

---

From the SelectedWorks of John Haiman

---

January 2010

# The creation of new words

Contact  
Author

Start Your Own  
SelectedWorks

Notify Me  
of New Work



Available at: [http://works.bepress.com/john\\_haiman/24](http://works.bepress.com/john_haiman/24)

# The creation of new words\*

JOHN HAIMAN

## *Abstract*

*Exaptation (Lass 1990) is the recycling of meaningless linguistic “junk”, which may be recycled to create entirely new words. Two cases are examined here:*

- *There are two auxiliary verbs do in modern English. The first (Chomskyan do) has the distribution that is famously outlined in Syntactic structures. The second, a homophonous auxiliary, occurs in (negative) imperatives, and was completely indistinguishable from the other periphrastic auxiliary in Elizabethan English. It is subject to very different constraints. The result is that English now has a prohibitive, as well as a declarative negative.*
- *Khmer has an enormous number of alliterating near-synonyms which differ only in one or two segments (the coda, or the entire rhyme) of the stressed final syllable. It is possible that many of these arose as (nearly?) meaningless “servant words” that accompanied an original root in symmetrical compounds.*

*Examples of this sort may call into question two dogmas of functional linguistics: that grammaticalization is irreversible, and the widely shared view that erosion is the only productive process of linguistic change.*

## **1. Introduction**

The terms “exaptation” and “degrammaticalization”, at least as used in linguistics (Lass 1990, 1997; Brinton and Traugott 2005: Ch. 3), refer to a variety of opportunistic processes of “promotion” which create words or at least meaningful morphemes from phonetic material that may have originally had no meaning. A commonly cited example is the *-oholic* of words like *work-oholic*. The examples which Lass cited in his seminal article might lead the casual reader to infer that the semantic movement of

such recycled material is merely “sideways”, a bound morpheme of one type simply becoming a bound morpheme of another. For example, an originally inflectional morpheme becomes isolated and thus reinterpreted as a derivational one, as is the case for the nominalizing suffix *-t* in words like *pursuit*. But this is clearly not always the case, as the promotion of *-oholic*, *-gate*, *-thon* and other comparable arriviste morphemes clearly demonstrate. I am not interested in parsing the differences between degrammaticalization and exaptation here — for my purposes they can be taken as synonymous, insofar as they relate to the promotion of phonetic material to the autonomous status of independent words in the speaker’s lexicon.<sup>1</sup> But I do wish to challenge the claim that promotion in general, under whatever name, is necessarily “magnitudes rarer than grammaticalization” (Lehmann 2002: 14–15), its apparent opposite.

The present discussion concerns two cases of this kind of recycling which result in semantic “promotion” — from phonetic to morphemic or lexical status.

## 2. Origin of the prohibitive in English

In modern English, there are two periphrastic verbs *do*. The first, made famous by Chomsky 1957, is syntactically in complementary distribution with “all other auxiliaries”: that is to say, with any or all of

*Be* in all its functions,  
*have* in its auxiliary perfective function,  
 and the modal verbs.

Being meaningless, it could be considered a purely decorative “peg” form, like the notorious peg syllable in Athabaskan languages (a meaningless “last resort” prefix which occurs on all verbs which would otherwise occur with none, cf. Rice 1989: 133; Faltz 1998: 524; Cook 2004: 14). The second *do* is found in imperatives, in particular, in negative imperatives (Schmerling 1982: 203; Warner 1991: 45, 61, 89). Unlike what I will call the Chomskyan *do*, imperative *do* (both positive and negative) exhibits the following entirely non-controversial peculiarities:<sup>2</sup>

- (A) It may co-occur with *be*, cf. Sadock and Zwicky 1985: 175 (and, possibly with the *have* auxiliary as well).
- (1) a. *You are not foolish.* (\**You do not be foolish.*)
  - b. *Don’t be foolish.* (\**Be not foolish.*)
  - c. *You have left.* (\**You do not have left.*)
  - d. *Don’t have left without cleaning up.*  
           (\**Have not left without cleaning up.*)

- (B) It does not have the *raison d'être* of the (1957) Chomskyan *do*, insofar as there is no “orphaned” finite inflection that it “supports” (cf. Schmerling 1982: 203). The imperative verb in English occurs as a bare infinitival stem. This distinction is also exemplified in (1), as well as in
- (2) a. *Don't anybody move!* (\**Doesn't anybody move.*)  
 b. *Be nice!* (\**Are nice.*)
- (C) Unlike the Chomskyan *do*, it is never “available” for tag formation:
- (3) a. \**Don't have a cow, do you?* (cf. *You don't have a clue, do you?*)  
 b. *Don't have a cow, will you?* (cf. \**You don't have a clue, will you?*)
- (D) When the subject is retained, it precedes that subject (Bergs 2005; van der Auwera 2008: 19):
- (4) a. *Don't you step on my blue suede shoes.*  
 b. \**You don't step on my blue suede shoes.*  
 (ungrammatical as an imperative)  
 c. *Don't anybody move.*  
 d. \**Anybody don't move.* (ungrammatical on all readings)

It is worth noting that the sentence-initial position of negative imperative *don't* in (4) cannot be accommodated or explained by the standard English rule of subject-verb inversion (cf. Rupp 1999: Ch. 4).<sup>3</sup> This rule, as is well-known, makes a distinction between *do not* and *don't*, such that the cliticized negative moves with the auxiliary into pre-subject position in questions, while the non-reduced *not* generally does not:<sup>4</sup>

- (5) a. *Don't you like pigs?* (\**Do you n't like pigs?*)  
 b. *Do you not like pigs?* (\**Do not you like pigs?*)  
 c. *Why don't you like pigs?* (\**Why do you n't like pigs?*)  
 d. *Why do you not like pigs?* (\**Why do not you like pigs?*)

If sentence-initial negative imperative in (4) owed its position to subject-verb inversion, we would expect to encounter pairs similar to (5) in negative imperatives, but we do not:

- (6) a. *Don't you step on my blue suede shoes.*  
 b. \**Do you not step on my blue suede shoes.*  
 c. *Don't anybody move.*  
 d. \**Do anybody not move.*

Instead, the non-reduced negative is simply incompatible with any retained subject:

- (7) a. \**Do not you step on my blue suede shoes.*<sup>5</sup>  
 b. \**Do not anybody move.*

Incidentally, the pre-sentential position of the imperative auxiliary *don't* provides a small amount of empirical justification for Klima's (1964: 298–300) hypothesis that the negative is associated with this initial position at some level.

All of these facts are totally uncontroversial in English (cf. Jespersen 1940: 443, 468, 473; Visser 1978: 1539–1542; Quirk et al. 1985: 11.24), but their characterization has never been clear, nor has the contrast been generally treated as worthy of theoretical interest. Following Jespersen (1940: 504), Ellegård (1953), and Chomsky (1957), formal synchronic, diachronic and discourse studies like Klima (1964, see in particular p. 258), Stein (1990), Rissanen (1991), Tottie (1991), Kroch (1989a, 1989b); Han (1998), etc. treat *do* support as a single unified phenomenon and devote little or no discussion to the differences noted here.

Among theoreticians, Schmerling (1982) and Warner (1993) seem to be alone in explicitly proposing that the imperative *do* is simply a different word from Chomskyan *do*, and Warner characterizes the distribution of the former as “thoroughly idiosyncratic” (Warner 1993: 61). Sadock and Zwicky (1985: 175) and Bergs (2005) make the insightful suggestion that I enlarge on here, that the syntactic distribution of this second *do* happens to allow modern English to make a relatively covert distinction between factual and prohibitive negation. The *do not* of *Do not leave me* and the *do not* of *I do not love you* are different syntagms.

