, with *mega-* and *tera-* being large and *micro-*, *nano-*, and *pico-* being small. According to the OED (1989), hyper- carries the additional meaning as a prefix with "the prepositional force of 'over, beyond, or above'" and is "rarely [found in] verbs; e.g. hyperdeify" This correlates with the findings in my data set of only two noted occurrences of an evaluative affix attaching to a verb: hyper- corrects in the context "As a result, the speaker (mistakenly) *hyper-corrects* the /jʒ/ phonetic sequence back to /jz/"; *hyperextend* in the context "…posture by forcing subjects to *hyperextend* the knee…".

Evaluative affixes typically attach to nouns and attachment to verbs is rare. This may be due to the fact that evaluative affixes tend not to change the category of the base. For example, it is possible to attach the prefix *mini-* to the base kick to form *mini-kick*. However, native speakers are unlikely to say "I'm going to *mini-kick* you" and would rather change the would-be verb to a deverbal-noun with the construction "I'm going to give you a *mini-kick*". This changes the verb of the sentence to give rather than *mini-kick*. The same could be said of all the evaluative prefixes, as in "take an *ultra-nap*" or "go for a *mega-run*". Therefore, there are very few occurrences of evaluative affixes attaching to and producing verbs. Instead, evaluative affixes tend to attach to deverbal-nouns and produce deverbal-nouns.

Productivity

My research is primarily concerned with the productivity of evaluative affixation in English. There are many competing definitions and theories of productivity that are fundamentally different, which leaves "studies of productivity […] in a rather poor state" (Bauer 2001, p.25). Some of the definitions of productivity and the scholars behind them are as follows: multiple scholars agree that productivity should be measured in terms of frequency of something in the dataset, but there is a disagreement as to which frequency should be measured (i.e. the frequency of types, tokens, available bases, etc.): Lieber—in terms of number of available bases; Aronoff— proportion of actual words versus total words possible; Harris & Aronoff—probability of new words actually occurring; and Rainer - number of new words produced in a specified time period (Bauer 2001, p.25). For the purposes of this research, I will be using Bauer's *Morphological Productivity* (2001) to define productivity as a measure of the availability and profitability of a morphological process.

Availability is a binary state: either a process is available and alive, or it is unavailable and dead. If a morphological process is available, it may be utilised to create new words, or coinages. An example of an unavailable process is that of the suffix *-th*, as in warmth, strength, and width. English no longer uses *-th* to create new words, and therefore *-th* is now unavailable (Lieber 2009a, p.61).[4] Additionally, availability of a morphological process may change depending on the time period under observation; originally the suffix *-th* must have been available, but in the present state of Modern English it is unavailable. Bauer asserts "statements of availability are temporally located … [w]hat is available in one period may not be in the next" (Bauer 2001, p.205).

The profitability of a morphological process is the "extent to which its availability is exploited in language use" (Bauer 2001, p.211). However, in order to determine the extent of exploitation of availability, one must count the number of coinages. If a process is unavailable, there will be no profitability, and therefore no current productivity. These two ideas come together to mean productivity is the measure of how profitable an available morphological process is in a living language. Bauer (2001, p.41) sums this up nicely:

Productivity is all about potential. A process is productive if it has the potential to lead to new coinages, or to the extent to which it does lead to new coinages. We are aware of productivity only through the new coinages.

Table 5 demonstrates productivity existing within the confines of both availability and profitability. As demonstrated, if a process is unavailable, it is also unprofitable and, therefore, not productive. When a process is available, the most productivity occurs when there is a high profitability. The example of the suffix *-th* has been shown to be unavailable and is therefore unproductive. The suffix *-ness* is both available and highly profitable and is, therefore, highly productive.

