Productivity of compounds: an application of construction morphology

Yoshihiko Asao

Kyoto University

September 26, 2008
ICCG-5 @ Austin, TX

Introduction

- Productivity of compounds cannot be adequately described with existing measures
- The idea of construction morphology is useful to describe the productivity of compounds
- These ideas will be illustrated by investigating Japanese verbal compounds

Productivity of compounds

Question:

How to describe productivity of compounds?

Productivity of compounds

- Productivity of compounds has often been described in informal ways
- Are quantitative measures available?
- As for derivational affixes, quantitative measures for productivity have been proposed

Productivity of compounds

Baayen's productivity measure for derivational affixes (Baayen & Lieber 1991; Baayen 1992)

$$\mathcal{P} = \frac{n_1}{N}$$

where

- n_1: the number of hapax legomena (words that occur only once)
- N: the token frequency

Productivity of compounds

$\mathcal{P}$ can be naively applied for compounds

	$\mathcal{P}$
nominal compounds	0.225
-tje	0.253
-ing	0.038
-heid	0.114
-er	0.076

(Baayen 1993: 183)

But what does this figure exactly mean?

Productivity of compounds

The high value of $\mathcal{P}$ for nominal compounds can mean:

- any pairs of nouns are able to form compounds (when semantic or other conditions are met)

But it is also possible that

- Only a few nouns can occur as second constituents of compounds, and they are very productive

This point is illustrated by Japanese verbal compounds

Japanese verbal compounds

compound verbs

- hasiri-tudukeru (run-continue) ’keep running’
- osi-taosu (push-fell) ‘push down’

deverbal compounds

- sara-arai (dish-wash) ‘dishwashing’
- te-dukuri (hand-make) ‘handmade’
- mahô-tukai (magic-use) ‘magician’

Japanese verbal compounds

By adopting Pinker and Prince's (1991) Dual Mechanism Theory, Ito & Sugioka (2002) claim that both compound verbs and deverbal compounds can be divided into two subtypes: rule-based and analogy-based.

Japanese verbal compounds

productive (rule-based) compound verbs

- -hazimeru ‘begin’ 0.069
- -tudukeru ‘continue’ 0.052
- -eru ‘be able’ 0.040

An example:

tabe-hazimeru
eat-begin
‘begin to eat’

Another example:

guguri-hazimeru
google-begin
‘begin to google’

Japanese verbal compounds

argument compounds: rule-based

- gomi-sute (garbage-throw.away) ‘garbage disposal’
- mado-sime (window-close) ‘window closing’

adjunct compounds: analogy-based

- te-gaki (hand-write) ‘handwritten’
- hi-yake (sun-burn) ‘sunburn’

Japanese verbal compounds

But some adjunct compounds are as productive as productive compound verbs, which are considered to be rule-based:

- -umare ‘be born’ 0.073
- -gurasi ‘live’ 0.058

An example:

kare-wa Tôkyô-umare da.
he-TOP Tokyo-be.born be

‘He was born in Tokyo.’

Another example:

kare-wa Tekisasu-umare da.
he-TOP Texas-be.born be

‘He was born in Texas.’

Japanese verbal compounds

Are these productive adjunct compounds rule-based?

No, according to the previous studies:

> ..productivity of these adjunct compounds are not due to the rule operating on abstract categories but rather they are formed by analogy based on a specific head, ..
>
> (Sugioka 1996: 237)

How about compound verbs like -hazimeru (begin), which are considered to be rule-based but nonetheless have specific heads?

Japanese verbal compounds

To summarize,

- productive compound verbs like -hazimeru ‘begin’ are considered to be rule-based because their first constituents are open-ended, while
- adjunct deverbal compounds like -umare ‘be born’ are considered to be analogy-based because their second constituents are lexically specified

In fact, both classes behave in the same way

Japanese verbal compounds

This confusion arises because the rule/lexicon dichotomy does not work — they are partially lexical and partially schematic

Construction Morphology

- Linguistic knowledge can be described by a network of constructions
- A construction is a pair of form and meaning
  e.g. [[S] [V] [O_1] [O_2] / S CAUSES O_1 TO RECEIVE O_2 BY VING]
  (cf. Goldberg 1995)
- The idea of construction grammar is also useful in morphology (Booij 2005, 2007)
  e.g. [[V]-er / ONE WHO Vs]

Construction Morphology

Japanese verbal compounds are instances of partially lexically specified constructions, or constructional idioms

- [[V]-hazimeru / BEGIN TO V]
- [[N]-umare / BORN IN N]

Argument deverbal compounds, such as mado-sime (window-close, 'window closing'), are instances of the fully schematic construction

- [[N]-[V] / VING N]

Construction Morphology

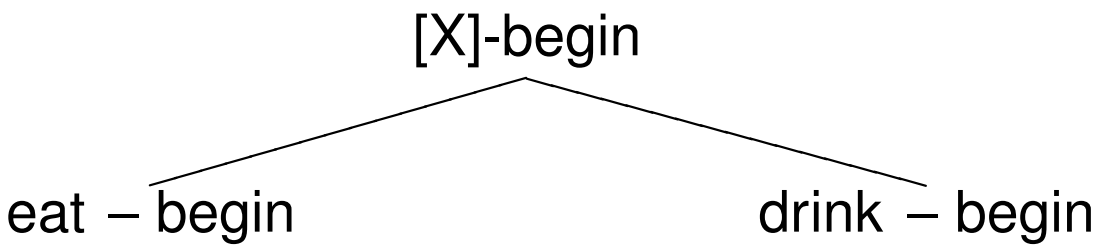
- The productivity of compounds can be replaced by the strength or degree of entrenchment of each construction
- How to calculate the degree of entrenchment?

