TRUBETZKOY’S ORPHAN

PROCEEDINGS OF THE MONTRÉAL ROUNDTABLE
“MORPHONOLOGY: CONTEMPORARY RESPONSES”
(Montréal, September 30 - October 2, 1994)

Edited by
RAJENDRA SINGH

Université de Montréal

With the collaboration of
RICHARD DESROCHERS

Université de Montréal

Volume 144

Rajendra Singh (ed.)

Trubetzkoy’s Orphan
Proceedings of the Montréal Roundtable
“MorphonoLOGY: Contemporary responses”
We simply need to know more about the lexicon before we can make further progress in other fields.
D.L. Bolinger (1973:9)
linguist could discern as structure, but in the perspective taken here, the question of what is a viable synchronic pattern is taken as a serious empirical question. Decades of research in generative grammar, with its emphasis on competence, have failed to turn up answers to the basic questions we will address here: how do language users acquire, internalize and generalize over morphological and morphophonological patterns?

The core issue in synchronic morphophonology is the nature of productivity and what factors determine the relative degrees of productivity present at a given stage of a language. In this paper it will be argued that 'productivity', defined as the likelihood that a pattern will apply to a new form, is a direct reflection of the type frequency of that pattern.

In connection with productivity, we will examine the related notion, 'regularity', defined as the relative lack of lexical idiosyncrasy, and a third notion, sometimes thought to be related to the others, 'fusion', defined as the extent to which the phonological shape of two morphemes are co-mingled or co-determined.

In natural language — and in most theories of natural language — productivity, regularity and lack of fusion tend to be characteristic of the same patterns, and their opposites, lack of productivity, irregularity and high fusion tend to characterize another set of patterns. I will argue that the links among these three properties of morphological and morphophonemic patterns are diachronic in nature and that synchronically they are relatively independent of one another. In the associative Network model argued for here, synchronic productivity and regularity result from the way language is used, and the difference between productive, regular patterns and unproductive, irregular ones is a quantitative rather than a qualitative difference. This model will be compared with Level-Ordered morphology and the Dual-Processing model proposed by Pinker, Marcus and others, both of which hypothesize innate, structural differences among types of morphological and morphophonological rules.

1. **The Network Model**

   In the model I have proposed in various works (Bybee 1985, 1988, 1991), irregular, unproductive morphology and regular, productive morphology are not categorically distinguished, but rather represent two ends of a continuum. The representation of these types of morphology is similar — whatever generalizations exist emerge from lexical patterns — and the productivity of any pattern is predictable from its distribution in the language. In many ways this model resembles a connectionist model since it proposes that generalizations arise from the patterns in an associative network; however this model is more complex than existing connectionist models in that it incorporates the notions of lexical item with a wide variety of inherent properties, lexical strength due to token frequency, and morphological analysis resulting from sets of interconnections among items. The three relevant properties of the model for present purposes are explained below. Note that the first and third of these properties relate language use to properties of the lexical representation.

   (i) Words entered in the lexicon have varying degrees of lexical strength, due primarily to their token frequency. Words with high lexical strength are easy to access, serve as the bases of morphological relations and exhibit an autonomy that makes them resistant to morphophonological change and prone to semantic independence.

   (ii) Words entered in the lexicon are related to other words via sets of lexical connections between identical and similar phonological and semantic features. Parallel phonological and semantic connections constitute morphological relations. These connections among items have the effect of yielding a morphological analysis, as shown in Figure 1.

   (iii) Sets of words having similar patterns of semantic and phonological connections reinforce one another and create emergent generalizations describable as schemas. New items or items whose connections are not known or are weak can be fit into these schemas. The likelihood of the schema being extended to new items is directly dependent upon the defining properties of the schema and its strength, the latter property

---

1 'Distribution' includes more than simple type frequency; it also must take account of

---

lexical features such as syntactic category, phonological shape and other subdivisions of the lexicon, such as borrowed vs. native items.
being derivable from the number of items that reinforce the schema. Thus productivity is a direct consequence of type frequency.

A consequence of the notion of lexical strength is that morphologically complex words can be represented lexically if they have sufficient token frequency. Thus some regular and productive formations are represented lexically, in association with the forms to which they are related. This means that there are two ways of producing a regular form: by direct lexical access of that form as a lexical unit, and by access of a base form and the schema for the morphological pattern (Bybee 1985, 1994).

Given that morphological patterns are represented as lexical connections, there is no particular advantage to having, for instance, affixes rather than stem changes, since they are both represented in the lexicon (rather than in a rules component) by lexical connections (Compare figures 1 and 2). On the other hand, high type frequency will give whatever patterns do exist a strong representation, making them highly available for new formations. Rubba (1993) presents an associative network analysis of Modern Aramaic which demonstrates that many of the problems encountered in analyzing Semitic languages with their interwoven lexical and grammatical morphemes are easily accommodated in an associative network model.

