

**Putting Blends in their Place**

Karen Steffen Chung

National Taiwan University  
Department of Foreign Languages and Literatures

karchung@ntu.edu.tw

1 Roosevelt Road, Section 4  
Taipei, Taiwan 10617

+886 2 3366-5534

## Abstract

Morphology is most often approached as a synchronic discipline (Katamba & Stonham 2006:3), yet it is clear that everything in any language has a history and did not always exist in its current form. This paper will discuss first, the notion of morphemes as a synchronic construct, and proposes that as we currently define it, a “morpheme” is an entirely *subjective* notion, depending on the specific language user’s perceptions at one specific point in time. We should accordingly work this into our theories and teaching texts, rather than tiptoeing around it awkwardly and not treating the resulting confusion explicitly.

Second, the paper addresses the issue of what word formation types we deal with in morphology, and which ones we “leave to the lexicon.” We propose that all types of word formation strategies should be treated in morphology, even when they seem to be highly irregular, more “creative” formations. English blends like *sunbrella* and *locavores* are taken as a case in point. The fact that blends are a very popular word formation device, with new ones appearing daily in the media, and the fact that they are in general easily understood and accepted, says that they obviously follow patterns, or they could not be imitated. Rather than insisting on linear, serial-processing type rules for what we treat in morphology, this paper suggests using other tools to identify trends and patterns among the examples of a word type. These tools can include examinations of larger units of language such as texts and genres, to better understand where and under what circumstances blends, for example, are, or are not, likely to be coined.

Finally, the paper looks at blends from the viewpoint of their *information value* within a given context. As certain word-concepts become more established, their information value goes down. This makes them susceptible to the infusion of new semantic material, in response to some kind of need, with the result that they become more specific in meaning. For example, the coalescence of Old English *hlaf* ‘loaf’ and *weard* ‘guard’ into *hlāford*, later *lord*, certainly reflected a need for frequent mention of the referent of the item, probably at a time when ‘guarding loaves’ was no longer very relevant to the idea of a ‘master’ or ‘leader’. The phonetic simplification gradually obscured the original component structure, until it was finally interpreted as a simplex word. It was time to “forget” about its composition, since it no longer merited the valuable brain resources that would be required to maintain popular awareness of it. Modern blends in fact follow a similar pattern. Once a word-concept is well established and its information value index begins to fall, it can be clipped and still recognized, thus raising its economy-of-effort index. Blends provide an adjustment that accommodates the brain’s need for a denser information load by shortening certain very familiar concepts which require a shorter processing time.

## Putting Blends in their Place

Karen Steffen Chung  
National Taiwan University  
Department of Foreign Languages and Literatures

*Men ever had, and ever will have leave,  
To coin new words well suited to the age,  
Words are like Leaves, some wither every year,  
And every year a younger Race succeeds.*  
–Horace, poet and satirist (65-8 BCE)

Although our topic is “blends”, the creation and use of blends in English is in fact only going to be cited as a case in point. On a more meta level, this paper will pose and briefly discuss these three questions:

**First, when does a morpheme stop being a morpheme, and why?**

**Second, are ostensibly irregular word formation processes to be excluded from morphological analysis and consigned “to the lexicon”?**

**Third, what connection is there between information value, blends, and the grain size of words?**

These are questions that are inevitably part of any study of morphology, although they are not always formally and explicitly addressed. In this very preliminary exploration of these issues, we will try to propose some new areas and directions that may hold considerable potential for further development of the field of morphology.

**One: When does a morpheme stop being a morpheme, and why?**

*Words, when written, crystallize history;  
their very structure gives permanence to the unchangeable past.*  
–Francis Bacon, essayist, philosopher, and statesman (1561-1626)

Until the 1940s, morphology was a part of philology; it was concerned with the history of complex words, or words that “were once complex” but are no longer recognizable as such. In the 1940s, morphology changed course, and became mainly a *synchronic* discipline. Instead of digging into the etymologies of words, it was required to make judgments on what was “recognizable” as being a “smallest unit of meaning”, based on the contemporary language. This meant that as the original forms and meanings of word components changed

and thus became more *opaque* to modern speakers of the language, they accordingly lost their status as “morphemes”, either coalescing with other components and often leaving puzzling fragments behind, or simply disappearing.

Opting for a synchronic approach was a conscious *decision* made for the field as a whole. Morphology *could* have continued developing with a historical orientation, but another path was chosen, and we still mostly continue along this path today. In her *History of the language sciences*, Auroux (2006: 2325ff) cites the “impact of Saussurean structuralism, mediated by Charles Bally (1865-1947) in 1944 and primarily represented in the work of Hans Marchand (1907-1978) and his pupils on the one hand, and generativism” on the other, as the key forces that molded morphology into what it is today. The work of tracing word origins now falls under the categories of etymology, historical linguistics, lexicology, and/or lexicography. In many ways, it does make sense. Morphology is mainly about the rules behind how words are put together, and rules are expected to be consistent and transparent.

One important question, however, has not been fully addressed all these years, though it is on occasion brought up: How is it to be decided what is and what is not a “meaningful component” of a word? Meaningful to whom? At what point in their education and life? Assuming what kind of background, life experience, and interests?

Let’s look first at a sample of how a few linguistics and morphology textbooks handle – or skirt around – this question:

(1) Bloomfield (1933; 1984: 61): “A linguistic form **which bears no partial phonetic-semantic resemblance to any other form**, is a simple form or *morpheme*.” [Emphasis added, here and in the following.] No further discussion regarding possible differences in judgment as to what is or isn’t a morpheme is appended. His inclusion of the word “partial” is interesting, and it covers phonologically-similar allomorphy.