Table 5: Availability and profitability

	Available	*Unavailable*
Low profitability	minimally productive	not productive
High profitability	maximally productive	not productive

Along with difficulties in defining productivity, there are many different methods of measuring productivity. One way productivity can be measured is to compare one corpus over time from an older version to the present and calculating the new additions (Bauer 2001, p.157). While this approach would certainly highlight new formations, it is impractical for this research as the COCA is a new corpus and time-constraints rule out the ability to wait a significant amount of time for new additions.

Another measure of productivity, the *Equation of P*, is a measure of productivity crafted by Baayen (1989) and shown as: $P = n1/N$ where n1= the number of hapax legomena (words in the data-set with a frequency of 1) and N = the total number of tokens for the affix. The equation of P ($P = ni/N$) operates under the theory that "the more productive a process is, the more new words it will give rise to and the more chance that these items will occur in a corpus with a very low token frequency, sometimes only once" (Lieber 2009b). These one-frequency words, hapax

[4] A reviewer points to the word *coolth* which, on the surface, appears to be a new coinage but has been attested as far back as the 16th century, according to the Oxford English Dictionary.

legomena (or hapaxes for short), are the key to measuring productivity in this equation.

An additional measure of productivity is "global productivity", known as P*, in which both P and type-frequency are charted. *Types* are individual words using a specific process (in this case an affix) recorded in a corpus. Global productivity produces a scatter plot in which processes occurring with low P and low type-frequency are shown in the bottom left. The high P-rated and higher type-frequency affixes appear in the upper right. This measure quickly shows which affixes are the most and least productive, given extreme positions on the chart. However, the fault of this measure is that "it is not possible to weight the relative contributions of [the two criteria] in such a chart" (Bauer 2001, p.154). This measure cannot compare productivity between two affixes when one experiences high P and low type-frequency and the other experiences low P and high type-frequency, meaning each affix is in a similar location along a different axis. In these situations, it renders the question "is x more productive than y?" meaningless (Bauer 2001, p.154). Global productivity is more suited to a quick view of overall productivity and would not allow for an in-depth analysis and comparison of evaluative affixes.

In my research I have focused on one of the most commonly known measures of productivity, the Equation of P, to measure the productivity of evaluative affixes. It is generally held that they exhibit very low productivity, as shown in an introductory text book to Linguistics by Fromkin, Rodman, and Hyams: "Other derivational morphemes in English are not very productive, such as the suffixes meaning 'diminutive', as in the words pig + let and sap + ling"(2003, p. 90). Although evaluative affixes may be less productive than other affixes, they are still productive. Until now, there has been no comprehensive list of English evaluative affixes compiled with their respective measures of productivity. This has left little means to discuss exactly how productive these affixes are. Table 6 displays the productivity of each evaluative affix using Baayen's measure of productivity.

Table 6 Measure of productivity of evaluative affixes

Affix	*Types*	*Hapaxes*	*Tokens*	*Productivity (P)*
Suffixes				
ette	181	72	18,742	0.00384
let	136	54	10,621	0.00508
ling	89	37	11,717	0.00316
y	741	236	303,625	0.00078
Prefixes				
hyper	1,050	614	9,211	0.06666
mega	1,022	650	5,461	0.11903
micro	1,882	987	31,108	0.03173
mini	1,879	1,241	7,759	0.15994
nano	488	279	4,148	0.06726
pico	37	13	124	0.10484
tera	26	14	198	0.07071
uber	114	108	159	0.67925
ultra	1,016	594	7,938	0.07483

Data analysis

There are two observations immediately available in this data: productivity of prefixes versus suffixes and productivity of native versus non-native affixes. As this data demonstrates, prefixes are more productive than suffixes; the suffixes exhibit P of no greater than 0.005, less than P of any prefix. Also shown, native affixes are less productive than non-native affixes. Of the two native affixes, *-y* and *-ling, -ling* experiences P = 0.0031. Even if prefixes and suffixes are examined separately, the French suffixes still exhibit more productivity than the native English suffixes, with *-let* at 0.005 and *-ette* at 0.0038, while *-ling* falls at 0.0031.