Simulation

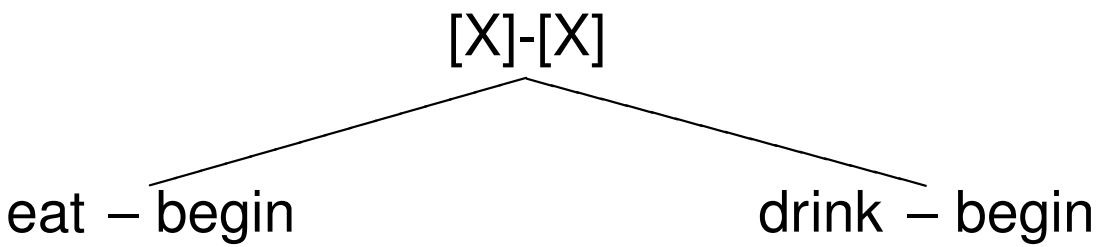
A simple simulation model

- Compounds are given to the model one by one, based on actual frequency data
- A construction is established when commonality between compounds is found by the model
- A construction gains strength when it sanctions a compound
- Constructions must be maximally specific; the most specific construction is chosen whenever more than one constructions are available

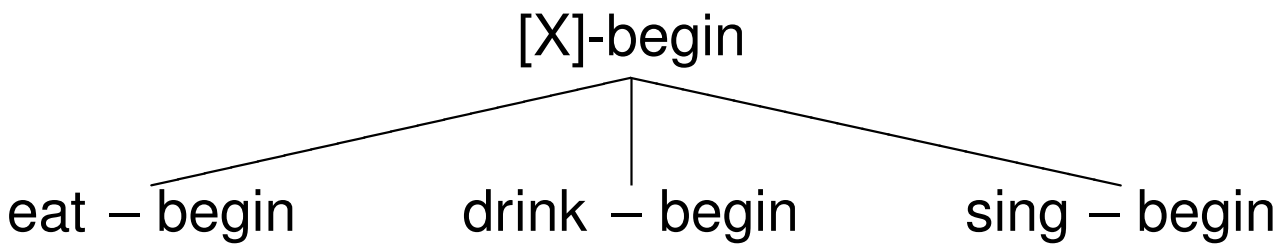
Simulation



Simulation



Simulation



Simulation

Procedure

- 10,000 compound verbs are randomly extracted from a newspaper corpus
- Compounds are given to the model
- The degree of entrenchment is calculated for each established construction
- The same procedure was then applied to deverbal compounds

Simulation

Results

- [[N]-[V]] has much higher degree of entrenchment (348) than [[V]-[V]] do (120)
- The constructions [[V]-hazimeru (begin)] (37) and [[N]-umare (be.born)] (22) can be found among the most entrenched partially lexical constructions
- However, some results does not accord well with native speakers' intuition and Baayen's $\mathcal{P}$; for example, [[V]-komu (get.in)] has higher degree of entrenchment than [[V]-hazimeru (begin)], although it is known to be unproductive. This is possibly because the data is too small

Conclusions

- Productivity of compounds cannot be fully described by a one-dimensional scale
- Construction Morphology is useful to describe productivity of compounds
- Simulation studies can model how constructions are acquired and used

My slides will be available at
http://ling.bun.kyoto-u.ac.jp/~asaokitan/

Search my name on the web

Thank you!

References I

Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In Booij, G., & van Marle, J. (Eds.), Yearbook of Morphology 1991, pp. 109–149. Kluwer, Dordrecht.

Baayen, R. H. (1993). Frequency, transparency and productivity. In Booij, G., & van Marle, J. (Eds.), Yearbook of Morphology 1992, pp. 181–208. Kluwer, Dordrecht.

Baayen, R. H., & Lieber, R. (1991). Productivity and English derivation: A corpus based study. Linguistics, 29, 801–843.

Booij, G. (2005). Compounding and derivation: evidence for Construction Morphology. In Dressler, W. U., Rainer, F., Kastovsky, D., & Pfeiffer, O. (Eds.), Morphology and Its Demarcations. John Benjamins, Amsterdam and Philadelphia.

References II

Booij, G. (2007). Construction Morphology and the Lexicon. In Montermini, F., Boyé, G., & Harbout, N. (Eds.), Selected Proceedings of the 5th Décembrettes, pp. 34–44 Somerville, MA. Morphology in Toulouse.

Goldberg, A. E. (1995). Constructions: a Construction Grammar Approach to Argument Structure. University of Chicago Press, Chicago.

Ito, T., & Sugioka, Y. (2002). Go no Sikumi to Gokeisei [Structure of Words and Word Formation], Vol. 16 of Eigogaku Monograph Series. Kenkyûsya, Tokyo.

Kageyama, T. (1989). The place of morphology in the grammar. In Booij, G., & van Marle, J. (Eds.), Yearbook of Morphology, Vol. 2, pp. 73–94. Foris Publications, Dordrecht.

References III

Kageyama, T. (1999). Word Formation. In Tsujimura, N. (Ed.), The Handbook of Japanese Linguistics, pp. 297–325. Blackwell, Oxford.

Pinker, S., & Prince, A. (1991). Regular and Irregular Morphology and the Psychological Status of Rules of Grammar. In BLS, Vol. 17, pp. 230–251.

Sugioka, Y. (1996). Regularity in inflection vs. derivation: rule vs. analogy in deverbal compound formation. Acta Linguistica, 45, 231–253.