(2) Fromkin, Rodman, Hyams (2003: 81-82): “A morpheme was defined [on p. 76] as the basic element of meaning, a phonological form that is arbitrarily united with a particular meaning and that cannot be analyzed into simpler elements.”...“Bound forms like *huckle-*, *boysen-* [as in *huckleberry*, *boysenberry*], and *luke-* [as in *lukewarm*] require a redefinition of the concept of morpheme. Some morphemes have no meaning in isolation but acquire meaning only in combination with other specific morphemes.” ... “...for modern English speakers, Latinate morphemes such as *-ceive* and *-mit* **have no independent meaning**. Their meaning depends on the entire word in which they occur... While some of these words may have historical origins, **there is no present meaningful connection.**”

(3) O’Grady (2001: 138): “Because the components of words like *receive* and *deceive* carry no identifiable meaning **for the average speaker of English** (who has no knowledge of Latin), **we will not treat them as separate morphemes** in this book. Thus **we take the word *receive* to consist of a single morpheme.**” p. 171: Note (2): “An interesting fact about these forms is that although *ceive* and *mit* have no identifiable meaning, they undergo certain

alternations that suggest that **they have a special status in the grammar**. Thus the *ceive* in words like *receive* and *deceive* becomes *cept* in *receptive* and *deceptive*, while the *mit* in words like *submit* and *permit* becomes *miss* in *submissive* and *permissive*.”

(4) Haspelmath (2002: 16): “Morphemes can be defined as the **smallest meaningful constituents** of a linguistic expression.”...“When we try to divide *chameleon* further (e.g. *chalmeleon*), we do not obtain parts that can be said to be meaningful, either because they are not found in any other words (as seems to be the case with *meleon*), or because the other words in which they occur do not share any aspect of meaning with *chameleon* (cf. *charisma*, *Canadian*, *caboodle*, *capacity*, in which it would be theoretically possible to identify a word part *chalca*). Thus, *chameleon* cannot be segmented into several morphemes, it is **monomorphemic**. Morphemes are the ultimate elements of morphological analysis; they are, so to speak, **morphological atoms**.”

(5) Katamba (1993; 2006: 3): Katamba, referring to Max Müller (1899), says: “His specific claim was that the study of the 400-500 basic roots of the Indo-European ancestor of many of the languages of Europe and Asia was the key to understanding the origin of human language.” And, “Such evolutionary pretensions were abandoned very early on in the history of morphology. Since then morphology has been regarded as an essentially **synchronic** discipline, that is to say, a discipline focusing on the study of word-structure **at one stage in the life of a language** rather than on the evolution of words. But, in spite of the **unanimous** agreement among linguists on this point, morphology has had a **chequered career** in twentieth-century linguistics, as we shall see.” His definition of a morpheme is (p. 24): “the smallest difference in the shape of a word that correlates with the smallest difference in word or sentence meaning or in grammatical structure;” and in the glossary of the same work: “An abstract entity that expresses a single concept within a word.” This definition covers the notion of subtractive morphemes.

(6) Stockwell & Minkova (2009: 78): “...allomorphy varies from being transparent synchronically to being **discoverable** only with reference to a **good etymological dictionary**. Therefore, **what may be considered two separate morphemes or two allomorphs or the same morpheme in Present-Day English will vary from speaker to speaker**.”

“There are many fairly extreme examples of obscured common origin where the divergence is so radical that the relationship of allomorphy cannot be invoked without reference to sophisticated historical knowledge. This is the story of the IE *\*bha* root which, as we saw before, surfaces in a number of common words: *bandit*, *fame*, *infant*, *phonetic*, and *symphony* all contain this root (in the forms  $\sqrt{ban}$ ,  $\sqrt{fa}$ , and  $\sqrt{phon}$ ). [Author’s note:  $\sqrt{\quad}$  stands for “root”.] These **roots** are **cognates**. But would we invoke the relationship of allomorphy between them? **Not automatically, and probably not at all for the ordinary speaker of the language, though it is always interesting to know what the etymological relationship is.**” This source encourages the reader to dig deeply into the histories of the component parts of words, going even beyond the history of English all the way to reconstructed Indo-European

roots.

From this very small sampling, we can see that there is a broad spectrum of divergent stands among different researchers. Still, no clear formula is offered regarding how to decide what exactly counts as a morpheme, according to what kind of person.

In his *Glossary of Morphology*, Bauer (2004: 70-72) admirably describes a broad range of different conceptions of what a morpheme is, tracing it to its beginnings with Baudouin de Courtenay in about 1880, who Bauer says did not define it precisely. He touches on Bloomfield, who would call the English plural ending *-s* as in *books* and *-en* as in *oxen* “synonymous morphemes”, explaining that they can be viewed as either allomorphs, under a **meaning-based system**, or as different morphemes, in a more **form-based system**. Bauer also mentions how generative scholars like Aronoff have “tried to redefine the morpheme with a formal side as an alternative to a meaning side”, which recognizes the status of *mit* and *miss* as in *permit* and *permissible* as morphemes and captures thus their allomorphic relationship (in contrast to O’Grady, who prefers to consider *permit* monomorphemic). Bauer also introduces Anderson’s “A-morphous morphology” which espouses a “general rejection...of the classical morpheme” and claims that “properties of individual lexical items...are not available to syntactic operations” and that “words do not in general have an internal morphological structure for phonological and morphological rules to refer to” (p. 5). Anderson “argues that the form of words is better treated by means of ordered rules operating simultaneously on the semantics, external syntax, and phonological content of the input to the rules” (Sadock 1994: 327). One way to deal with a slippery construct like a “morpheme” is to get rid of it altogether.

Bauer concludes this long entry of central importance to his work thus: “While **it is impossible to give any definition which encompasses all of these various concepts**, perhaps we can define a morpheme relatively uncontroversially as ‘a set of signs whose members are synonymous and in complementary distribution’, while adding to that definition (i) that a set may contain only one member; (ii) that in the most clear-cut instances, members of the set are clearly phonologically related to each other; and (iii) that the conditioning factor(s) for the complementary distribution can be stated in phonological rather than in lexical terms. Thus the plural endings on *cows* /kaʊz/, *horses* /hɔːsɪz/ and *cats* /kæts/ clearly resemble each other (in sharing a sibilant fricative) and are conditioned by the last sound in the **base**, and clearly meet the definition.” He incorporates further flexibility into his definition by adding, “However, if *oxen* and *children* and *alumni* are added to the list, the conditions are less clearly met. And while /waɪf/ and /waɪvz/ are clearly phonologically similar, /beə/ and /bɜː/ (in *bear* and *birth*) are less obviously so, but **there is no well-defined cut-off point on the scale of phonological relatedness.**”

What this sample of viewpoints suggests is that we are saying things in the margins, parentheses and footnotes that we are not – but certainly *should* – be dealing with in our

definitions and main expositions. Rather than hedging, or giving up entirely, can we find a way to accommodate the messy parts into the core of what we are trying to do?