In proposing that negative imperative auxiliary *do (not)* is simply a different morpheme from its homophone, the Chomskyan *do (not)*, I follow Schmerling (1982), Sadock and Zwicky (1985), Warner (1993), Bergs (2005), and van der Auwera (2008). I propose further that the two auxiliaries were indeed once indistinguishable, but underwent a split which seems to have occurred over the 17th century. Ellegård (1953: 162) notes that in the English of 1600, and in fact up to that time, the frequency of incidence of *do* in negative imperatives (e.g., *Do not fear*) was almost exactly the same as the incidence of *do* in affirmative declarative sentences (e.g., *I do deny this*) — that is, extremely low, and never higher than 10% of all possible cases. The rest of the time, *do*-less versions ruled: *Fear not* and *I deny this*. Had statistical tendencies of 1600 been predictive of future analogical leveling (as, incidentally, they were for all cases of Chomskyan *do*), both would have disappeared entirely, and yet we know that this did not happen. By 1700, *do* in unmarked assertive sentences had in fact disappeared entirely yielding the modern pattern (Ellegård 1953: 162).

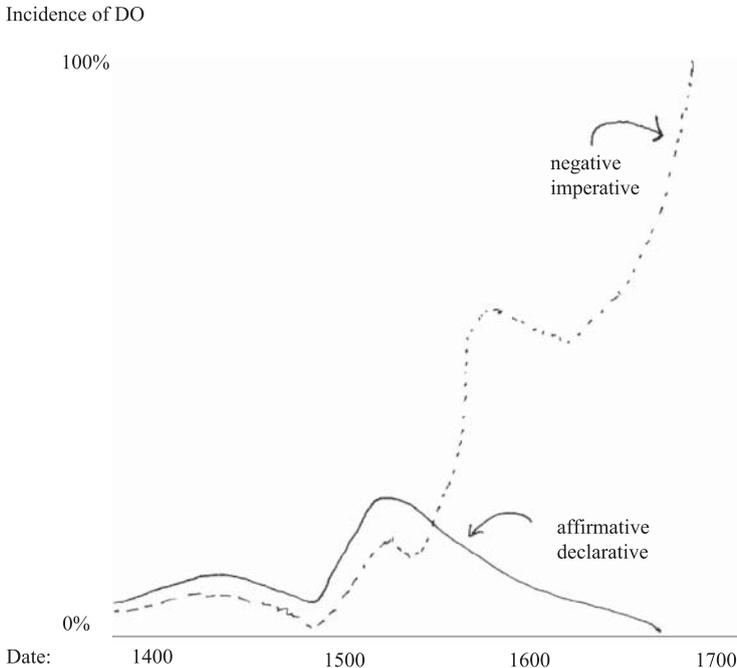


Figure 1. *DO over time* (after Ellegård 1953: 162)

But the occurrence of *do* in negative imperatives, entirely alone in bucking this statistical trend, had suddenly shot up steeply, so that its use was (and remains) virtually exceptionless:

All through the sixteenth century, the frequency [of *do* in negative imperatives] is remarkably low, but around 1600 the *do* forms gain ground rapidly, so that in the seventeenth century, the development is roughly the same as in negative declarative sentences (Ellegård 1953: 178). The type [of negative imperative] without *do* drops into disuse after the end of the seventeenth century (Visser 1978: 1542).

The same development marks the co-occurrence of *do* and *be* in negative imperatives. Visser (1978) notes only two examples in all of Shakespeare, and a third in Beaumont and Fletcher's *Knight of the Burning Pestle* of 1613. After 1693, the attestations of *don't be* are dense. For some reason, *do* in imperatives was subject to different tendencies from the Chomskyan *do*, eventually becoming a form so distinct from Chomskyan *do* that one authority asserts that "there is no worthwhile generalization" to be made about these two forms at all (Warner 1993: 90). How did it become so "thoroughly idiosyncratic"?

It is notable that these 17th century developments, leading to the facts summarized in (1) to (4), have allowed English to join most other languages in having a distinct expression of the prohibitive, or the negative in imperatives. However, since “imperative *do*” occurs in both positive *and* negative imperatives, we cannot yet call “imperative *do*” part of a dedicated prohibitive construction. But there is solid evidence that imperative *do* in positive imperatives like (8a) is a later backformation:

- (8) a. *Do (you) follow me.*  
 b. *Don't (you) follow me.*

That is, (8a) was constructed by analogy with the pattern (8b) in the negative imperative. As we know, only the form without the subject pronoun is now acceptable:

- (9) a. *Do (\*you) follow me.*  
 b. *Do (\*you) be quiet.*

The evidence for the back-formation hypothesis is partly semantic. Chomskyan periphrastic *do* in so-called emphatic sentences is generally (if not always) synonymous with a double negative. Sentences like

- (10) a. *I do (so) like pigs.*  
 b. *I do (so) believe you.*

and others like them are typically uttered in response to an expressed or (perhaps only tenuously) implied denial (e.g., *You don't like pigs*), and their purpose is to deny the denial. Thus Chomsky's EMPH (ex hypothesi =NEG NEG) morpheme is plausibly moved to the same site as NEG. But double negation is simply not the meaning of *do* in emphatic sentences like (9). Rather, they have a hint of a “polite British” flavor, in making an “urgent [but softened] request or entreaty” (Jespersen 1940: 504), quite unlike (10).<sup>6</sup>

And partly the evidence for the backformation hypothesis is historical. Ellegård (1953) notes that while the occurrence of *do not* in imperatives was low until 1600, the occurrence of *do* in positive imperatives was virtually nil, in fact comparable to the distribution of *do* with *to*-infinitives. He enumerates all cases before 1600 that he unearthed in a mere two pages (Ellegård 1953: 175–176), and notes that the rapid spread of negative imperative *don't* throughout the 17th century “should naturally be seen against the background of the near absence of affirmative imperative *do*” (Ellegård 1953: 178).

We can then account for *do* in modern positive imperatives like (10) by saying that it is a purely formal — and later — analogical generalization

of *do* in negative imperatives. And it is primarily the genesis of this latter structure which engages our attention.

To answer the question what possible motivation there may have been for the special behavior of *do* (resulting in a separate homophonous form) in the negative imperative, it is clear that comparison with other languages will yield little, as long as we seek analogues of the periphrastic *do* itself. As is well-known, periphrastic *do* is a relatively exotic curiosity of English (cf. Dahl 1979: 85).

But the comparative study of negatives in general may be more rewarding. As noted most forcefully by Wierzbicka (1972: Ch. X), but also by a number of other students of typology (Bhat 2004: 1209; Croft 1991: 14–15; Horn 1989: 447; Payne 1985: 223; Sadock and Zwicky 1985: 175; Schmerling 1982; and most recently van der Auwera et al. in their entry in Haspelmath et al. 2005) a large number of languages, perhaps even the vast majority, make an overt distinction between factual (Latin *non*, Greek *ou(k)*) and “prohibitive” (Latin *noli*, Greek *mee*) negatives. Languages around the world from Cree to Dyrbal seem to give overt expression to this difference between “say that . . . not” and “want/order that . . . not”. (In languages like Greek and Korean, the two are expressed by totally unrelated morphemes, while in others like Latin and Hungarian, the two morphemes are clearly related, but different.) The widespread existence of such a distinction does not call into question the fundamental relationship between the two negatives, nor make it surprising that a contrast between them could fail to be expressed, and (unlike Wierzbicka 1972) I do not have an interest in trying to “derive” one of these from the other semantically. But it seems plausible that once they had the periphrastic *do* to play with, speakers of 17th century English created exactly the same very widespread *non/noli* distinction covertly, resulting in the distinct patterns noted in (1)–(4), (7).

Indeed, it may be that English is not the only language to enforce this distinction in a relatively covert fashion. As Sadock and Zwicky (1985) point out, three quarters of the languages in their sample distinguish factual and prohibitive negation in some way — a statistical assessment that is precisely echoed in van der Auwera et al.’s entry in the much larger *World atlas of linguistic structures* of Haspelmath et al. 2005. This distinction is made even in other modern Indo-European languages which seem to get along without it. The inherited Indo-European prohibitive particle *mV-*, found in Ancient Greek *mee*, apparently survives today only in some Iranian languages like Kurdish and Zazaki (e.g., Sandonato 1994: 134), and the Latin negative verb *noli* has been entirely lost. But . . .

The Romance languages, although they lost the morphological *non/noli* distinction, often behave as if they retained it syntactically (cf.