However, while this data provides some answers, it raises questions about the Equation of P and the idea of productivity. Many of the productivity ratings of evaluative affixes are counter-intuitive. Additionally, the productivity for the suffix *-y* is vastly different from my early predictions; while intuitively *-y* should be the most productive given the vast array of possible bases with which to attach, it in fact exhibits the lowest productivity of all the affixes.

When examining the number of word types in a corpus, it must be understood that some affixation processes apply to a larger range of bases than others. For example, the prefixes *pico-* and *tera-* have a pragmatic restriction in that they are typically used only in the scientific community as a means of measurement, such as *picosecond, picogram, terabyte, terajoule.* The prefixes *micro-* and *nano-* are also used in the scientific community (*microgram, nanobot, nanosecond*) but as technology has become more common and the average person is exposed to these prefixes more frequently, they have expanded their range of bases and are attachable to additional, non-scientific words (*microcootie, microdecision, nanoboat, nanodeath*). As marketing companies and advertisers are more familiar with these prefixes, they may invent product names incorporating these prefixes, further fuelling their public acceptance. The more often these affixes are used, the more profitable they become and the more productive they are perceived to be. To demonstrate this, an average person would most likely have an idea of what a "*nano-scratch*" or a "*micro-touch*" might be (an action so slight, it was barely felt or seen); contrast this with a "*pico-scratch*" or a "*tera-touch*", which are less transparent.

Table 7

Affix	*Types*	*Hapaxes*	*Tokens*	*P*
micro	1,882	987	31,108	0.03173
nano	488	279	4,148	0.06726
pico	37	13	124	0.10484
tera	26	14	198	0.07071

In this sense, *micro-* and *nano-* are less restricted than *pico-* and *tera-* and apply to a larger range of bases. Intuitively, if an affix may be used on a wider range of bases and average native-speakers begin producing new words with the affix, the productivity should be higher than that of an affix with a very restricted set of bases, rarely used to create new, analysable words. The problem arises when P is measured for these four prefixes. Using the equation of P, both *pico-* and *tera-* experience a higher level of productivity than *micro-* or *nano-*.

Bauer (2001) notes that Aronoff (1976, p.36) believes it to be unfair to consider type frequency when considering productivity, due to the restrictions of word formation contained within each affix. Bauer also concedes that higher type frequency "does have something to do with the perceived productivity of a particular process" (2001, p. 145). The Equation of P is counter-intuitive to our perceived productivity because of the disparity between the number of types created using an affix and the actual level of productivity measured, according to the equation.

Table 8

Affix	*Types*	*Hapaxes*	*Tokens*	P
uber-	114	108	159	0.67925

This same problem of counter-intuitive results occurs with the affix awarded the highest productivity, *uber-*. *Uber-* is recorded as having the second lowest number of tokens in the corpus, yet it is rated the most productive because nearly 70% of the tokens are hapaxes. This measure of P is counter-intuitive for two reasons: the affix is barely recorded in the corpus and this affix is rarely used in native-English. If uber- had a total of 1,000 tokens and 679 were hapaxes, then the measure of P of 0.67925 would feel more accurate because the affix *uber-* would be understood to be used frequently with established words as well as to create new words. In the real world outside of the COCA, the affix *uber-* became popular with "geeks" and computer games through the use of slang, such as the well known phrase (to gamers) "*uber-leet*" which translates as "*super elite*". As this prefix gained popularity, it grew to be used by the general public as an alternative choice for super, which is what the German word *über* actually means. This spike in popularity of creating new types is shown by the number of hapaxes, but the limited number of total tokens also shows that the affix has not gained full acceptance and is not commonly used. Thus the measure of P in this instance is overstated; it does not reflect the reality of the true, low productivity of the affix.