Bauer has made an excellent point regarding phonological relatedness, i.e. the lack of a clear cut-off point, and it is one we would do well to expand on in our growing conception of morphemes in general, and what does and does not constitute one. Although we strive for rigor and preciseness, and like clear, explicit, consistent rules, what we have on our hands here is clearly *fuzzy*, a continuum. Because of this, I propose that there *is* no one way to definitively define a morpheme in English.

We will at this point introduce an analogy from computer science, one that could be very powerful in aiding our understanding of morphemes. It is the notion of serial or digital processing and parallel or network processing. Vertosick (2002:142) summarizes their differences thus: “Networks *invent their own rules as they go*, while digital computers *need to have the rules spelled out for them by programmers.*”

Serial processors are strictly rule-governed, and every step must be inputted without a single mistake by the programmer. The advantage is accuracy and consistency. You will always get the unambiguously correct answer from a **bug-free** serial processor.

The “bug-free” condition results however in steep demands on the programmer. The program fails utterly when it encounters the smallest bug. In addition, it is very **slow**, since it can only process one instruction at a time, given in the correct order. It can shine its light only on the step it is currently working on, and does not allow anything else to be running in the background that might affect the outcome. You get a unique, single answer with no ambiguities, fuzzy edges, or alternative possibilities.

Correspondingly, the disadvantages of parallel processors or networks (“parallel” here does not refer to an array of serial processors working together, but a network-type system) include lower accuracy and less consistency. It offers in exchange, however, huge advantages.

Conventional morphology – and syntax, and other linguistic subfields – is basically serial in nature. Something either is or isn’t a morpheme, we argue a bit about whose opinion counts, the “average native speaker without previous Latin training” wins, then we tally up the results: *shampoo-1, birth·day-2, tele·vis·ion-3, re·bell·ious·ness-4*.

Corpus linguistics began to add weight in linguistics to the parallel/network processing approach. Instead of saying that a usage is absolutely grammatical or not, for example, you can say like Algeo (2006: 14) that the “CIC [Cambridge International Corpus] has more than 7 times as many tokens of *fitted* [rather than *fit* for the past tense and past participle of *fit*] in British as in American texts.” No one claims that either form is the only “correct” one, or that one is *always* the “British” and the other the “American” form. Both are clearly used by both sides, but with a clear majority preferring one of the forms on each side. Instead of becoming a binary yes-no choice, we are now able to give a *scalar* answer to the question about which form is correct, for whom, and how often. Quoting Vertosick again (2002:191): “Networks deal best with ***patterned information*** and ***prefer good solutions to exact ones.***”

I suggest that it is time to adopt a similar approach in morphology.

Returning to that “average speaker of English who has no knowledge of Latin” – exactly what kind of person did they have in mind? A 1979 survey found that about 4.1 Americans had some familiarity with Latin (Eddy 1980); one 2008 source (Stonehill College) says currently only 1.9 percent of the US population have studied Latin. The ratio is likely higher in the UK, though it has been declining there as well. So the number of native English speakers with formal Latin study is not large; however, anybody who knows English, perforce knows a good deal of Latin, whether they identify it as such and whether are aware of it or not. And the higher the level of education, the more one is exposed to Latinate and Greek word components.

While the “average” English speaker may not be able to analyze the components of any given English word on the spot, they very likely have an awareness of at least some of the parts. This awareness, even when it is not outright confident knowledge without checking elsewhere, recalls Aronoff’s form-based approach to morphemes, with the additional feature that the speaker often does have some intuition of the meaning, or at least knows that it *has* a meaning in Latin or Greek, and that they need only get to a good dictionary to find out what it is.

I propose that this knowledge that something *has* meaning, even when that meaning is vague or not immediately known, and even when the speaker cannot parse the word perfectly correctly, still **counts as some level of morpheme recognition**.

Haspelmath gives the example of *chameleon* as a single-morpheme word. A competent English speaker might perhaps call this into doubt, for two reasons: (1) its spelling makes it look a bit exotic, and you might suspect it has word components you can’t easily recognize with a knowledge of just English; and (2) it is a long, polysyllabic word. The native vocabulary of English tends to be monosyllabic; Latinate and Greek vocabulary is often disyllabic. When you see a four-syllable word, you may have a strong intuition that it must be divisible somehow. Humans being the perverse creatures we are, seeing a claim like Haspelmath’s might incite a good many to go and check a dictionary. And indeed, the etymology given in Merriam-Webster is: **chameleon**; often attributive; Etymology: Middle English *camelion*, from Middle French, from Latin *chamaeleon*, from Greek *chamaileōn*, from *chamai* on the ground + *leōn* lion — more at *humble* Date: 14th century.

One might well forget the first part, but there is a good chance the *leon* = *lion* part will stick, because of its similarity to a familiar word – and this word can never again be monomorphemic for this particular speaker, unless they simply forget about what they learned or someday suffer from dementia. You can say, “Well, then this person is an *exception* – *most* people...” The point is, “most people” is not “everybody”, and when you systematically exclude large numbers of people, in this case simply because they have a higher level of bookish education, there may in fact be a problem with your basic working assumptions and your resulting definitions.



Other words will have differing levels of transparency for people with different academic and professional specialties. We can in fact choose to say “the average speaker”, as we have either explicitly or implicitly been doing for years. But if we do so, we should **define precisely what we mean by an “average speaker.”** A way to start might be by taking a survey of a wide swath of native English speakers to reveal their understanding of word composition. And then to organize this information by specialty, age, education, geographical location, social class, and so on, showing typical degrees of transparency for each group, if there is indeed consistency within the groups we have delineated.

Armed with this information, we could still only say, we found this set of levels of word component transparency among *contemporary* English speakers in these categories; our knowledge would have a fairly limited depth timewise. But this would at least be a more representative assessment of the actual current situation, to be preferred over constant reference to that assumed “average speaker of English.”