Zanuttini 1997). Italian, for example, uses only the one negative morpheme *non* in both statements and commands, but in commands, the verb is in the infinitive (or, in the polite form, the subjunctive) mood, exactly as if the verb were acting as the complement of a higher verb:

- (11) a. *Non fuma-re*  
not smoke-infinitive (familiar imperative)  
b. *Non fum-i*  
not smoke-3sg.subjunctive (polite imperative)  
c. (*Tu*) *non fum-i* (familiar 2sg. indicative)  
d. (*Lei*) *non fum-a* (polite 3sg. indicative)

The Slavic languages (like Germanic) apparently never had the morphological distinction between constative and prohibitive negation, but Bulgarian, like English, seems to have home-grown its own, and in much the same way, by using a verb meaning ‘do’. In addition to the inherited *ne* negative, there is an innovative negative imperative compound *ne + dej* ‘not +do + imperative’ followed by the bare verb stem (Bulgarian has no infinitive) which is possible in both familiar and polite imperatives (Antonova et al. 1997: 196–197).

- (12) a. *Ne-de-j* (-te) *plaka* (novel form)  
not-do-imperative (plural) cry  
b. *Ne plače* (-te) (inherited form)  
not cry.imperative (plural)

The Bulgarian negative imperative “verb” \**dej-* (currently attested only in derivatives like *dejno* ‘actively’, *dejanie* ‘deed’, *dejstvuvam* ‘(I) act’, and *dejstvie* ‘action’) of (12a) is cognate with *do*,<sup>7</sup> which makes *nedej* the literal equivalent of the English negative imperative *don’t*, but there is of course no Chomskyan *do* in the language.

The specific negative imperative function of the *don’t* auxiliary is implicitly and unofficially recognized by many editors of bilingual dictionaries of languages from Cree (Wolfart and Ahenakew 1998: 33) through Bulgarian (Atanassova et al. 1988: I, 502) to Nunggubuyu (Heath 1982: 259), who simply gloss their prohibitive particles as ‘don’t!’. Unlike Chomskyan *do*, which serves a purely decorative function, imperative *don’t* makes a meaningful contribution to the sentences in which it appears.

It is impossible to draw from these facts the conclusion that the imperative *do(n’t)* in English arose because speakers absolutely had to mark a distinction between prohibitive and factual negation. There are many languages (including presumably earlier forms of English) which seem to manage without it. But perhaps such languages are less numerous than

we think, fewer even than the 23% recognized in van der Auwera et al.'s (2005) sample from the *World Atlas of Linguistic Structures*. After all, Modern English itself would once have seemed to be such a language (not even Jespersen seems to have remarked on the significance of the distinction), and in general the extent of the phenomenon may be generally underreported. We note that typological studies of negation in Kahrel and van den Berg (1994) overlook its existence in two of the languages described in chapter-length studies (Hungarian and Mandarin), although the distinction in these languages is morphologically overt (Mandarin factual *bu* vs. prohibitive *bie*, Hungarian factual *nem* vs. prohibitive *ne*). Thus we should consider the further possibility that it may be worthwhile to carry out more work to unearth the contrast in other languages where the distinction may exist covertly. It is certainly worth noting that even morphologically impoverished languages like Cambodian, which mark virtually no grammatical categories (and do so half-heartedly), enforce this one without exception: the factual preverbal particles are *pum* (< *pum* 'reject'), *awt* (< *awt* 'lack'), *mwn*, occurring in almost entirely free variation, but the prohibitive *kom* (< *kom* 'shut') is reserved for negative commands, and a few related irrealis functions.

The remarks in this section have been a confirmation of and enlargement on insights of Warner 1993 (that there are two periphrastic *do*'s in English), and of Schmerling (1982), Sadock and Zwicky (1985), Bergs (2005), and van der Auwera (2008) (that the distributional contrasts between the two enable English to distinguish factual and prohibitive negatives). But they also provide evidence for a relatively unscrutinized mechanism for the creation of grammatical categories. It has been a commonplace of linguistics since Bopp (in fact, since the 1500s) that such categories can be created by the phonetic and semantic erosion or "demotion" of words — this is the conventional meaning of grammaticalization. In fact, many of the negative imperative morphemes that have etymologies derive via grammaticalization from main verbs such as 'close' 'lack', or 'reject' (thus Cambodian) or 'stop' (cf. Dixon 1972: 112; Heine and Kuteva 2002: 283–284; van der Auwera 2006), and the ultimate origin of the Chomskyan auxiliary *do* from the main verb meaning 'make' or 'perform' is also uncontroversial.

What we see in the cases of English and Bulgarian, on the other hand, is the creation of a negative imperative morpheme via a novel kind of bricolage, or tinkering (cf. Jacob 1977; Lass 1990, 1997). In the course of the 17th century, a meaningless morpheme with a random, later a purely syntactic, distribution was "available" for exploitation. Speakers started putting it to a new use and created the prohibitive, a previously unattested grammatical category in English.

### 3. The origin of “servant” words and near synonyms in Khmer

The lexicon of Khmer is dense with alliterating near-synonyms like the following (culled from Headley et al. 1977 passim):<sup>8</sup>

- (14) a. ‘short and fat’
- |                           |                                |
|---------------------------|--------------------------------|
| <i>Kawntanj</i>           | ‘short and fat; stocky; broad’ |
| <i>Kawntaw:</i>           | ‘small and fat’                |
| <i>Kawnthat</i>           | ‘fatty’ (pejorative)           |
| <i>Kawnthia</i>           | ‘small and broad’              |
| <i>Kawnthok</i>           | ‘obese’                        |
| <i>Kawnthoc</i>           | ‘fleshy, fat’                  |
| <i>Kawnthol</i>           | ‘plump, fat, obese’            |
| <i>Kawnteunj kawntonj</i> | ‘short, squat, fat’            |
- b. ‘curved’
- |                       |                                |
|-----------------------|--------------------------------|
| <i>Kngiang</i>        | ‘bent to one side’             |
| <i>Kngong</i>         | ‘bent like a hook’             |
| <i>Kngawng</i>        | ‘very curved’ (said of sticks) |
| <i>Kngawk</i>         | ‘bent, curving’                |
| <i>Kngol</i>          | ‘bent over, stooped’           |
| <i>Kngeung kngang</i> | ‘bent out of shape’            |
- c. ‘in one’s sleep’
- |                          |                                     |
|--------------------------|-------------------------------------|
| <i>Maumial</i>           | ‘hallucination as one falls asleep’ |
| <i>Maumeu:</i>           | ‘walk, talk in one’s sleep’         |
| <i>Maume:ng</i>          | ‘dazed’                             |
| <i>Maumeu: maumiaj</i>   | ‘talk in one’s sleep’               |
| <i>Maumeung maumeang</i> | ‘half asleep’                       |
| <i>Maumi: maumeu:</i>    | ‘dazed’                             |
- d. ‘torn’
- |                        |                       |
|------------------------|-----------------------|
| <i>Raujaj</i>          | ‘torn apart, in rags’ |
| <i>Raujej</i>          | ‘badly torn, ragged’  |
| <i>Raujeah</i>         | ‘torn apart’          |
| <i>Rauji:k raujiak</i> | ‘ragged, tattered’    |
| <i>Raujiak</i>         | ‘torn apart, ragged’  |
- e. ‘dangle’
- |                           |                            |
|---------------------------|----------------------------|
| <i>Rauja:</i>             | ‘dangling earrings’        |
| <i>Raujee:ng raujo:ng</i> | ‘dangle’                   |
| <i>Rauji:ng raujo:ng</i>  | ‘droop, hang unevenly’     |
| <i>Raujian</i>            | ‘hang down, droop, dangle’ |
| <i>Raujo:n</i>            | ‘hang down, dangle’        |
| <i>Raujum</i>             | ‘droop, hang down’         |

