Productivity and the Lexicon

Andrea D. Sims
The Ohio State University
A definition

Productivity of a morphological pattern = the likelihood of it being applied to new bases to create new lexemes (= derivational) or new word-forms (= inflectional)
Measuring productivity

- The most productive morphological patterns occur disproportionately often among the lowest frequency words

- “Hapax-based productivity measures”
  - Hapax (legomenon) = A word that occurs exactly once in some corpus
Distribution of word frequencies

- In any corpus, there are many hapaxese!
  - Approximately 40-50% of all words

Based on the book *Alice in Wonderland*
source: Baayen (2001)

![Figure 1.3: The frequency spectrum of Alice in Wonderland (m: frequency class; V(m, N): number of types with frequency m).]
Measuring productivity

- Semantically similar morphological patterns can differ in productivity

Growth rate of vocab

Based on the novel Moby Dick

source: Baayen (2001)
Measuring productivity

- And differ in productivity over time

Figure 6.1: Relative profitability of five nominalisation processes

New entries in the Oxford English Dictionary

source: Bauer (2001)
Factors

- Many factors affect productivity
  - Selectional restrictions
  - Semantics/pragmatics of the resultant word
  - Social and stylistic factors
    - Text type, perception of nativeness/foreignness, prescriptivism, etc.
  - Structure of entries in the lexicon and word processing
    - Blocking effects, and much more
<table>
<thead>
<tr>
<th>-er derived word</th>
<th>Blocking word</th>
</tr>
</thead>
<tbody>
<tr>
<td>cycler</td>
<td>cyclist</td>
</tr>
<tr>
<td>batter (in cricket)</td>
<td>batsman</td>
</tr>
<tr>
<td>typer</td>
<td>typist</td>
</tr>
<tr>
<td>studier</td>
<td>student</td>
</tr>
<tr>
<td>stealer</td>
<td>thief</td>
</tr>
<tr>
<td>deliverer</td>
<td>delivery person</td>
</tr>
<tr>
<td>shop assister</td>
<td>shop assistant</td>
</tr>
<tr>
<td>lift attendner</td>
<td>lift attendent</td>
</tr>
</tbody>
</table>
Productivity and the lexicon

- Hypothesis: Productivity is a function of the resting activation level of the morphological pattern in the mental lexicon
  - Resting activation level = extent to which a lexical entry is activated in the mind when not receiving stimulation

- Productivity is thus a reflection of the storage and processing of complex words
Item-and-Arrangement yet again

- **Primitive elements** = morphemes
  - Morphemes = lexical bundles of form + meaning
  - Lexicon contains morphemes

- **Operation type** = concatenation

- **Conditions** = mostly affix-driven selectional restrictions (i.e. affixes select bases with certain properties)

- **Output** = meaning-adding ("incremental")
Distributed Morphology

- **Primitive elements** = morphemes
  - Morphemes = abstract sets of morphosyntactic values
  - Lexicon contains morphophonological forms that realize morphemes
    - Minimal lexicon – roots + affixes separately

- **Operation type** = concatenation

- **Conditions** = mostly affix-driven selectional restrictions (i.e. affixes select bases with certain properties)

- **Output** = meaning-realizing (“realizational”)

Word and Paradigm redux

- **Primitive elements** = words
  - Lexicon = whole words, and maybe also entries for generalizations made over whole words (realizational rules)

- **Operation type** = processes
  - Functions over stems that may include concatenation, but are not limited to this

- **Conditions** = affix-driven selectional restrictions, but less limited by this

- **Output** = meaning-realizing (“realizational”)
The Big Question

- What does productivity indicate about the structure of the lexicon?
  - And vice versa?