For those who are getting ready to complain that this is far too much trouble, and we shouldn’t expend resources on expensive and time-consuming “frivolous” social science surveys like this, our one way out is to say, a morpheme is an entirely *subjective* notion **depending on the analysis of one particular user on one particular occasion or time.**

The notion of “morpheme” is subjective not only because different individuals will have different levels of recognition of “the smallest difference of meaning or grammatical function” in a given form, but also because of a lack of agreement regarding just what – how big or how small a unit – constitutes a unit of difference.

Some cases seem quite straightforward, for example, compounds, like *bookcase* and *eyelash*, in which the parts of the whole, the compound, are themselves whole words, i.e. free morphemes.

Transparency tends to fall and controversy may also tend to increase when it comes to other types of units that form a word. Lehrer (2007: 121) says this about “splinters”, i.e. bound morphemes found in blends: “When a splinter becomes so common that people start using it frequently, it may lose its connection with the source word and can be considered as a morpheme in its own right. Of course, since there is a scale from a completely novel splinter to a completely conventional morpheme, the transition from splinter to independent morphemehood is a diachronic process.” Her examples of more conventional splinters, all word final in these cases, include the *gate* in *Watergate* (*Irangate*, *Contragate*), (*a*)*holic* from *alcoholic* (*chocoholic*, *shopaholic*), and *thon* from *marathon* (*walkathon*, *bikathon*). So Lehrer subscribes to the idea of a continuum of units used for word formation, and the “morpheme” in its most typically imagined form is but one point on this continuum. Into the same category as splinters she puts neo-classical combining forms, which are often the result of splinters, such *bio-*, *Euro-* and *eco-* (2000:152).

Derivational affixes, which are supposed to be bound forms, and may be less transparent, weaker semantically, and more abstract than whole words and splinters, are very much a part

of the same continuum. Booij (2002:130) has found that: “Derivational affixes are pieces of morphological structure, just like the constituents of compounds. This position leads us to expect that compounding and derivational affixation do not differ in accessibility for rules of grammar. This expectation appears to be borne out by the facts.” Lehrer (2000:151) further distinguishes between “function” affixes, such as nominalizing suffixes [-*ation*, -*ness*, etc.], which exist for “grammatical purposes”, and “content” affixes, such as the suffixes [-*able*, -*ful*, -*less*, -*free*, -*ess*, *ly*], and most prefixes, which have more specific meanings.

Sometimes a word formation component may be considered “smaller than a morpheme” or “sub-morphemic”, with a meaning that has largely faded historically from memory. A case in point is phonaestemes, about which Knowles (1987:114) says: “In our examples *gleam*, *glimmer*, *gloom*, the initial *gl-* is almost certainly a ‘dead’ morpheme, of which we have only a vague idea of the meaning, and which is reduced phonetically to the point where it just forms part of another syllable. The rest of these words, -*eam*, -*immer*, -*oom*, may also have consisted of well defined morphemes at one time.”

Shisler (1997) says that some linguists consider phonaestemes to be a kind of morpheme; he feels they are not, because they, unlike morphemes, do not play a syntactic role in a word, can appear anywhere in a word, and because their semantic content is not as “potent” as morphemes. Yet it is clear that they carry meaning, albeit vague meaning, and there is some degree of consistency across the words in which they appear.

An example of a unit perhaps a step up from phonaestemes yet falling somewhere below unambiguously recognized morphemes is the -*le* ending in words like *bundle*. Taken alone, we might not hesitate to decide that the word *bundle* is monomorphemic. But if we see or think of it together with words like *kernel*, *nozzle*, *axle*, *navel*, *nipple*, *pimple*, *mongrel*, we may sense some kind of connection between the words due to their final syllable. It may not fully rise to the level of consciousness, but human brains are designed to be good at pattern recognition, and there is clearly some kind of pattern here. In fact it is an ancient Germanic and Indo-European diminutive, *lo* (Skeat 1910: 102).

A speaker might also wonder what the -*en* represents when comparing *maiden* and *maid*, *chicken* and *chick*. In fact it too is a diminutive ending that goes back to the Old English period, and if recognized outright as such, similar to the -*y* in *doggy*, we would consider it a discrete morpheme. This one is further confusing because *chicken* now is the generic term, and *chick* refers to a baby chicken; the diminutive and non-diminutive forms have in effect switched places. Even with the confusion, a speaker will know the two are related and that -*en* is some kind of suffix, though with a meaning even less concrete than *ceive* and *cept* in *receive/reception*.

Then there is *conversion* or *zero morphemes*, in which a difference in meaning is marked by no change at all in the form of a word, e.g. *table* can be a noun; the same word form can also be a denominal verb, as in *to table a motion*. Whether conversion constitutes a distinct morpheme or not is controversial; it is in any case close to the extreme end of the continuum.

Even before this though come *subtractive morphemes*, in which material is *taken away* to mark a morphological operation; Sproat (1992: 64) cites examples from the Muskogean language Koasati, but calls the phenomenon “problematic.” He says, “Although there exist formal phonological mechanisms where, by adding phonological material, one can actually end up removing phonological material, such an approach would, on the face of it, seem to render vacuous any claim that morphological processes must consist of the addition of phonological material. So, it seems as if some morphological processes do not involve morphemes if morphemes, and in particular affixes, are defined as above. Unfortunately, I can do no more than note this as a problem for morphological theory, leaving it unresolved at this point.”

The real issue, however, isn’t whether a phonaestheme, so-called subtractive morpheme, or other unit of word formation is a kind of morpheme or not; the important point is that each occupies a stretch on the same continuum of meaning-carrying elements in word formation that includes canonical morphemes along with many other types of units. And they should formally and explicitly take their place there in our treatments of English word formation.

So our conclusion is: Rather than quibbling about what is and is not a morpheme synchronically, and letting everyone walk away with the uneasy feeling that we either missed one, or called up a little too much historical knowledge that we embarrassingly, un-representatively, and un-averagely happened to have, we should instead come clean and simply recognize the concept of “morpheme” as the *subjective* construct it is. We should not mention it only in passing, brushing off the inevitable confusion, hesitation, and conflicting opinions that ensue, and should not give morpheme counting exercises to our students that pretend to have one correct answer, implying that any right-minded and intelligent person would or should get the same answer.