- f. ‘sloppy’  
*Rauta:j* ‘stretch out sloppily, hang down low’  
*Rauta:k* ‘hang down sloppily’  
*Rautee:k rauta:k* ‘very sloppy’  
*Rautee:j rauta:j* ‘very long, dragging down’  
*Rautaek rautaok* ‘very sloppy’  
*Rautaeng rautaong* ‘hang down unevenly’  
*Rautaok* ‘hang down unevenly, droop’
- g. ‘squirm’  
*Rauvi:k rauveu:k* ‘twitch, squirm’  
*Rauvi:m rauviam* ‘wiggle; be scarred all over; confused’  
‘unclear’  
*Rauviam* ‘squirm, wriggle’  
*Rauvwc* ‘wag, swish, flutter’  
*Rauvwc rauviam* ‘bother, annoy’  
*Rauveu:k* ‘squirm’
- h. ‘limp’  
*Rauvee:t rava:t* ‘soft and flexible’  
*Rauvaet rauvaw:t* ‘soft and flexible’  
*Rauveuk rauvawk* ‘soft and flexible’  
*Ravi:t rauviat* ‘fragile, feeble, limp’  
*Rauvwn rauvoan* ‘limp, unsteady’
- i. ‘bumpy’  
*Raukhonj* ‘rough, rugged’  
*Raukheunj raukhonj* ‘rugged, rough’  
*Raukwp raukup* ‘bumpy, rough’  
*Rauki:nj* ‘warped, uneven’  
*Raukwng* ‘pocked, covered with pimples’  
*Raukup* ‘bumpy, rugged’  
*Raukeu:l* ‘rocky, bumpy’  
*Raukee:ng rauko:ng* ‘warped’  
*Rauko:l* ‘bumpy, rocky, rough’

There is widespread disagreement among native speakers as to the existence and the meaning of these forms. The examples above are drawn from the standard bilingual Khmer-English dictionary, which is based on the Buddhist Institute’s monolingual Khmer dictionary by Chuon Nath (1938), supplemented by other dictionaries. Other Khmer wordlists, like that of Chun Leuh (2007), lack some of these forms, but include many others that are not included in Headley. Finally, two consultants I have

worked with over the last twelve years recognize partially overlapping but non-identical sets of such words.<sup>9</sup> The fact that native speakers' familiarity with these differs so much more than for the "normal" vocabulary, suggests that perhaps they belong to the class of original or nonce creations (Paul 1995 [1880]: Ch. 9): that is, they have neither a common etymology, nor have they become completely established in the common core vocabulary.

In this, they seem to be reminiscent of the ideophone sets of languages like Korean (Diffloth 1972), where a series of semantically related reduplicated words are produced by ablaut alternations (both vocalic and consonantal) of a common root. However, there is no productive vocalic or consonantal ablaut process in Khmer for the generation of such forms, and possibly for this reason, many native speakers consider these forms to be unanalyzable word roots — whose close phonetic resemblance to each other must be treated as coincidental. Familiar considerations impel us however to consider the members of each family as flakes off a common core.

The novel mechanisms whereby these sets may have been generated in Khmer are in fact partially open to inspection.

First, it must be observed that the near-synonym sets above include both monolexemic words and a number of "Symmetrical Two-word Expressions" (henceforth STE) that also consist of alliterating roots: *Kawnteunj kawntonj* 'short, squat, fat', *Maumeu: maumiaj* 'talk in one's sleep', *Kngeung kngang* 'bent out of shape', *Raukheunj raukhonj* 'rugged, rough'.

Alliterating STE are extremely numerous in Khmer, and may constitute as many as 10% of the lexical items in the language. Semantically, they be assigned to three classes:

- (a) the two words are (or once were) near-synonyms: these are then genuine compounds like English *kith and kin*.<sup>10</sup>
- (b) only one of the words has meaning: the other (which may precede or follow its companion) has no independent meaning, and is called a *bo'ri'va: sap* or "servant" word,<sup>11</sup> as in the English twin forms *jibber jabber*, *mish mash*.
- (c) the compound is a single word neither of whose parts is (now) meaningful or can occur by itself, like English *spic and span*, *hoity-toity*, or *higgledy-piggledy*.

Judgments on which class these examples belong to are also fluid. For some native speakers, a given two-morpheme alliterating expression may be a synonym compound, for others a word + servant word combination (different speakers may also disagree on which is the genuine word, and

which is the servant), while for yet others, the whole is an unanalyzable single unit.<sup>12</sup>

In Ourn and Haiman (2000), where we listed several hundred such STE, we assumed that grammaticalization or semantic erosion was the sole historical origin of such pairs. Following conventional wisdom, we supposed that meaningful compounds like *last and final* could become fossilized particularly easily when its members were already similar, so that one, or in some cases, both parts lost their original meaning and capacity for occurring as independent words (cf. also Wälchli 2005). In other words, we posited a grammaticalization trajectory (a) > (b) > (c). This of course is the history of expressions like *kith and kin*.

In more recent work we have come to question our original conclusion, and in Haiman and Ourn (2009) and Haiman (2008) we have provided three arguments that in many such symmetrical compounds, the direction of development was the opposite one: a meaningless partner is somehow generated from “whole cloth” for a base word, on the basis of criteria that are originally purely phonological. Given the independently attested *Drang nach Parallelismus* in Khmer (Ourn and Haiman 2000; Haiman and Ourn 2003, 2009),<sup>13</sup> it is not surprising to find that a root *X* may accompany a typically alliterating near identical *x* — a companion which may be assembled or conscripted from a variety of other sources. We distinguished a number of source strategies for creating decorative symmetry, of which two are repeated here.

The first, or “conscription” strategy is one whereby a root may (apparently) simply hijack an alliterating word whose independent meaning is completely irrelevant. Thus *bawnlawm* ‘confuse’ occurs with the servant word *bawnlae* ‘vegetables’, to form the decorative pair *bawnlae bawnlawm* ‘confuse’.

The second, or “Procrustean” adaptation strategy is the one whereby the word occurs with a near synonym, which is however tricked out with (or, less frequently, deprived of) extra and meaningless phonetic material, with the result that it alliterates with the dominant word. Thus the word *prawhael* ‘about, approximately like’ is nearly synonymous with *hak* ‘like, similar’. In a symmetrical coordination, the latter word appears with a meaningless “prefix” *praw*.<sup>14</sup> The resulting decorative compound *prawhak prawhael* means ‘about, approximately like’. See Haiman and Ourn (2009) for many examples of conscription and a smaller number of examples of adaptation.

I propose the following scenario for the plethora of near-synonyms exemplified in (14). First, a word is accompanied by a decorative servant word. Later, this originally decorative partner may take on a semantic and syntactic life of its own, and eventually become a separate word: *Xx*

(a decorative compound) becomes *X* and *x* (two separate dictionary entries for synonymous words), and later *X* and *Y* (two separate words). The density of near synonyms in Khmer would then be explained by the prevalence of servant word STE's, whose meaningless member may come to stand on its own: originally as a nearly meaningless decorative echo, then a perfect synonym of the word it once accompanied, and then later, via the well-known process of "repartition" (Bréal 1897; Lass 1997: 342–352) as a near synonym.

One of the main questions that faces us is the origin of most of the servant word members of STE, like the initial member of *Raukheunj raukheunj* 'rugged, rough'. Strategies like "Conscription" and "Procrustean adaptation" may not take us very far, because in general, most servant words cannot thus be traced back to other extant words. Recall, moreover, that there *is* no productive mechanism of reduplication via vocalic or consonantal ablaut for the production of these minimally differing forms, which is certainly one plausible reason why so many speakers insist on their unanalyzability, in spite of their suspiciously common occurrence.

At least one further novel mechanism does however exist for the production of such forms — perhaps for the majority. There is a well-known rhyme-swapping verbal game in Khmer (similar to the creation of playful Spoonerisms in expressions like *tee many Martoonis*) known as *piak kunloah kat* (translated by Stephanie Farmer as "word slice and dice"). In this game, the initial and final words in a compound or a common expression swap the codas of their stressed (final) syllables. The ludic function of this game is patent, and insisted upon by all consultants. (One possible payoff is the disguising of taboo forms, cf. Hoonchamlong 1991: 12 for Thai.) In addition, the game functions as a kind of "intelligence test" (Noeurng Ourn p.c.) inasmuch as "points" are awarded for lapidary deformations that turn out to be surprisingly apt or true. For example, the expression

- (15) a. *Kru: prawpaun l'aw:*  
 Teacher wife beautiful  
 'The teacher has a beautiful wife' can become
- b. *Kraw: prawpaun l'u:*  
 Poor wife complain  
 'The poor man's wife complains.'