- And by extension, about what kind of morphological theory is best?
Pinker’s (1991) Dual-Route Model

- Only simplex words and irregular derived words are stored in the lexicon
  - Connected by associative network

- Regular derived words are stored/accessed according to component morphemes

- Postulation: The lexicon is optimized for storage efficiency (i.e. minimal amount of memory space)
  - Notice the implicit evaluation metric!
Evidence

- Regularization through derivation: "verbs intuitively perceived as derived from nouns of adjectives are always regular"
  - E.g. grandstanced, flyed out, high-sticked

- Lexical compounding can have internal inflection only if it is irregular
  - mice-infested vs. ??rats-infested
  - teethmarks vs. ??clawsmarks
  - men-bashing vs. ??guys-bashing
The Problem

Evidence that *some* regular forms are composed by rule rather than being directly stored/accessed in the lexicon does not mean that *all* regular forms are composed by rule.
A different hypothesis

- The lexicon is fundamentally **word-based**
  - Morphologically regular words may be stored 'whole'
  - Some words may still be faster to process via 'parts'

- Morphological rules are emergent from word-based lexical entries
  - Via associative network of connections among lexical entries
  - Morphological rules as 'redundancy rules'

- No special status for irregulars vs. regulars
  - Or for concatenation (morphemes) vs. non-concatenative processes
Are neutral and non-neutral derivational affixes in English structured differently in the lexicon?

Neutral = does not trigger allomorphy in base
- E.g. -en, -ize, -ness, -able, -ment, -er

Non-neutral = does (sometimes) trigger allomorphy in base
- E.g. -ion, -al\textsubscript{N} –al\textsubscript{V} –ity, -ous, -ic
Alegre and Gordon (1999)

- Study 1: Analysis of phonological similarity of words with given affix
  - Lexical gangs = “... sets of words with shared phonological and semantic properties that influence morphological productivity” (Pierrehumbert 2012)

- Results: “All nonneutral affixes display a strong [lexical] gang organization. The same is true for two neutral affixes: -en and -ize... being nonneutral is a sufficient but not a necessary condition to attract gang clustering” (349).
**Alegre and Gordon (1999)**

### TABLE 1
Gang Structure in Derivational Forms

<table>
<thead>
<tr>
<th>Affix</th>
<th>Type</th>
<th>No. of types analyzed</th>
<th>No. of gangs identified</th>
<th>% of types in gangs</th>
<th>% of types with partial overlap</th>
<th>% of types with no associative support</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ion</td>
<td>Nonneutral</td>
<td>477</td>
<td>9</td>
<td>87</td>
<td>10.5</td>
<td>2.5</td>
</tr>
<tr>
<td>-alN</td>
<td>Nonneutral</td>
<td>389</td>
<td>15</td>
<td>89</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>-alN</td>
<td>Nonneutral</td>
<td>36</td>
<td>5</td>
<td>89</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>-ity</td>
<td>Nonneutral</td>
<td>195</td>
<td>9</td>
<td>90</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>-ous</td>
<td>Nonneutral</td>
<td>85</td>
<td>3</td>
<td>77</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>-ic</td>
<td>Nonneutral</td>
<td>167</td>
<td>4</td>
<td>86</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>-en</td>
<td>Neutral</td>
<td>30</td>
<td>1</td>
<td>100</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>-ize</td>
<td>Neutral</td>
<td>76</td>
<td>5</td>
<td>70</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>-ness</td>
<td>Neutral</td>
<td>318</td>
<td>5</td>
<td>17</td>
<td>—</td>
<td>83</td>
</tr>
<tr>
<td>-able</td>
<td>Neutral</td>
<td>302</td>
<td>6</td>
<td>20</td>
<td>—</td>
<td>80</td>
</tr>
<tr>
<td>-ment</td>
<td>Neutral</td>
<td>246</td>
<td>5</td>
<td>20</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>-er</td>
<td>Neutral</td>
<td>516</td>
<td>3</td>
<td>11</td>
<td>89</td>
<td>89</td>
</tr>
</tbody>
</table>
Alegre and Gordon (1999)

- **Study 2: Rating of novel forms (productivity!)**
  - Stimulus design: 2 (gang affix) x 3 (similarity to attested forms)
    - Gang affix: yes/no
    - Similarity: near/intermediate/distant

- **Results: Gang x similarity interaction**
  - The similarity effect for derived forms was significant for Gang affixes (-ion, -al_{N}, -al_{V}, -en) but not for the No-Gang affixes (-er, -ness, -able)
Ratings of nonwords (wugs) based on magnitude estimate
- Anchor = ‘tralden’ = 100