What is and is not a morpheme – at least in English – is in the eye, ear and mind of the speaker, listener, reader and writer, and it will vary over the span of a lifetime in the same individual. Many different sizes and types of units carry semantic and grammatical meaning of varying transparency, concreteness and integrity. We should make this relativity, subjectivity and scalarity a central part of our definition of a morpheme, and take it as a starting point in all of our morphological investigations, if the morpheme is to continue playing the central role that it currently does. It is perhaps the only way to achieve consistency in our work, and the only truly solid – and honest – basis on which to build the morphological edifice.

## **Two: Are ostensibly irregular word formation processes to be excluded from morphological analysis and left “to the lexicon”?**

*As we must account for every idle word, so must we account for every idle silence.*

–Benjamin Franklin (American statesman and inventor, 1706-1790)

At its heart, this question is about whether we in morphology choose to sift out and then deal only with regular, rule-governed behavior in word composition. If we choose to do this, though, what is to be done with all the leftovers? Are irregular word formation processes – blends for example – to be excluded from morphology and relegated to the “lexicon”, simply because of their low predictability? Who then deals with these, and how?

In the never-ending quest for “productive” rules in morphology, we may sometimes feel frustration when we encounter partial productivity, or are at a loss to formulate rigorous rules that can be applied to produce novel derivations. But what do we do with that partial productivity, where some kind of pattern is visible across a subset of the data, but it is mixed together with apparent idiosyncrasy?

Demanding an absolute one-to-one correspondence between form and meaning for each morpheme is an idealistic approach that simply doesn’t reflect the way things work in the real world, at least not in English. We may find it a convenience to act as though a morpheme is something most people will easily agree on. Often it does happen; these are the tidy textbook cases. But very often it doesn’t.

I suggest that insisting on the tidy, textbook example approach is like hiring a cleaning lady who “doesn’t do windows”, or move heavy furniture to clean behind and under it, or enter the teenage daughter’s room. With so many items intentionally and systematically ignored, you wonder just how much use it is to pay someone to come in at all.

Rather than feeling frustration whenever we find a lack of consistency, we can view irregularity as an opportunity to open up and explore an even more exciting kind of morphological inquiry.

Sun Tzu’s *Art of War* describes the principle of *qízhèngxiāngshēng* 奇正相生, which in Sun’s work means ‘to combine both overt and covert methods when attacking’, e.g. holding peace negotiations while engaging in sabotage. On a more literal and deeper level, it means ‘the regular and the irregular beget each other’, i.e. a period of regularity will eventually give way to irregularity, which will then again settle down into a new regularity, like times of national stability interrupted by periodic conflict and revolutions.

We can think of irregularity as “noise”, the unpredictable and unordered part of a signal that sometimes masks, interferes with, or distorts the message being sent. Vertosick (2002: 313) has this to say about irregularity or “noise”: “We condemn noise when it harms us, but noise created us all. And noise, like death, taxes and the poor, will be with us always.” It is the “mistakes” or irregularities in an ongoing pattern that enable new things to emerge. These

new things may eventually settle into a pattern, but they are sure to have flaws, dissonance, parts that are slightly “off” built into them – if they didn’t, they would lack vigor and the ability to spawn something new. For this reason, we shouldn’t just sweep them under the carpet or feel irritated at them for upsetting the beautiful order that would reign if only they weren’t there to spoil things. They are a vital source of life and regeneration.

“By embracing chaos,” Vertosick says (2002: 313), “life harnessed yet another source of noise and another way to bring diversity into computation so that natural selection could choose among many possibilities, not just a few.” Ignoring the irregularity that begets new patterns is like pretending the whole process of dating, love, sex and birth don’t exist, and only dealing with living beings once they are fully formed individuals, without accounting for how they came into being.

Since it is irregularity that begets new forms – some of which grow and prosper and others of which disappear almost as quickly as they surface – it should be made **an inherent part of morphological study and theory**. As things are, it often seems to get treated as a spanner in the works. But just as a system of any kind needs redundancy, noise, and time for rest – sometimes even whimsy – built into it, so does language, and word formation. We should **incorporate** it into our theories and descriptions rather than trying to tiptoe around it and pretend no one notices or cares about it too much, or tossing it over the fence to the lexicon, like a junk pile for misfits and factory seconds.

Rather than – or perhaps in conjunction with – using “productivity” as our main index of morphological word formation, by for example counting *hapax legomena*, we could consider evaluating the *entropy* of morphological forms, or of a text, though exactly how this would be done remains to be worked out. According to Pierce (1980: 22), in information theory, entropy is a measure of the amount of uncertainty in information conveyed from a message source. “In general, the meaning is that an increase in entropy means a decrease in order. But, when we ask what order means, we must in some way equate it with knowledge. Even a very complex arrangement of molecules can scarcely be disordered if we know the position and velocity of every one.” And (p. 23): “The more detailed our knowledge is concerning a physical system, the less uncertainty we have concerning it (concerning the location of molecules, for instance) and the less the entropy is. Conversely, more uncertainty means more entropy.” And further, “The entropy of communication theory is a measure of this uncertainty and the uncertainty, or **entropy, is taken as the measure of the amount of information conveyed by a message from a source**. The more we know about what message a source will produce, the less uncertainty, the less the entropy, and the less the information.”

Reducing entropy to as low a level as possible, to where there is no redundancy, and the signal is maximally efficient, means a low to zero tolerance for errors. But room for creativity and *play* is needed, and we also need to formally work this into our theories and descriptions. Creativity and play by nature have low predictability and high (but not limitless) entropy, or they wouldn’t fulfill their function. Jokes that are predictable no longer qualify as being jokes.

The same is true of clever turns of phrase, metaphors, and novel word formations, like blends. Without entropy, uncertainty, noise, language would become stale and fall into deep, monotonous ruts. To meet the powerful human need for variety, surprise, play, continual rejuvenation and freshness, random and odd bits that don't quite fit in to established patterns need to be scattered around here and there for possible use. The availability of these bits gives us the ability to be creative and experiment with new combinations – which may sometimes lead to new patches of regularity that may or may not gain momentum and eventually become part of the morphological establishment.