Both utterances are meaningful, and the combination of the two makes a rueful comment on how sex and status go together. But in most cases, the result of the deformation is simply nonsense:

(16)	Common Expression	Result of Rhyme-Swapping
a.	<i>Ps- a: mian c- ao</i> Market have thief 'There are thieves in the market'	→ <i>ps- ao mian c- aa</i> (nonsense) have (nonsense)
b.	<i>Tul-iaj th-om</i> Roomy big	→ <i>tul-om th-iaj</i> (nonsense) (nonsense)
c.	<i>J-om sr-ae</i> Weep cry out	→ <i>j-e:k sr-om</i> (nonsense) (nonsense)
d.	<i>Rauh-ae</i> <i>proh th-oj</i> Tear because soft 'soft and frayed'	→ <i>rauh-oj proh th-ae</i> (nonsense) because (nonsense)
e.	<i>Rauh-iang daoj pum</i> Indistinct by not <i>prakaw:d bawnt-euc</i> exact a bit 'hard to make out exactly'	→ <i>rauh-euc daoh pum</i> (nonsense) by not <i>prakaw:d bawnt-iang</i> exact (nonsense)
f.	<i>R-eut cawmnaw:ng</i> Tighten knot <i>poat craeun tr-uat</i> around much repeat 'tie a rope tightly around'	→ <i>r-uat cawmnaw:ng</i> (nonsense) knot <i>craeun tr-eut poat</i> around much (nonsense)
g.	<i>R-uk daoj co:l</i> Invade by enter <i>tauntr-ian</i> invade	→ <i>r-ian daoj</i> (nonsense) by <i>co:l tauntr-uk</i> enter (nonsense)
h.	<i>rum'-awk cae pawprae</i> complain grumble <i>t'o:nj t'-ae</i> complain	→ <i>raum'-ae cae pawprae</i> (nonsense) grumble <i>t'onj t'-awk</i> (nonsense)

The novel *piak kunloah kat* 'cut up words', (cf. Ieu Koeus 1995[1945]: 286–291) generated in this way are, then, usually nonsense.<sup>15</sup> But sometimes these nonsense forms are dragooned into a higher service, so to speak: in English, we are familiar with some cut up blends as semantically motivated portmanteaux: *smog*, *brunch*, *animatronic*, *cyborg*, *telethon*, and *mimbo* (= *m-ale* + *b-imbo*, courtesy of Seinfeld). In Khmer we may encounter them much more frequently as servant words. *Psao* does not denote a semantic blend of 'market' and 'thief': but *Psa: psao* is a synonym of *psa:*. The nonsense word *tulom* does not denote a blend of *thom* 'big' and *tuliaj* 'roomy', but *tulom tuliaj* is a synonym of *tuliaj*. So too *jom je:k* 'weep' is a synonym of *jom*, *rawhaek rauhoj* 'tear' a synonym of *rauhaek*, *rauheuc rauhiang* 'hear vague rumors' a synonym of *rauhiang*, *reut*

*ruat* ‘tighten around’ a synonym of *reut*, *ruk rian* ‘invade’ a synonym of *ruk*, *ruap ruam* ‘unite’ a synonym of *ruam*, and *rum’ae rum’awk* ‘complain’ a synonym of *rum’awk*.

Chun-Leuh (2007: 190–277), following Sisowath (n.d.), attempts to account for *all* servant words in this fashion, and not all of his ingenious “etymologies” seem equally plausible to native speakers. Some of them, however, are accepted by these native speakers as plausible, insofar as the “source” is a common expression, and they are familiar with this deformation as the source from their own personal experience (which includes having been so taught).

In Ourn and Haiman (2000), we attempted to account for the typologically odd fact that servant words in Khmer and other SE Asian languages are almost always alliterating forms (like *spic ’n span*, alike at the beginning), rather than “rhyming” (like *higgledy-piggledy*, alike at the end), as in most languages. (Ablauting forms, like *jibber jabber* are alike in both onset and final segments, and thus simultaneously both alliterating and “rhyming”. These are common in languages of both types.) We claimed that this tendency was consonant with the fact that Khmer was a prefixing and infixing language, and proposed that the prevalence of prefixation would favor alliteration, while the prevalence of suffixation (the situation in most languages) would favor rhyme. In both cases, contrastive focus would rest on those parts of the word that were most likely to receive stress, and to differ segmentally. I am not prepared to reject this as a possible explanation, but a simpler one is now available. If servant words are largely generated by rhyme swapping, as Sisowath (n.d.) and Chun-Leuh (2007) propose, then it follows of course that they will alliterate, rather than rhyme, with the source word.

Once generated (through conscription, adaptation, or via the *piak kun-loah kat* word game), junior servant words can gradually achieve both semantic and syntactic independence. Let us assume that initially they are both entirely meaningless (decorative), and confined to one combination (like the notorious *cran-* of English). The language offers us as number of transitional types between such a totally dependent status and the autonomous entries of a dictionary.

First, there are junior words which have no syntactic independence, but which do make a semantic contribution in the symmetrical compounds where they occur. Consider the triplet:

- (17) a. *tmeung*  
       ‘immobile’  
       b. *tmeung tmeu:j*  
           immobile  
           ‘immobile, unresponsive’

- c. *tmeung tmeang*  
 immobile  
 'ignore, act innocent, play dumb'

(It is possible that in [17c], the servant word is a procrustean adaptation of yet another servant word \**meang*, which occurs in the symmetrical compound *mi:ng meang* 'ignorant'.)

Note that while neither *tmeu:j* nor *tmeang* can occur alone (they are cranberry morphs), they do make a semantic contribution to the compounds (17b) and (17c), which are not totally synonymous. There are many servant words of this sort.

Now consider the triplet

- (18) a. *nwaj*  
 'tired'  
 b. *nwaj na:j*  
 tired  
 'mildly bored'  
 c. *cwn na:j*  
 sick  
 'sick and tired, bored to death'

It is widely agreed that *na:j* is a servant word.<sup>16</sup> But note that in addition to appearing with the alliterating *nwaj* (a Thai borrowing meaning 'tired') in (18b), it also occurs as an accompaniment to a totally different root in (18c). It is still a servant word, but with a new master. Again, there are many servant words of this sort.

Another alternative scaffolding in which a servant word may occur is a four-word symmetrical phrasal compound. The decorative compound

- (19) a. *raubo:t raubeh*  
 slip out

consists of a meaningful root and a decorative word (which may be a procrustean adaptation of a hijacked root *beh* 'pick, pluck'). But *raubo:t* occurs in an idiomatic expression

- b. *kham cee:k raubo:t*  
 bite banana slip out

'[be so old and toothless that when one] bites [on a] banana [it] slips out'. This idiomatic expression may be conjoined with a meaningless decorative partner in

- c. *kham cee:k raubo:t kham po:t raubeh*  
 bite banana slip out bite corn  
 'be old and toothless (=19b)'.

The decorative word *raubeh* occurs not only with another word (as in [19a]), but in a four-word expression (as in [19c]). These are steps towards syntactic (as well as semantic) autonomy.

Finally, a servant word may entirely replace the meaningful word with which it was originally partnered. One might compare this with an understudy emerging into a performing role. I will provide two examples of this. In conventional Khmer, *jo:n* exists only as a servant word with *jau:k* in the expression

- (20) a. *jau:k jo:n kawmmaeut*  
           take       birth  
           ‘start life’

But my consultant Mr. Ourn was (only) mildly scandalized to encounter it in

- b. *jo:n kawmmaeut*  
    ‘start life’

A decorative word has here replaced its master.

In conventional Khmer, *lia* ‘take leave’ has been conscripted to act as the servant word of *lu:t* ‘great-’ in expressions like

- (21) a. *cav lu:t lia*  
           grandchild great

but (21a) is not yet common. What does exist in common parlance is the four-word expression:

- b. *cav lu:t cav lia*  
    grandchild great grandchild  
    ‘great- and greatgreatgrandchildren’

What may be just around the corner is a contrast between

- c. *cav lu:t* ‘greatgrandchildren’  
   d. *cav lia* ‘greatgreatgrandchildren’

If this happens, a decorative word will have come to contrast semantically with its erstwhile master.