Figure 2: Network representation of a schema for a vowel change class

---

2 Productivity, Regularity and Fusion

2.1 Productivity

Morphological productivity is the extent to which a morphological pattern applies to new forms. Relative productivity may be measured among different morphemes or different allomorphic patterns. At the morphemic level, word-formation devices (or derivational patterns) can differ in their productivity. A well-known example of a productive derivational suffix in English is -ness, which is far more productive than the derivational suffix -dom, for instance.

Productivity of allomorphic patterns is important in inflection. Inflections are by definition obligatory, so for every inflectional morpheme, there must be at least one productive pattern, one way of inflecting new words. Different patterns for the same morpheme can be more or less productive; English plural -s is the only productive pluralization allomorph, though other expressions of plural in nouns exist.

Generative theories provide no way of predicting the productivity of a given morphological or allomorphic pattern. The two generative theories to be discussed below (Level-Ordering and Dual-Processing) make distinctions among rule types that correlate partially with productivity, but these distinctions are structural in nature and do not provide an explanation for why one pattern is productive and another is not. On the other hand, MacWhinney (1978), Bybee (1985, 1988) and Baayen & Lieber (1991) claim that productivity is directly related to the type frequency of the pattern. Bybee (1988, 1995) notes further that the openness of the defining features determining the domain of the pattern also contributes to productivity; that is, a pattern with no phonological or semantic restrictions is more likely to be or become productive than one with such restrictions.

Thus in the Network model, productivity is directly related to properties of the use of a pattern, primarily to the number of different types to which the pattern applies. The type frequency is represented in the strength of the schema: schemas that apply to large numbers of items are highly reinforced and thus highly available for future uses.

2.2 Regularity

Often the terms ‘regular’ and ‘productive’ are used interchangeably, and while productive patterns are usually regular, it is worth pointing out a difference in focus in the two terms, especially given the use of the regular/irregular distinction in some recent theories (such as the Dual-Processing model, see below). ‘Regular’ is contrasted with ‘irregular’, where the latter indicates that a pattern is characterized by lexical idiosyncrasies. Lexical idiosyncrasies may be measured in degrees on two dimensions: the extent to which the pattern is applicable only to arbitrary lexical classes and the extent to which it deviates from the regular pattern. If English Past tense -ed is regular (despite phonologically conditioned allomorphy), the verbs which use this suffix and a vowel change (kept, slept, left, etc.) are somewhat irregular,
verbs that use only a vowel change (bit, drove, struck) are more irregular, and verbs that have both vowel and consonant changes are the most irregular (thought, taught, went). Thus irregularity is defined language-internally in relation to the most regular pattern.

2.3 Fusion
Some morphological patterns involve a greater degree of fusion between the stem and the grammatical morpheme than others, which are more agglutinative in nature (Bybee 1985). The other end of this scale is called ‘morphotactic transparency’ by Dressler (1985a). When I use the term ‘transparency’ in this paper, I will be referring to morphotactic transparency. Some indicators of fusion (lack of transparency) are allomorphic changes conditioned in the stem by the affix, or the reverse, allomorphic changes conditioned in the affix by the stem, as well as actual phonological fusion at the boundary between the two. Even greater fusion would be indicated by the use of a stem-change to express the grammatical category, or the interweaving of lexical and grammatical morphemes. The most extreme case of fusion would be the total replacement of the stem, as in suppletion.

It should be noted that the distribution of high fusion in a language is not arbitrary. Two gradient factors affect fusion. One is the semantic relevance of a grammatical category to the stem — higher relevance categories are more fused with the stem (Bybee 1985). The other is degree of grammaticization; as grammaticization proceeds over time, affixes become more fused with stems, as we shall see below.

3. The Relations among Productivity, Regularity and Fusion
The hypothesis presented here is that the common co-occurrence of productivity, regularity and transparency in the same morphological patterns is a result of diachronic patterns of development and is not a psychologically valid correspondence that should be modeled for synchronic morphology. However, because these three properties tend to be associated in natural language such that morphological or allomorphic processes that are productive tend also to be lexically regular and exhibit a relatively low degree of fusion, while unproductive ones tend to be irregular and characterized by higher fusion, some theories build parts of this association into their models. I will argue that a model, such as the Network model, that does not insist on any of these correlations is more accurate than ones that do.

The most complete association of the three factors is hypothesized by Dressler in the theory of Natural Phonology and Morphology, which correlates morphotactic transparency (lack of fusion) with productivity, and with regularity (lack of arbitrary allomorphy) (1985a:316-22). According to Dressler, transparent morpheme combinations are easier to process (319), which, I surmise, in his view leads to their greater productivity. Lack of allomorphy and lack of fusion both result from the same general principle of semiotic transparency, or the one-to-one relation between meaning and form. Operations that are more natural on one scale will also be more natural on other scales. Thus the three properties discussed here all correlate in Natural Morphology.