Plag (2004:23) says that “The productivity of a word-formation process can be defined as its general potential to be used to create new words and as the degree to which this potential is exploited by the speakers. Productivity can be assessed by various measures, both corpus-based and dictionary-based. Productivity emerges from the mental lexicon as the result of parsability, relative frequency, semantic and phonological transparency.”

Plag seems to be talking mainly about the more straightforward rule-governed kind of productivity; but we also need to address the other kind of productivity, namely, free-form new creations from found bits and locally available materials. In addition to studying, analyzing and incorporating the *zhèng* 正 ‘straight, regular, orthodox’, we should also be trying to do more than we currently are with the *qí* 奇 ‘irregular’ part of word formation. But how can we go about it?

Blends are nothing new; Pound's 1914 book on blends says, “It seems time that specific attention be called to the contemporary popularity of blends, and to the freedom felt in their coinage, not only in the factitious creations of the lettered class, and in folk-forms, but in scientific nomenclature, in trade terms, and in arbitrarily made baptismal names and place-names.” Two things to note are, first, that blends were already popular enough in 1914 for Pound to write a substantial, dedicated work on them; second is the mention of some of the contexts in which blends are especially common: “scientific nomenclature, in trade terms, and in arbitrarily made baptismal names and place-names.” There are certainly many other **contexts in which blends are common; identifying** these should be another focus of our work.

One problem is that when we stick too closely within the bounds of our own field and deal mainly with data at the word level, we are unable to capture larger trends and groupings. To address this we can perhaps combine morphology with other approaches to linguistic analysis, like stylistics and text analysis. **Genre and text analysis** are traditionally thought of as belonging mainly to literary and discourse analysis. But these tools can also be applied to a higher level of morphology: we can discover patterns of what kinds of words and formations we can expect from what genre or text type. From the *Method in text-analysis* syllabus (2007): “Genre will tend to define a particular way of speaking or writing and to shape the vocabulary, including how frequently particular words appear.”

We can determine, for example, the percentage of Greek, Latinate and other learned

vocabulary for each type, that of native Anglo-Saxon words, of proper nouns, of more recently foreign borrowings, and the number of neologisms. Among neologisms, blends have become extremely popular and common in the media. But how much are they actually used, per piece and per genre?

Using such methods as concordancing, we can study writings in various genres to discover where what kind of blends are most often being coined and used. Some examples from news media writing include: editorials (*Minnewegian*, Minnesota + Norwegian, re *Fargo*, Palin accents), politics (*Obamacare*, Obama + health care, *Obamanauguration*, Obama + inauguration; *affluenza*, affluence + influenza), arts and entertainment news (*Fanilow*, fan of Barry Manilow; *Brangelina*, Angelia Jolie + Brad Pitt; *celebutante*, celebrity + debutante; *dramady*, drama + comedy; *wovel*, Web + novel; *televisual*, television + visual; *workation*, *staycation*, *nocation*, work/stay/no + vacation), computers/Internet/electronics (*webisodes*, web + episodes; *Crackberry*, crack + Blackberry, a forum for solving Blackberry problems), cooking (eggplant *Zentastic*, Zen + fantastic, *aprichoc* (cake), apricot + chocolate; *spork*, spoon + fork; *arugulance*, arugula + arrogance), sports, fashion (*guyliner*, guy + eyeliner), advertising, business (*Webinar*, Web + seminar), linguistics (*Blackcent*, black + accent, coined by John McWhorter; *Manglish*, mangled + English), miscellaneous nonce words (*huggle*, hug + cuddle; *gregacious*, gregarious + gracious; *testigroanials*, testimonial + groan), and so on.

It is clear that blends proliferate quickly in many new, growing areas, such as computer programming – online lists of, for example blogging terms, include large numbers of blends that seem to only continue growing, e.g. (*blogiversary*, anniversary of a blog; *blogosphere*, blog + sphere, *blognoscenti*, blog + cognoscenti; *blogorrhea*, blog + diarrhea). In the process of collecting examples of blends, we should also identify and track what kind of semantic, expressive and social needs the new word forms are a response to.

**Parts of speech** can also be taken into consideration. While nouns are especially common, blends can also be verbs, e.g. *churr*, chug + purr, the sounds baby grizzly bears make when nursing and contented; earlier formations include *don*, do + on, and *doff*, do + off. Collecting, organizing and analyzing this kind of information can help lower the entropy associated with the coining and use of novel word formations like blends.

It is important at the same time to note, for comparison, contexts where blends typically **do not appear** at all, or only very seldom, or where only very long-established and highly assimilated ones do. These might include more formal documents and writing styles, but corpus studies would be required to establish this.

Another approach would be to make a **structural comparison of blends in English and other languages**. Chinese in particular offers an interesting comparison: the approach used in forming blends in English is simply the normal way of forming complex words in Chinese. Chinese has a close single syllable-to-single morpheme correspondence; to combine any two (or more) ideas into a single concept, one needs only put them together in a compound, with no phonological modification, beyond inevitable tone sandhi, in the correct order, i.e. head

final for modification, VO for predication, i.e. phrases. Western European languages might tend to behave more like English, but what differences there are could be revealing.

We also need to better understand how and the extent to which the creation of neologisms such as blends are being **consciously and actively engaged in** and encouraged, for example, the *Washington Post* holds **word contests** on nonce word formation, which inevitably include lots of blends. There are many online lexica of new words, and seeing numerous examples of a given word type, like blends, is likely to further reinforce and promote a growing trend. The same is true of seeing blends or other formations in trendy, widely-admired writing, say, in *New York Times* editorials, or in magazines covering popular culture. Fanciful creations often circulate as e-mail forwards and on web pages.

To sum up: Blends have been formed for centuries, and they have become an extremely popular and common word formation strategy in English. Although they are a relatively irregular word formation strategy, people are obviously finding it easy and fun to form and use them, whether only for a single or more long-term use. In terms of how brains file and call up semantic notions, they have a great advantage over a monomorphemic neologism not as transparently related to what is referred to, since the fragments usually make the meaning clear.