Table 9

Affix	*Types*	*Hapaxes*	*Tokens*	*P*
-y	741	236	303,625	0.00078

The last major discrepancy between perceived productivity and the measure of P occurs with the results for the suffix *-y*. Due to its exceedingly large number of possible bases and the high frequency of use, *-y* has the highest record of tokens of all the evaluative affixes. One would expect -y to be highly productive, yet it experiences the lowest rating of P. This problem may be inherently due to the inadequacy of capturing evaluative uses of *-y*.

There are two reasons why I believe the productivity of *-y* is registering so low. First, the register (the environment of a word) in which we use *-y* to form hapaxes may be too difficult to capture in the corpus. Native-English users may make many words diminutive by adding the suffix *-y*; however, this is usually done in a very informal register or in a mocking manner. Consider an older brother mocking a younger sibling: "Do you have a little *crushy-wushy*?" The term "*crushy-wushy*" is a one-off, one-time-use word made up on the spot. Native speakers will understand that this term implies the older brother is teasing the younger sibling about liking someone. However, this rhyming word is highly unlikely to be recorded in a corpus because of its register; the corpus does not typically capture insults made to siblings, unless these are fictional insults written in a book which the corpus incorporates.

These types of derivatives using *-y* in a private register are not often recorded in the main sources of the COCA: literature and publicly accessible speech, such as news shows or television shows. Instead, these types of words remain one-time-use, informal, and often unrecorded even though they are valid examples of evaluative affixation.

The second reason for the low productivity is that *-y* is so highly used in forming some very well-known words, the frequency, and consequently the token count, is dramatically raised. The diminutive familiar name for Lawrence (*Larry*) was recorded 23,125 times and experiences the highest frequency of all familiar-name creations. There are many hypothetical reasons why Larry is the most common name; perhaps *Larry* is the name of an anchor for a news show that is fed into the COCA. This would raise the frequency of the word *Larry* every time he appeared. The massive number of hits for the name *Larry* detracts from the impact productivity of familiar names that are hapaxes, and therefore new creations, such as *Craiggy*, *Sibby*, or *Tiffy*.

To demonstrate the effect of high-frequency common words significantly detracting from P, I sorted the evaluative instances of -y into four categories: diminutives, familiar names, nicknames, and rhymes. I recorded the statistics of each group and then removed the top three highest-frequency words from each. This significantly lowered the token count from each group and lowered the total tokens of *-y* by almost 100,000 hits.

Table 10: New -Y Suffixation

Affix	*Types*	*Hapaxes*	*Tokens*	*P*
-y original	741	236	303,625	0.00078
-y new	729	236	216,942	0.00109
-y diminutives	311	127	44,758	0.00284
-y new dims.	308	127	24,247	0.00524
-y familiar names	183	17	251,591	0.00007
-y new fams.	180	17	189,679	0.00009
-y nicknames	109	29	5,937	0.00489
-y new nicknames	106	29	2,108	0.01376
-y rhymes	138	63	1,339	0.04705
-y new rhymes	135	63	908	0.06938

This measure improved P in each category and raised the overall rating of P for the suffix *-y*. However, despite *-y* being one of the most frequently used evaluative affixes, with a very large range of bases, and despite this effort to equalise the "unfairness" of the high frequency words lowering the productivity of *-y*, it is still the lowest ranking affix in productivity.

Conclusion

This paper has demonstrated that English contains and uses evaluative morphology. If a rough total is taken of all the tokens demonstrating evaluative affixation (roughly 410,811 at the time of collection) and that total is divided by the total count of all tokens in the COCA (roughly 400,000,000), we see that 0.103% of all tokens are

evaluative. This percentage may be low but it validates the life of evaluative affixation in the English language.