In a model with level-ordering of phonological and morphological rules, both phonological and morphological rules are sequenced in such a way that the ones with the most lexical restrictions are also the ones that have greater access to the stem, and thus create the greatest fusion (Kiparsky 1982a). In this model, then, irregularity and fusion are correlated. It is also true of this model that the rules in later levels are more productive, a property derivable from the generality of rules. By the Elsewhere Condition more specific rules apply first and more general rules apply later to all forms not affected by previous rules.

A more recent model which I will call the ‘Dual-Processing model’ makes a strict distinction between regular and irregular morphology and claims that the two types of morphology are processed in entirely different ways (Pinker 1991, Marcus et al. 1992, Marcus et al. 1995, Clahsen & Rothweiler 1992). Regular morphology is handled by concatenating, symbol-manipulating rules which countenance no lexical restrictions, while irregular morphology is represented in the lexicon by means of associative networks, as proposed in Bybee & Slobin (1982b) and Bybee & Moder (1983). The regular rules apply in various ‘default’ or ‘emergency’ circumstances, characterized as cases in which the inflected form is unknown (as in the case of new or derived forms, memory lapses, etc.). In this model, regularity and productivity are equated. However, no way of predicting or explaining which patterns are productive and which are not is offered. Nothing explicit is said about fusion, but Marcus et al. (1995) do characterize regular rules as ‘symbol-concatenating’ computations, which suggests that highly fusional patterns, such as those involving stem change, do not qualify for such rules.

3.1 The Diachronic Source of Morphology
The vast majority of affixes in the languages of the world evolve from independent words by the gradual process of ‘grammaticization’ or ‘grammaticalization’ (Heine & Reh 1984, Heine et al. 1991, Bybee et al. 1994). In the progression from a lexical morpheme to a grammatical one, changes occur...
in the phonological shape of the morpheme, its meaning and its grammatical behavior. A well-documented instance of this type of change is the development of the future tense in Romance languages such as Spanish and French. A periphrastic construction in Latin consisting of an inflected auxiliary *habere* "to have" and an infinitive yielded a meaning of obligation or predestination:

```
\text{amar \ he}
```

"I will/shall love"

The auxiliary reduces phonologically and comes to consistently appear after the infinitive (where previously it could occur in various places in the clause). In Old Spanish we find the construction indicating future:

```
\text{amar \ lo \ he}
```

"I will love him"

Later this possibility disappears and the auxiliary becomes an actual suffix to the verb:

```
lo \ amaré
```

"I will love him"

In this process the grammaticizing morpheme undergoes phonological reduction (e.g., from *habeo* to *he* to *é*), its position becomes fixed, it fuses with the verb, and the whole construction takes on a more abstract, grammatical meaning.

A similar process leads to the development of derivational affixes. However, in this case the process begins with compounding. If the same element occurs in a number of compounds, it can reduce phonologically and change semantically in such a way that it becomes a derivational affix. For instance, the Modern English suffix -ly derives from a noun, which in Old English was *lic(e)* meaning 'body.' The compound *mann-lice* originally meant 'having the body or appearance of a man' whence it generalized to 'having the characteristics of a man,' the modern sense of *manly.* Since -lic was used in so many combinations, it lost its stress and reduced to -ly by losing its final consonant. Its meaning had already generalized in Old English to sometimes just mean 'pertaining to' as in the form *heofon-lice* 'heavenly.' In Modern English -ly is used to derive adjectives, as in friendly, and to derive manner adverbs, as in cleverly, but it occurs in many other uses as well: consider daily, weekly, cowardly, possibly and so on. Most derivational affixes in English and other languages can similarly be traced back to independent words where evidence is available.

The process of grammaticization is not discrete, but continuous: grammaticization in the form of semantic change and further phonological reduction and fusion continues even after grammatical status is achieved, and even after affixation occurs. This means that we can categorize morphemes for their 'degree of grammaticization.' Non-affixed forms such as auxiliaries are less grammaticized than affixes; affixes that are more reduced (e.g., shorter), that cause changes in the stem or undergo changes caused by the stem are more grammaticized than morphemes that are unvarying. As one instance of this continuing development, consider the Spanish future forms discussed above. Some time after affixation had occurred, the new suffixes began to condition changes in certain verbs. Thus the combination *veniré* "I will come" changed to *vendré; quereré* "I will want" changed to *querré; teneré* "I will have" changed to *tendré. *Such changes can be taken to indicate increased fusion between the stem and the suffix.