Since blends are so widely coined and enjoyed, they certainly have rules and patterns, and for this reason they should not simply be relegated “to the lexicon”. Ways of approaching the formation of blends and other less strictly regular word formation processes include text and genre analysis of blend use in, for example, the various types of news media English, such as political news, editorials, columns, and entertainment. Where blends typically do *not* appear is also relevant and worth studying. Analyses can also focus on what parts of speech blends tend to be. We may gain insights from comparing English blends to equivalents in other languages, especially typologically divergent ones such as Chinese, where a strategy similar to blend formation is a very common, everyday way to coin new compounds. And finally we can study the ways blend formation are encouraged and spread, including overt ways such as contests and e-mail forwards.



### **Three: What connection is there between information value, blends, and the grain size of words?**

*Do not say a little in many words but a great deal in a few.*

–Pythagoras (Greek philosopher and mathematician, BC 580-500)

Finally, we will consider the role of *information value* in blend formation.

New and old information are marked in discourse in a number of ways that fall onto a continuum. The more familiar something is, or the more it is a given, the more likely we are to somehow reduce or compact it, so that attention in the discourse can be focused on new information, that is, parts of the utterance with a higher information value. Cutting back on the time, energy, relative length, volume and so forth on some parts of the discourse frees up listener memory and attention resources for more novel and interesting items. And our minds seem to prefer a pattern of denser bits of information padded with fluffier ones, rather than a steady stream of equally weighted ones. This provides more variety; and the lighter bits give the brain time to absorb what has just gone by, and to prepare for the next relatively substantive item.

Once a word-concept is well established and its information value index begins to fall, it can be clipped and still recognized, thus raising its economy-of-effort index. Blends are another possible outcome – in essence they are a combination of two clipped forms. But blends have the advantage of tending to maintain the prosodic integrity of (usually) the second component word while raising the overall information value of the single item.

Blends, such as *sunbrella*, *snackrifice*, *mocktails*, *chillax*, *locavores*, *femisphere*, *newstalgia* reflect a need for more information concentrated in less space. They are a handy way of coming up with a unique name for a useful referent with a minimum of brain resources expended on onomasiology – the components are already in the context; all that has to be done is to combine them in a phonotactically allowable way that also leaves the word components recognizable to most people, or at least people familiar with the word's context and background. If the elements are not known to an individual – and blends often require a fairly good acquaintance with current events and popular culture which not everyone necessarily has – they may be motivated to Google it, learn about the components, and why they are put together. Blends, especially when they are first coined – and many never survive beyond their initial coining and end up as nonce usages – can require considerable conscious deciphering on the part of the listener or reader.

Brevity is the soul of wit, and the brevity that is achieved with blends helps to adjust the length of an input signal to the faster brain processing speed that occurs when something has become expected and predictable. Reading through the whole length of a familiar word or expression takes more time than is warranted by the content gained, so the brain is bored by the time it reaches the end of the word. Making the familiar shorter means the time required to

read it is a closer match to the quality of content obtained, based on what is already familiar to the brain. The brain is further challenged by having to put the two pieces together to figure out the puzzle, and is rewarded with a self-congratulatory feeling similar to the kind one gets upon figuring out a pun or joke. Emotional engagement certainly contributes to making the word more memorable and popular. It may fail if it's too opaque, or bore if it's too easy or predictable.

If the blend happens to catch on, over time there may be less need for the average user to be as aware of the compositionality of its components, and relative opacity may grow, though not necessarily at the same rate as the fall in information value or novelty. There is however a push in the opposite direction due to our natural need and desire to assign meaning to pieces over a certain length, based on familiar patterns, habit, and also as an aid to memory. Knowing each individual piece makes it easier to place the whole word or formation in one's existing semantic network. Over time, however, with the drop in information value of the whole, phonological simplification may occur, transparency may correspondingly drop, until the point where it is hardly or no longer at all recognizable as two morphemes.

One well-known example of this is the coalescence of Old English *hlaf* 'loaf' and *weard* 'guard' into *hlāford* (Barney 1977: 36), later *lord*; *lady* has a similar kind of history: *hlæfdige* 'bread-kneader' is composed of *hlaf* 'loaf of bread' and *dige*, 'female kneader.' This simplification certainly reflected a need for frequent mention of the referent of the item, probably at a time when 'guarding loaves' and 'kneading loaves' was no longer very relevant to the idea of a 'master/leader' and 'lady'. Similarly, 'daisy' comes from *dæges-ēage* 'day's eye,' In each case, the phonetic simplification gradually obscured the original component structure, until it was finally interpreted as a simplex word. It was time to "forget" about its composition, since it no longer merited the valuable brain resources that would be required to maintain popular awareness of it.

In sum, blends in fact straddle the line between productivity and creativity. They are creative, like a pun; yet they are at the same time productive, falling into the patterns of a very strong, widespread trend, though with a higher rate of entropy or uncertainty than more rule-governed kinds of word formations. Blends help raise a falling information value, or entropy, at a crucial point in a discourse, while often maintaining the prosodic integrity of the second item used in their composition, offering the listener or reader a sense of recognition and pleasure from deciphering them. Patterns of information value rises and falls in discourse and over time are a field ripe for further study and development.

## Conclusions:

First, we should begin to treat the notion of “morpheme” as the subjective, scalar construct that it is. And rather than quibbling over what is or is not a morpheme, for example phonaesthemes, we should allow each unit of word composition that carries any degree of meaning at all to take its rightful place on the continuum of meaning-bearing units of word formation that includes but is not restricted to morphemes as traditionally defined and understood.

Second, word formation patterns that are not rigorously rule-governed but that still exhibit identifiable patterns should not be simply relegated to the “lexicon”; we should instead draw from an arsenal of different tools, such as corpus studies of text and genre analysis, to identify larger, often looser, patterns of use and composition.

Finally, we should pay special attention to the rises and falls of information value of individual word formations within discourse and over history to better understand them and extract more generalizations about their nature and how they work and fit into the larger picture of word formation.

## Acknowledgment

Big thanks to my daughter Nadia Yun Chung 鍾韻 for her editing help and for preparing the Power Point presentation of this paper.