This data demonstrates the existence of evaluative affixation in contemporary English. However, discovering the problems of applying P to this data has shed light on the inherent difficulties of quantitatively measuring productivity in general. The struggle to understand the productivity rating of *-y* outlines a major hurdle for accurately measuring productivity. If a process occurs outside the register of the corpus it is not being accurately recorded. If it experiences a very high frequency for some familiar derivatives, the productivity is not accurately calculated, as these derivatives detract from the impact of many newly created words, shown as hapaxes in the corpus. Another problem occurs when comparing similar affixes that experience different restrictions, as demonstrated in the comparison between micro- and *nano*- versus *pico*- and *tera*-. These affixes reveal the paradox in the Equation of P: the more well known an affix is and intuitively perceived as productive, the less productivity according to the measure of P the affix may actually experience in native-English. This problem is also expressed in the extremely high value of P for the prefix *uber*-. The Equation of P will produce a high rating of productivity even if a process gives rise to very few types, as long as those types are hapaxes. This means affixes perceived as being less productive occur with fewer tokens, yet may have comparatively more hapaxes due to their limited presence in the corpus, and thus exhibit high productivity according to the Equation of P.

In order to more accurately calculate productivity, I believe another equation is needed. There is a need to address the problems encountered in this research, i.e. data being outside the register or clashes in perceived versus calculated productivity. I believe these fixes must be addressed in an equation, as online corpora will never be able to catch all spontaneous uses of evaluative affixation, and I cannot find fault in the data collected through the COCA. Additionally, English is constantly changing and future studies of evaluative affixation using online corpora will return different results. As these evaluative affixes take on new roles, their productivity will change. If the prefix *mega*- becomes "uncool" in the future or there is a technological breakthrough in science spurring a surge in the use of the prefix *pico*- by the general public, the productivity rating will change, and change drastically given enough time. Because productivity is ultimately the measure of the ability of a process to lead to the creation of new words, any changes in popularity in contemporary English will lead to changes in measures of productivity.

References

ARONOFF, M. (1976) *Word Formation in Generative Grammar*. Cambridge MA: MIT Press

BAAYEN, R. H. (1989) *A corpus-based approach to morphological productivity: statistical analysis and psycholinguistic interpretation.* Unpublished doctoral dissertation, Free University of Amsterdam

BAUER, L. (2001) *Morphological Productivity*. Cambridge: Cambridge University Press

CORPUS OF CONTEMPORARY AMERICAN ENGLISH. Available from http://www.americancorpus.org/ [accessed 7 June 2010]

DOSSENA, M. (1998) Diminutives in Scottish Standard English: A Case for 'Comparative Linguistics'? *Scottish Language* 17: 22-39

FROMKIN, V., RODMAN, R., & HYAMS, N. (2003) *An Introduction to Language* (7th ed.) (pp.88 - 90). Boston MA: Thomson/Heinle

LIEBER, R. (2009a) *Introducing Morphology*. Cambridge: Cambridge University Press

LIEBER, R. (2009b) *Personal Communication with author*, November 2009

OXFORD ENGLISH DICTIONARY ONLINE. Available from http://www.oed.com [accessed 1 April 2010]

Appendix A

List of familiar names created with –y:

Andy 11,097
Anny 40
Barney 3005
Barny 6
Becky 2,307
Benjy 208
Benny 2,521
Bethy 3
Betsey 256
Betsy 2,786
Betty 6,607
Billy 12,896
Bobby 10,549
Brucey 2
Buffy 1,131
Cammy 18
Chrissy 326
Christy 1,165
Cindy 3,876
Cissy 2
Cliffy 8
Craiggy1
Craigy 1
Cristy 10
Danny 8,873
Davey 44
Debby 219
Dickey 640
Dicky 209
Dommy 1
Donny 1,064
Dotty 137
Eddy 1,369
Elly 318
Elsy 5
Emmy 34
Emy 14
Etty 10
Fergy 5
Franky 40
Franny 261
Freddy 1,471
Fredy 11
Gabby 381
Genny 29
Georgey 3
Georgy 257
Gerty 16
Gibby 116
Gilly 51
Ginny 767
Gordy 1,271
Greggy 25
Gussy 6
Gwenny 2
Harry 20,994
Herby 80
Hildy 70
Horty 37
Indy 1,624
Izzy 573
Jackey 6
Jacky 98
Jakey 14
Janey 213
Jebby 2
Jeffy 12
Jemmy 64
Jenny 4,549
Jerry 16,650
Jessy 111
Jezzy 1
Jilly 708
Jimmy 14,158
Jody 1,783
Joey 3,899
Johnny 8,667
Jonesey 6
Jonesy 35
Jonny 291