Group 1: dighten, peaten, thitten, totten, vaughten
Group 2A: balten, gleeten, nilten, ploaten, pratten
Group 2B: boppen, dauppen, pipen, neapen, vappen
Group 3: cliven, diffen, dussen, naffen, plarcen
Group 4: blizen, flotchen, meechen, sorzen, zinthen
Group 5: arpen, elzen, orthen, flimperen, hickelen, breenen, roren, nirmen, beelananen, prilen
Acceptability of -en words according to template

<table>
<thead>
<tr>
<th>Template</th>
<th>Average ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td>2A</td>
<td>110</td>
</tr>
<tr>
<td>2B</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>70</td>
</tr>
</tbody>
</table>
FIG. 4. Acceptability ratings for derivational forms for individual affixes.
Alegre and Gordon (1999)

- Gang clustering among non-neutral (and some neutral) affixes indicates a word-based pattern of storage
  - Logic: Phonological similarity effects cannot exist if the affix is abstracted away from the word-forms

- On the other hand, lack of gang clustering among remaining neutral affixes might be indication that not all words are stored in the lexicon
  - Or more precisely, that not all words are accessed during lexical processing via whole-word entries
Discussion of analytic exercise 5

- In the final analytic exercise, you looked at the productivity of English past tense formation (i.e. inflection)

- Is the productivity of irregular past tense patterns gradient or categorical?

- What about regular past tense patterns?

- What does this suggest about word-based vs. morpheme-based storage in the lexicon?
Follow up question 1

- Are all regular (and irregular) words stored in the lexicon and accessed as whole words?

- In other words, what is the balance between whole-word storage/access vs. morpheme-based storage/access?
Plag and Baayen (2009)

- Are there whole-word frequency effects for words with regular derivational suffixes?

- Investigated the processing of 2,529 derived English words containing only root + suffix
  - -th, -en, -ment, -or, -ster, -ary(N), -ian, -er, -ette-, -ary(Adj), -ive, -ist, -ee, -ish, -ess, -age, -ly(Adj), -ery, -ling, -ship, -dom, -hood, -less, -ous, -ful(Adj), -fold, -wise, -ly(Adv), -ful(N), -ism, -ness

- Measures: Word naming latencies and lexical decision latencies
Plag and Baayen (2009)

- Strong effect of derived word frequency in both lexical decision (left) and word naming (right) tasks

![Graph showing the relationship between derived frequency and log lexical decision latency on the left.](image1)

![Graph showing the relationship between derived frequency and log naming latency on the right.](image2)
Plag and Baayen (2009)

- Predicted bias in favor of whole-word storage of derived words in English
- Some affixes occur mostly in words predicted to be stored
- Storage-dominant = fastest to process

![Graph showing median naming latency](image)

- Latinate word types predicted to be parsed
- Germanic word types predicted to be parsed
Lexical processing involves a balance between direct access (i.e. via whole words) and computation (via “morphemes”)
- Item-by-item, but with aggregate effects for English suffixes
- Inherent bias (in English) towards storage of and access via whole word representations
- Even for regular derived words!
- Postulation: The lexicon (and lexical access) are optimized for efficiency of access (speed!), rather than efficiency of storage
Follow up question 2

How does this related to productivity?
Hay and Baayen (2002)

- Number of hapaxes (V1) (a measure of productivity) vs. number of tokens/types estimated to be parsed during lexical access. Each dot is an English suffix.
Interpretation

- More access to lexical entry for affix pattern (parsing) → more productivity of affix
The Big Points

 Psycholinguistic evidence: lexicon is a network in which whole words are frequently stored and connected to each other associatively
  - Based on phonological, morphological, and/or semantic similarity…

 Generalizations about word-form relatedness can be abstracted from these lexical entries + associative connections (e.g., un-\(X_{adj}\)). These are morphological patterns
The Big Points

- Some morphological patterns are more likely to be activated during lexical access than others
  - But not straightforwardly related to regular vs. irregular

- Amount of activation determines the productivity of the pattern

- Productivity is thus a product of the structure of the lexicon (and word processing), and conversely, is informative about the lexicon
References