## References:

- Ahlsén, Elisabeth. *Introduction to Neurolinguistics*. Amsterdam & Philadelphia: John Benjamins, 2006. (p. 90)
- Algeo, John. *British or American English? A Handbook of word and grammar patterns*. Studies in English Language. Cambridge: Cambridge University Press, 2006.
- Anderson, Stephen R. *A-Morphous Morphology*. Cambridge: Cambridge University Press, 1992.
- Aronoff, Mark. *Word Formation in Generative Grammar*. Linguistic Inquiry Monograph 1. Cambridge, MA: MIT Press, 1976.
- Barney, Stephen A. *Word-hoard: An Introduction to Old English Vocabulary*. New Haven and London: Yale University Press, 1977.
- Bauer 2004. *A Glossary of Morphology*. Washington D.C.: Georgetown University Press, 2004.

Bloomfield, Leonard. *Language*. With a new forward by C. F. Hockett. Chicago & London: University of Chicago Press, 1933; 1984.

Booij, Geert. "Compounding and Derivation." In: Dressler 2004.

Dressler, Wolfgang U., Oskar E. Pfeiffer, Markus Pöchtrager, & John R. Rennison, ed. *Morphological Analysis in Comparison*. Amsterdam & Philadelphia: John Benjamins, 2000.

Dressler, Wolfgang U., Dieter Kastovsky, Oskar E. Pfeiffer, & Franz Rainer, ed. *Morphology and its Demarcations*. Amsterdam & Philadelphia: John Benjamins, 2004.

Eddy, Peter A. "Foreign Languages in the USA: A National Survey of American Attitudes and Experience." *The Modern Language Journal*, Vol. 64, No. 1 (Spring, 1980), pp. 58-63  
<http://www.jstor.org/stable/324754>

"Method in text-analysis: An introduction." AV1000: Fundamentals of the digital humanities, 2007. Online course syllabus, accessed at:  
<http://www.cch.kcl.ac.uk/legacy/teaching/av1000/textanalysis/method.html#Rhetoric>

Grzeda, Joachim. "Some Thoughts on a Cognitive Onomasiological Approach to Word-Formation with Special Reference to English." *Onomasiology Online* 3 (2002).  
<http://www1.ku-eichstaett.de/SLF/EngluVglSW/grzega1023.pdf>

Joseph, Brian D. Ch. 18. "Diachronic Morphology", in Andrew Spencer and Arnold M. Zwicky, ed. *Handbook of Morphology*. Blackwell Handbooks in Linguistics. Oxford: Blackwell Publishers, 1998.

Katamba, Francis and John Stonham. *Morphology*. 2<sup>nd</sup> ed. New York: Palgrave Macmillan, 2006.

Knowles, Gerald. *Patterns of Spoken English: An introduction to English phonetics*. London & New York: Longman, 1987. pp113-116.

Lehrer, Adrienne. "Are Affixes Signs?" In Dressler, Wolfgang U., Oskar E. Pfeiffer, Markus A. Pöchtrager and John R. Rennison, ed. 2000. p. 143-154.

Lehrer, Adrienne. "Blendalicious." In Munat 2007, p. 115-133.

Libben, Gary. "Why Study Compound Processing? An overview of the issues." In Libben 2006.

Libben, Gary & Gonia Jarema, ed. *The Representation and Processing of Compound Words*. Oxford & New York: Oxford University Press, 2006.

Merriam-Webster Online Dictionary: <http://www.merriam-webster.com/dictionary/>

Munat, Judith, ed. *Lexical Creativity, Texts and Contexts*. Amsterdam & Philadelphia: John Benjamins, 2007.

O'Grady, William. *Contemporary Linguistics*. 4<sup>th</sup> ed. London: Pearson ESL, 2001.

Pierce, John R. *An Introduction to Information Theory: Symbols, Signals and Noise*. 2<sup>nd</sup> rev. ed. New York: Dover, 1961, 1980.

Plag, Ingo. "Productivity". To appear in: *Encyclopedia of Language and Linguistics* 2nd Ed., Elsevier May 17, 2004 File accessed at:

<http://www2.uni-siegen.de/~engspra/Papers/Morphology/productivity-ell2.pdf>

Pound, Louise. *Blends: Their Relation to English Word Formation*. Heidelberg, Carl Winter's Universitätsbuchhandlung, 1914. Accessed online from:

<http://www.archive.org/details/blendstheirrelat00pounuoft> p. 3-5

Sadock, Jerrold M. Review of Stephen R. Anderson, *A-Morphous Morphology*, Cambridge: Cambridge University Press, 1992. In *Natural Language & Linguistic Theory*. Springer Netherlands. Volume 13, Number 2, May, 1995. Pages 327-341.

Sandra, Dominiek & Marcus Taft, ed. *Morphological Structure, Lexical Representation and Lexical Access*. A Special Issue of *Language and Cognitive Processes*. Hove UK & Hillsdale USA: L. Erlbaum Associates, 1994.

Shisler, Benjamin K. "The Influence of Phonesthesia on the English Language", 1997.

Accessed online at: <http://www.geocities.com/SoHo/Studios/9783/phonpap1.html>

Skeat, Walter W. *A Primer of English Etymology*. Orig. pub. by Clarendon Press, 1910; reprinted by New York: Cosimo Books, 2005.

Sproat, Richard. *Morphology and Computation*. Cambridge MA & London: MIT Press, 1992.

Stekauer, Pavol. "Fundamental Principles of an Onomasiological Theory of English Word-Formation." *Onomasiology Online* 2 (2001).

<http://www.ku-eichstaett.de/SLF/EngluVglSW/stekauer1011.pdf>

Stonehill College Website. "The Dead Language That Never Dies." November 17, 2008.

<http://www.stonehill.edu/x15538.xml>

Vertosick, Frank T., Jr. *The Genius Within: Discovering the Intelligence of Every Living Thing*. New York: Harcourt, 2002.

Zwitserslood, Pienie. "The Role of Semantic Transparency in the Processing and Representation of Dutch Compounds." In Sandra & Taft 1994. pp. 341-368.