Jordy	572	Maddy	88	Ronny	52
Jorgy	1	Manny	365	Roxy	97
Jory	431	Margey	8	Ruthy	8
Joshy	1	Margy	17	Sally	60
Juby	12	Marky	8	Sammy	3
Judy	8,047	Marty	761	Samy	5
Jussy	11	Marvy	1	Scotty	176
Kandy	71	Matty	503	Shelley	1
Karly	45	Maudy	31	Shermy	1
Kassy	9	Maury	416	Sherry	2,055
Katey	14	Maxy	20	Sibby	1
Kathy	5,622	Meggy	14	Siddy	1
Katy	1,601	Mickey	4,299	Silvy	2
Kenny	5,579	Mikey	92	Smitty	35
Kevy	1	Milly	256	Sophy	25
Kimmy	62	Mindy	111	Stacey	81
Kirsty	155	Missy	976	Stacy	269
Kovy	3	Molly	5,030	Susy	17
Krissy	74	Mondy	1	Suzy	449
Kristy	263	Mordy	30	Tammy	250
Larry	23,125	Morty	39	Terry	171
Lenny	1,496	Pammy	12	Tery	4
Letty	178	Patty	264	Theddy	17
Lexy	8	Petey	10	Tiffy	1
Libbey	59	Raddy	1	Timmy	3
Libby	2,516	Raffy	6	Toddy	187
Liddy	387	Ralphy	4	Tommy	8
Lilly	1,905	Reesy	4	Tony	17,793
Lily	5,380	Reesy	4	Tony	310
Lindy	433	Richey	8	Trudy	113
Lissy	35	Richy	3	Vicky	1
Lizzy	276	Ricky	1	Vinny	66
Lonny	117	Robby	96	Willy	88
Lotty	41	Roby	22		
Lucy	6,089	Rommy	6		

Appendix B

List of nicknames created with -y:

Baghdaddy	3	Cubby	57	Knitsy	12
Baldy	35	Darkey	4	Krusty	15
Bapsy	1	Darky	19	Laffy	2
Barfy	1	Diddy	205	Lammy	1
Barny	1	Diggy	3	Lanny	26
Batty	102	Doddy	3	Lefty	383
B-diddy	1	Dopey	258	Liggy	1
Beany	27	Dougy	8	Limpy	16
Binzy	1	Fabby	5	Lippy	26
Bitsey	3	Figgy	1	Livvy	34
Blacky	19	Fitty	8	Longy	25
Blunty	2	Fitzy	56	Loudy	1
Bluppy	1	Flanny	1	Mimsy	3
Bomby	5	Flopsy	2	Mitzy	4
Bonesy	3	Frenchy	15	Moochy	7
Boobsy	1	Frischy	1	Moony	17
Bootsy	4	Fucky	1	Muggsy	4
Bridey	31	Gammy	7	Mugsy	8
Browny	4	Gampy	2	Muttsy	7
Buggsy	1	Ganny	12	Nebby	2
Butchy	6	Gimpy	34	Netty	9
Butsy	1	Goofy	86	Nevy	2
Buzzy	123	Gorty	36	Newty	1
Chappy	3	Grinchy	4	Norby	2
Chazy	2	Gumby	68	Pigsy	1
Chazzy	2	Hammy	25	Rosey	4
Chicky	70	Higgy	4	Rossy	1
Chinky	3	Humpy	29	Rummy	1
Clippy	7	Iggy	237	Santy	2
Cricky	1	Juggy	2	Scaredy	2