It is interesting that the Khmer phenomenon comes closest to exaptation in the biological literature, where novelty arises from modifications of redundant and imperfect copies of some basic structure, and then is put to novel uses (Lass 1990: 80, citing Gould and Vrba 1982: 10–11; cf. Mayr 2001: 38 *passim*).

#### 4. Conclusion

Linguistic exaptation, two very disparate examples of which have been presented here, is not a unified process: rather, it is the label given to the end result where, by whatever mechanism, originally redundant or meaningless phonetic strings have been opportunistically recycled into morphological material.<sup>17</sup> And indeed, the mechanisms that were employed in these two cases, to the extent that we can comment on them at all, are surely distinct.

The reanalysis of minimally contrasting near-synonyms as distinct words with different meanings in Khmer, may be “explained” by the cognitive drive to try to find differences of meaning whenever there is a difference in form between two signs (Bréal 1897),<sup>18</sup> I have commented on a similar motivation for comparable changes elsewhere (Haiman 1998; Haiman 2003; Haiman and Ourn 2003). The reanalysis of the inherited Romanian infinitival suffix *-re* as a nominalization (*fi* ‘to be’, but *fi-re* ‘existence’, etc.), the reanalysis of the inherited Khmer anacrusic-syllable coda *-Vm* (a phonetic string without meaning which is typically elided in casual speech) as a nominalization or a causative morpheme (*criang* ‘sing’, but *c-awm-riang* ‘song’; *slap* ‘die’, but *s-awm-lap* ‘kill, etc.) may be compatible with three universal facts:

- (a) Languages do seem to swarm with synonyms and near-synonyms;
- (b) Often in conformity with the drive to eliminate purposeless variation, the synonyms are semantically or stylistically distinguished (cf. Bolinger 1975: 422 for pairs like *ignorant/igernut*, *partner/pardner*);
- (c) A plausible basis for making the distinction is that often “more form” signals more meaning (cf. Haiman 2008).

In the cases just alluded to, phonetic alternants exist simply as the result of careful and casual pronunciations of the same words. In the case of the servant words discussed in Section 3, the alternants may be motivated by a pressure for decorative symmetry, and exist as a consequence of entirely different processes, but once they exist for any reason, such forms are available for exaptation.

But the mechanics of the reanalysis of the auxiliary *do* in English remain mysterious. One can grope for an explanation. For example, one could note that the auxiliary *do* still had a vestigial semantic function in Elizabethan English (cf. Stein 1990), as illustrated by a handful of minimal contrast pairs like Othello’s “not that you love me not, but that you do not love me”. One could assert that the “active” meaning of *do* in this

case was what was exploited in the creation of the prohibitive in English, and a corresponding use of the cognate root in Bulgarian. But all we can say for sure is that the end result of a number of changes over the 17th century English was this: there was once a single periphrastic verb *do*, whose distribution in 1600 AD was more or less stochastic, and whose meaning was close to nil. There were by 1700 AD two contrasting periphrastic verbs *do*, whose distribution is fixed. The first is a meaningless purely grammatical cipher, while the second is part of a novel prohibitive marker.

How does the existence of exaptation comport with the widespread (but contested) dogma that the opposite process of grammaticalization is a unidirectional and irreversible process (e.g., Heine and Kuteva 2002: 4 and the extensive literature cited there)? One could argue, in favor of that doctrine, that in fact the two processes have nothing to do with each other: new morphological material that is built from phonetic rubble is not identical to its diachronic source. Nor is it in any of the cases examined here. (So, for example, the English prohibitive has nothing in common semantically with the main verb that is synonymous with ‘make’; the rubble which creates new words in Cambodian originates as the result of conscription, adaptation, or a verbal game.)

It is as if the semantic “ricochet” of exaptation did not fly back directly into the face of its source.

In the development of *do* in English, grammaticalization reduces a meaningful main verb to the familiar Chomskyan auxiliary, and exaptation makes of this auxiliary a meaningful prohibitive. The claim that

Source:

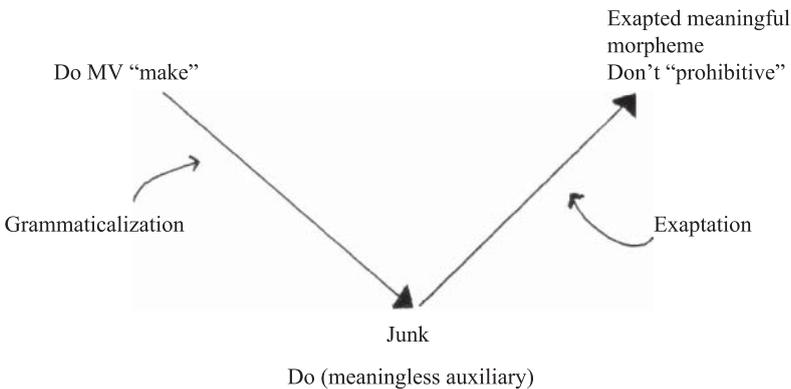


Figure 2. *The trajectories of grammaticalization and exaptation*

grammaticalization is unidirectional is tantamount to the claim that “the ricochet” of exaptation can never retrace the grammaticalization path. If, however, grammaticalization is unidirectional only because speakers wish to be novel (Haspelmath 1999), then nothing precludes a direct ricochet after the original meaning of the junk in question is no longer remembered.

But this may be no more than a statistical accident. Once junk is truly junk, there is no memory of its semantic origins, and thus no a priori reason that it could not be opportunistically recycled to have any kind of meaning at all — including, trivially and quite by accident, meanings that it had once long before.

And how does the nature of exaptation comport with Lehmann’s earlier cited dictum that grammaticalization is “orders of magnitude” more common than its converse? It may be that the contrast between the two is not one of frequency but of regularity. Grammaticalization, like analogy, like sound change, has its worn and familiar grooves which have been discussed in great detail in recent literature (cf. Bybee et al. 1994; Heine and Kuteva 2002). We can not talk of laws, but there are definitely regularities. Exaptation, on the other hand, is above all an opportunistic process: there is no predicting how a particular piece of junk will get to be recycled, if it does at all. This may lead to an appearance, but perhaps only the appearance, of infrequency.

The critical fact is that exaptation (or degrammaticalization or whatever else one calls it) can be recognized and set against sound change and grammaticalization as a process which builds up, rather than tears down, linguistic structures. This is surely a good thing: Any realistic uniformitarian model of language change must recognize both kinds of processes.

*Received 18 March 2008*

*Revised version received*

*12 February 2009*

*Macalester College*

## Notes

\* Correspondence address: Linguistics Program, Macalester College, 1600 Grand Avenue, St. Paul, MN 55105 USA. E-mail: haiman@macalester.edu.

1. It is interesting that Lass 1997, the apostle of exaptation, does not mention regrammaticalization or degrammaticalization at all, and that Brinton and Traugott barely mention exaptation in passing (2005: Fn 80), as a synonym of Greenberg’s “regrammaticalization”. There may be useful distinctions one could make between the two, but I