Schmidtty	2	Spleeny	1	Whitey	75
Sharkey	11	Steffy	1	Wifey	30
Sissy	996	Sussy	1	Wolfy	2
Smoochy	59	Swammy	1	Zhidy	1
Sniffy	1	Taddy	1	Zippy	18
Snuffy	5	Thinny	1		
Sonny	2,450	Tweety	5		

Appendix C

List of rhymes created -y:

acey-deucey	3	dotty-spotty	1	fatsy-watsy	1
achy-breaky	7	dressy bessy	7	fishy-wishy	1
airy-fairy	8	ducky-wucky	1	flippy-floppy	1
annie banany	1	easy-peasy	8	floaty-boaty	1
artsy-fartsy	8	eeney meeny	1	foxy knoxy	3
artsy-shmartsy	1	eeney-meaney	1	foxy loxy	36
arty farty	3	eensy beensy	1	freaky deacky	1
arty-farty	1	eensy teensy	1	freaky deaky	1
bunny cunny	2	eensy weensy	4	freaky-deaky	4
casey-dacy	1	eensy-beansy	1	freaky-deeky	1
catty latty	12	eensy-weensy	3	freddy-weddy	1
chewy-gooey	2	eeny meeny	9	fuddy duddy	6
chunky-funky	1	eenyteeny	1	fuddy-duddy	27
crepey-drapey	1	eenyweeny	1	fuddyduddy	3
curly-wurly	1	eeny-weeny	9	funky-punky	1
delay-shlemay	1	fancy-dancy	2	fuzzy-wuzzy	4
dillydally	7	fancyschmancy	1	geezy-peezy	1
dilly-dally	9	fancy-schmancy	11	goody-goody	46
dippy-hippy	1	fancy-shmancy	6	gretzky betzky	1
doggy loggy	35	fancy-smancy	3	groovy-doovy	2
dollsy wollsy	1	fatsy patsy	2	handydandy	1

handy-dandy 14
hanky-panky 76
heady-weddy 2
herby curby 1
hibbity-dibbity 1
hickory-dickory 1
higgledy-piggeldy 2
higgledypiggeldy1
higgledy-piggledly1
higgledy-piggledy 25
hinky-dinky 1
hippie dippy 1
hocus-y pocus-y 1
hoitytoity 1
hoity-toity 35
hokey-pokey 18
hokeypokey 4
holy-moly 2
honey-bunny 1
honky-tonky 3
hoochy-coochy 1
hoochy-koochy1
hootchy-cootchy 2
hotsy totsy 1
housy-wousy 1
hully-gully 3
humpty dumpty117
humpty-dumpty 14
hurdy-gurdy 16
hurlyburly 34
hurly-burly 76
iddy biddy 3
iggy-biggy 2
ipsy dipsy 1
itsy bitsy 235
itsy-bitsy 47
itty bitty 51
ittybitty 2
itty-bitty 79
itzy bitzy 2
ivey-divey 3
joshy-woshy 1
jussy-wussy 1
kissy-kissy 7
klunky-wunky 4
knicky-knacky2
loony-toony 1
loosey-goofy 1
looseygoosey 1
loosey-goosey 23
loosy-goosy 1
lopey-dopey 1
lovey dovey 10
mimsy-flimsy 1
namby-pamby 2
namby-pamby 8
nappy-nappy 2
nittygritty 1
okey dokey 13
onesey-twosey 2
oopsy-daisy 3
oopsy-doopsy 1
palsy-walsy 1
pitty-patty 1
popsy wopsy 4
puppy luppy 3
roly-poly 21
rooty-tooty 1
sexy-rexy 3
squeaky-deakey 1
tighty-whitey 3
topsy turvy 5
upsy-daisy 3
wakey-risey 1
wakey-wakey 1