- will not attempt any. I use “autonomy” in the same way as do Bybee and Brewer (2005 [1980]: 13).
2. A pitfall in writing about basic English syntax is that every path has been so well-trodden that it is unlikely that the investigator will make any new discoveries. The infinitesimal smidgen of novelty in the discussion of English in this section (if indeed there is any) is grossly outweighed by the prodigious amount of prior research that is still only partially recorded in the references. My thanks to Larry Horn, Tania Kuteva, Roger Lass, Jerry Sadock, Susan Schmerling, and Anthony Warner, for discussions, examples, counterexamples, and references. And my apologies to the legions of pioneers whose work in this area I have inadvertently missed and thus inadequately acknowledged.
  3. Appearances to the contrary. Susan Schmerling (p.c.) notes that in sentences like
    - a) *Don't anybody be late*
    - b) *Don't anybody have let the cat out**Don't* seems to act like the first auxiliary in being the only one subject to subject-verb inversion, exactly as in interrogatives that have undergone inversion. Nevertheless the objection to a unified treatment still stands. It is impossible to leave uncontracted *not* “behind” in such sentences:
    - c) *Did anybody not love him?*
    - d) *\*Do anybody not have let the cat out.*
  4. Larry Horn has pointed out (p.c.) that heavy NP shift may create structures like
    - a) *Do not all the considerations we have been discussing militate strongly against that solution?*
    - b) *Does not the very existence of all these “isms” indicate that, by his very nature, man is forever seeking to transcend his material-bound self?*
    - c) *Does not the fact that Christ exists mean that his sickness is not unto death?*
  5. Larry Horn points out that parenthetical vocatives (unlike subjects) are perfectly acceptable here (thus [a] and [b]), as are indeed all parentheticals in general (thus [c]):
    - a) *Do not, anyone, step on my blue suede shoes.*
    - b) *Do not, O Claudius, step on my blue suede shoes.*
    - c) *Do not, by the way, step on my blue suede shoes.*
  6. Susan Schmerling suggests that a common discourse function of imperative *do* is to signal a reprise of an earlier invitation or command — presumably because the first invitation went unheeded or was declined. If true, this observation would make imperative *do* semantically much closer to emphatic *do* than Jespersen (or I) would admit.
  7. Both deriving from I.E. *dhee-* ‘put’. My thanks to Tania Kuteva for all information and references pertaining to Bulgarian.
  8. The practical orthography employed for Khmer examples here was developed by Stephanie Farmer and me in 2007, and defended in Ourn and Haiman 2007.
  9. My thanks to Mr. Noeurng Ourn, who has been trying to teach me Khmer since 1996, and to Mr. Veasna Keat, who has continued with this task since 2007.
  10. Whether or not there is an explicit conjunction and between conjuncts is important for familiar reasons, but all the Khmer examples discussed here are asyndetic.
  11. The expression *bo'ri'va: sap* (Pali *parivrāśabda* ‘retinue word’) derives from the Khmer grammatical tradition. All Khmer authorities concur on the important point that servant words are meaningless morphemes with a primarily euphonic function and are incapable of standing alone. “A servant word is a sound which accompanies a base word in a compound, much like a retinue or entourage [accompanies a ruler] in order to make speech more pleasant to the listener’s ear and also to qualify the original word” (Chun-Leuh 2007: 190). Philip Jenner (p.c.) points out to me that the term expresses

- ‘serial words’, i.e. words in tight sequence, and was originally used for the two-word combination, rather than just for the decorative word.
12. Some examples: *jo:k jaul* ‘give preferential treatment to (e.g., regular customers)’ is not found at all in Headley 1977. Chun-Leuh 2007 identifies the second element as a servant word, Veasna Keat the first (p.c.). *Jo:l jo:k* ‘oscillate’ is identified by Chun-Leuh as a compound whose second element is a servant word; Mr. Keat recognizes it as a synonym compound. *Raukhaek raukha:k* is translated in Headley as ‘in ruins’, and by Keat as ‘sloppy’. Chun-Leuh identifies its first element as a servant word; Mr. Keat classifies it as one indissoluble word. *Bawngheun bawnghaoc* ‘waste’ is not listed in Headley 1977. Chun-Leuh and Keat agree on identifying the first element as a servant word, while Noeurng Ourn (p.c.) identifies the expression as a synonym compound. (There are of course also many instances where all hands agree.) The Rashomon effect is precisely what one should expect when dealing with words in transition.
  13. The predilection for such structures is condemned by Khmer speakers themselves as “ungrammatical” — explicitly, because it violates grammatical standards introduced by the French. For example, the definition (cited in a free translation in note 11 above) of *bo’ri’va:r sap* in Chun-Leuh (2007: 190), a scholarly work, is translated literally below (symmetrical compounds are bracketed): “sound [accompany join] base word in [group compound] each noun [compare compare like] (sic!) be retinue in order to facilitate [speech speak] and play music [voice sound] pleasant [harmonize follow] sound base word so that exist state coherence facilitate to ear of listener on the one hand in order to describe word explain clearly [feature quality] of that base word [more also].”  
 Now here it is as tidied up by Mr. Veasna Keat, who condemns this passage as “too much like the spoken language”(Mr. Keat’s changes are highlighted): “sound which [accompany join] with base word in [group compound] each word [compare like] be retinue in order to facilitate [speech speak] so-that exist pleasant sound to follow of base word [in order so-that] exist state coherence facilitate to ear of listener and in order explain feature or property of that base word [more also].”  
 Some (but far from all!) symmetrical compounds have been eliminated, and a handful of “signpost” words supplied. Notably, Mr. Keat inserts a new symmetrical compound ([in-order so that]) which was not in the original that he “corrected”, mainly to get rid of them. A fascinating study could be written about Khmer attitudes and language ideology, as reflected in statements and behavior like this.
  14. The prefix *pra-* ‘forth’ in Sanskrit may well be the origin of Khmer *praw-*. It is characterized as having a largely pleonastic (decorative) function there by Goldman and Sutherland 1987: 139. My thanks to Erik Davis for the reference.
  15. This game also exists in Thai, where it is known as *kham phuan* ‘flipped words’ (Hoonchamlong 1991: Section 2.1 and references; Iwasaki and Ingkaphirom 2005: 46–47). Thus *yii pun* ‘Japan’ can be made into *yun pii*. I have encountered no reference to the existence of this game in any Western descriptions of Khmer. Unfortunately, I still know almost nothing about the social context of *piak kunloah kat*: who engages in it and when.
  16. Nick Enfield informs me that *nvaj na:j* exists in Thai as well — and that apparently *na:j* is just as much a bound form there as it is in Khmer. This is an opportune moment to acknowledge that I have systematically ignored all questions of borrowing; and that borrowing may be yet another source for both *bo’ri’va: sap* and for near synonyms.
  17. Lass 1990 insisted that only junk could be recycled. In his later book he acknowledged that “all is not junk” (1997: 318), but what is recycled is “either serving some other purpose or serving no purpose at all” (1997: 316).
  18. Lass (1997: 342–352) is skeptical of this universal. It is indeed a dogma of linguistic faith that no true synonyms exist: “the mind shuns purposeless variety”. It is also a

fact, however, that the word *synonym* itself exists in many languages, that Roget's thesaurus, and innumerable pairs like *last and final* in English and other languages are constructed on the basis of the concept, and, finally, that if Bréal's principle of repartition describes a historical process, it must have had a synonym situation to apply to in the first place. The difference between finding "No synonyms" and finding "Scads of synonyms" be one of perspective. Given enough attention and encouragement, speakers can suss out semantic, social or stylistic differences between any two synonyms. But for other (ludic or esthetic) purposes, it is precisely their functional identity which may be highlighted.

## References

- Antonova, Julija N., Elga G. Kirjakova & Tatjana K. Nakova. 1997. *Blgarski ezik, blgarija, i blgarite* [The Bulgarian language, Bulgaria, and the Bulgarians]. 4th edn. Sophia: Institut za čuždestranni studenti.
- Atanassova, Teodora, Jana Molhova, Vladimir Phillipov & Marco Mincoff. 1988. *Blgarsko-anglični rečnik* [Bulgarian-English dictionary]. Sofia: Izd. Nauka i izkustvo.
- van der Auwera, Johan. 2006. Why languages prefer prohibitives. *Journal of Foreign Languages* 161(1). 3–25.
- van der Auwera, Johan, Ludo Lejeune & Valentin Goussev. 2005. Prohibitives In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *World atlas of linguistic structures*. Oxford University Press.
- Bergs, Alexander. 2005. "Don't you kill the messenger!": The structure and development of the prohibitive. Unpublished manuscript. Heinrich-Heine University Düsseldorf.
- Bhat, D. N. Shankara. 2004. Negation. In Geert Booij, Christian Lehman, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphologie/Morphology: Ein internationales Handbuch zur Flexion und Wortbildung / An International Handbook on Inflection and Word-Formation* (Handbücher zur Sprach- und Kommunikationswissenschaft/ Handbooks of Linguistics and Communications Science HSK 17/2), 1207–1212. Berlin & New York: Mouton de Gruyter.
- Bolinger, Dwight. 1975. *Aspects of language*. 2nd edn. New York: Harcourt Brace Jovanovich.
- Bréal, Michel. 1897. *Essai de sémantique*. Paris: Hachette.
- Brinton, Laurel & Elizabeth Traugott. 2005. *Lexicalization and language change*. Cambridge: University Press.
- Bybee, Joan & Alexandra Brewer. 2005 [1980]. Explanation on morphophonemics. In Joan Bybee (ed.), *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar*. Chicago: University of Chicago Press.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chun-Leuh. 2007. *Ve:jjaka:w Kmae*. [Cambodian grammar.] Phnom Penh: n.p.
- Cook, Eung-Do. 2004. *A grammar of Dene Suline (Chipewyan)*. (Algonquian and Iroquoian Linguistics Special Athabaskan Number Memoir 17). Winnipeg: Algonkian & Iroquoian Linguistics.
- Croft, William. 1991. The evolution of negation. *Journal of Linguistics* 27. 1–27.
- Dahl, Osten. 1979. Typology of sentence negation. *Linguistics* 17. 79–106.
- Diffloth, Gerard. 1972. Notes on expressive meaning. *CLS* 8: 440–447.

- Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland*. Cambridge University Press.
- Ellegård, Alvar. 1953. *The auxiliary DO: The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell.
- Faltz, Leonard. 1998. *The Navajo verb*. Albuquerque: University of New Mexico Press.
- Goldman, Robert P. & Sally J. Sutherland. 1987. *Devanapravesika: An introduction to the Sanskrit Language*. 2nd edn. Berkeley, CA: University of California Press.
- Gould, Steven & Elizabeth Vrba. 1982. Exaptation — a missing term in the science of form. *Paleobiology* 8(1), 4–15.
- Haiman, John. 1998. Possible origins of infixation in Khmer. *Studies in Language* 22, 597–617.
- Haiman, John. 2003. Explaining infixation. In John Moore & Maria Polinsky (eds.), *The nature of explanation in linguistic theory*, 105–120. Stanford, CA: CSLI.
- Haiman, John. 2008. In defence of iconicity. *Cognitive Linguistics* 19(1), 35–48.
- Haiman, John & Stephanie Farmer. 2007. A user-friendly Khmer-English dictionary. Unpublished manuscript, Macalester College.
- Haiman, John & Noeurng Ourn. 2003. Nouns, verbs, and syntactic backsliding in Khmer. *Studies in Language* 27(3), 505–528.
- Haiman, John & Noeurng Ourn. 2009. Decorative symmetry in ritual (and everyday) language. In Roberta Corrigan, Edith Moravcsik, Hamid Ouali & Kathleen M. Wheatley (eds.), *Formulaic language, volume 2: Acquisition, loss, psychological reality, and functional explanations*, 567–588. Amsterdam & Philadelphia: John Benjamins.
- Han, C.-Y. 1998. *The structure and interpretation of imperatives: Mood and force in universal grammar*. Philadelphia: University of Pennsylvania dissertation.
- Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37(6), 43–68.
- Haspelmath, Martin, Matthew Dryer, David Gil & Bernard Comrie (eds.). 2005. *World atlas of linguistic structures*. Oxford: Oxford University Press.
- Headley, Robert, Kylin Chhor, Lam Kheng Lim, Lim Hak Kheang & Chen Chun. 1977. *Cambodian-English dictionary*. Washington DC: Catholic University of America Press.
- Heath, Jeffrey. 1982. *Nunggubuyu dictionary*. Canberra: Australian Institute of Aboriginal Studies.
- Heine, Bernd & Tania Kuteva. 2002. *World Atlas of Grammaticalization*. Cambridge: Cambridge University Press.
- Hoonchamlong, Y. 1991. Tones in Thai: An autosegmental account. <http://yhoonchamlong.net/yuipublist.htm>
- Horn, Laurence. 1989 *A natural history of negation*. Chicago: University of Chicago Press.
- Ieu Koes. 1995[1945]. *Phiasa: kmae*. [The Cambodian language]. Phnom Penh: Buddhist Institute.
- Iwasaki, Shoichi & Preeya Ingkaphirom. 2005. *A reference grammar of Thai*. Cambridge: Cambridge University Press.
- Jacob, François. 1977. Evolution and tinkering. *Science* 196, 1161–1166.
- Jespersen, Otto. 1940. *A modern English grammar on historical principles, volume 5*. Copenhagen: Ejnar Munksgaard.
- Kahrel, Peter J. & René van den Berg (eds.). 1994. *Typological studies in negation*. Amsterdam & Philadelphia: John Benjamins.
- Klima, Edward. 1964. Negation in English. In Jerold Katz & Janet Fodor (eds.), *The structure of language*, 246–323. Englewood Cliffs, NJ: Prentice-Hall.
- Kroch, Anthony. 1989a. Function and meaning in the history of English: Periphrastic Do. In Ralph Fasold & Deborah Schiffrin (eds.), *Language variation and change* (Current Issues in Linguistics Theory 52), 133–172. Amsterdam & Philadelphia: John Benjamins.

- Kroch, Anthony. 1989b. Reflexes of grammar in patterns of language. *Language Variation and Change* 1(3). 199–244.
- Lass, Roger. 1990. How to do things with junk. *Journal of Linguistics* 26. 79–102.
- Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lehmann, Christian. 2002. New reflections on grammaticalization and lexicalization. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 1–18. Amsterdam & Philadelphia: John Benjamins. Amsterdam: Benjamins.
- Mayr, Ernst. 2001. *What evolution is*. New York: Basic Books.
- Ourn, Noeurng & John Haiman. 2000. Symmetrical compounds in Khmer. *Studies in Language* 24. 483–514.
- Ourn, Noeurng & John Haiman. 2007. Creative forces in Khmer. In Shoichi Iwasaki, Andrew Simpson, Karen Adams & Paul Sidwell (eds.), *Papers from the 13th meeting of the Southeast Asian Linguistics Society 2003*, 157–164. Canberra: Pacific Linguistics.
- Paul, Hermann. 1995[1880]. *Prinzipien der Sprachgeschichte*. Tübingen: Max Niemeyer.
- Payne, John. 1985. Negation. In Timothy Shopen (ed.), *Typology and syntactic description, volume 1: Clause structure*, 197–142. Cambridge: Cambridge University Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Rice, Keren. 1989. *A grammar of Slave*. Berlin & New York: Mouton de Gruyter.
- Rissanen, Matti. 1991. Spoken language and the history of DO-periphrasis. In Dieter Kastowsky (ed.), *Historical English syntax*, 321–342. Berlin & New York: Mouton de Gruyter.
- Rupp, Laura. 1999. *Aspects of the syntax of English imperatives*. Colchester: University of Essex dissertation.
- Sadock, Jerry & Arnold Zwicky. 1985. Speech act distinctions in syntax. In Timothy Shopen (ed.), *Typology and syntactic description, volume 1: Clause structure*, 155–196. Cambridge: Cambridge University Press.
- Sandonato, Marie. 1994. Zazaki. In Peter J. Kahrel & René van den Berg (eds.), *Typological studies in negation*, 125–142. Amsterdam & Philadelphia: John Benjamins.
- Schmerling, Susan. 1982. How imperatives are special and how they aren't. In Robinson Schneider, Kevin Tuite & Robert Chametzky (eds.) *Papers from the parasession on nondeclaratives*, 202–218. Chicago: Chicago Linguistics Society.
- Sisowath, M. (n.d.). *Bo'ri'va: sap knong phiasa: kmaer*. [Servant words in Cambodian]. n.p.
- Stein, Dieter. 1990. *The semantics of syntactic change*. Berlin & New York: Mouton de Gruyter.
- Tottie, Gunnel. 1991. *Negation in English speech and writing*. New York: Academic Press.
- Visser, F. Th. 1978. *An historical syntax of the English language, volume III*. Leiden: Brill.
- Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.
- Warner, Anthony. 1993. *English auxiliaries*. Cambridge: Cambridge University Press.
- Wierzbicka, Anna. 1972. *Semantic primitives*. Frankfurt a.M.: Athenäum.
- Wolfart, H. Christoph. & Freda Ahenakew. 1998. *The student's dictionary of literary Plains Cree*. (Algonquian and Iroquoian Linguistics Memoir 15). Winnipeg: Algonquian and Iroquoian Linguistics.
- Zanuttini, Raffaella. 1997. *Negation and clausal structure: A comparative study of Romance languages*. Oxford: Oxford University Press.

Copyright of Linguistics is the property of De Gruyter and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.