### 3.2 Inflection

In considering how productivity, regularity and degree of fusion are related, let us take up inflection first. A newly developed inflection was formerly a full word separate from the stem, so it will exhibit low degrees of fusion at first, and only when the affix and stem have been joined for some time will phonological processes begin to take their toll, gradually eroding the morphotactic transparency of the combination. Similarly, new formations will be relatively free of lexical idiosyncrasies, as these also take time to develop. Thus regularity and lack of fusion correlate for diachronic reasons.

Productivity can be said to be a prerequisite for the development of

---

2 There is at least one type of instance in which lexically conditioned allomorphy can be built in to the grammaticization process from the beginning: where two auxiliary constructions were present from the beginning, but merge into the same grammatical construction. Thus the use of both 'have' and 'be' auxiliaries in perfect and perfective constructions in French, German, Dutch and related languages could be regarded as a lexical idiosyncrasy or irregularity of some verbs.
inflection: affixation of a grammaticizing construction does not occur until that construction has achieved some measure of lexical generality, which implies that it can apply to new bases. When a new grammatical category is developing it is usually the case that it becomes fully general, applying to all lexical items of a general class (such as all verbs or nouns), and totally replacing older inflections. However, there are also cases where the new construction divides the lexical domain with an older one but without signaling any semantic difference. The ‘layering’ (Hopper 1991) thus created yields allomorphic variation between the old construction and the new one.

An illustration may be found in the English Past tense system, where the older means of signaling past remain in the verbs that have vowel changes in the stem, despite the development of a new synonymous construction in which the Preterite form of the ancestor of the verb do is suffixed to the verb.3 The newer construction is more productive, more regular and more transparent than the older one. Evidence suggests that the new construction began as a periphrastic causative construction, verb + dyde, and gradually spread to more verbs, losing its causative meaning, but retaining its past sense. Its uses included providing a Past form for new verbalizations and borrowings. Because the two means of forming Past in English represent an older and a younger layer, the former pattern is irregular, largely unproductive and has a greater degree of fusion. Thus these properties correlate for diachronic reasons.

Layering is a fairly common phenomena in inflectional languages. Another interesting example occurs in the Saharan language, Kanuri (Hutchison 1981). Compare the Class 1 verb bu- “eat” with the Class 2 verb fəlɛ- “show, point out”. The Class 2 verb root is easily separable from the suffixes, while the Class 1 root is not.

Class 1: bu- “eat”   Class 2: fəlɛ- “show, point out”
    bukĩn    fəlɛnɡĩn
    bũmũn    fəlɛnʃũmũn
    zẽwũn    fəlɛʃũn
    bũyẽn    fəlɛnɤn
    bũwũl    fəlɛmũwũl
    zũwũn    fəlɛzũl

3 The existence of ablauting verbs in Germanic or vowel and consonant changes in Semitic languages does not constitute a counter-example to the claim that all morphology develops via grammaticization. In all cases where the sources of stem changes are known, phonological changes conditioned by affixes create stem changes and then delete (e.g., umlaut in English and German). It is reasonable to assume that stem changes existing since prehistoric times have the same type of source.

Class 1 contains about 150 verbs, many of which have irregularities, and the class is unproductive; Class 2 comprises the rest of the verbs of the language, which are regular, and this class is productive. Class 2 was formed by adding the forms of the Class 1 verb nɡũn “say, think” to the verb root. Hutchison argues that the construction originated with onomatopoetic forms, parallel to English go boom, but with the form “say boom”. The construction provided a means of forming verbs from adverbs and nouns, and was used with compounds, derived forms and loan words from Arabic, Hausa and English. Because Class 2 represents a newer formation, it is more regular, less fused and more productive than the older Class 1. The Kanuri case is very similar to the case of English Past tense.

Two factors determine productivity: type frequency and the lack of restrictions, both phonological and semantic, on the application of a pattern. When a new construction is developing, it does not always have a high type frequency from the beginning. However, if it is free of phonological and semantic restrictions, it will be available to apply to new words, derived and borrowed, and its type frequency will in this way increase, making it more and more productive.

3.3 Derivational Morphology

For derivational morphology, as mentioned above, productivity applies to individual morphemes. Relative productivity is determined by type frequency and the relative lack of phonological and semantic restrictions on application. Since derivational morphology does not have to be obligatory, it is possible to have derivational patterns that apply in a very restricted domain, such as the English suffix -dom, which forms an abstract mass noun from a concrete one indicating a person’s position or status. The new abstract noun indicates either the domain, realm, or condition bestowed by that status (kingdom, serfdom). In a few cases it is used on adjectives to produce an abstract noun (freedom, wisdom, boredom). Its productivity is restricted by its meaning and the semantic conditions on its application. It does show some productivity, as evidenced in a recent issue of Time magazine, where the word supermodeldom was coined to describe the status and domain of a 16-year-old model, whose perks included getting invited to more parties.

In comparison, the well-studied suffix -ness seems to have very few semantic, phonological or morphological restrictions concerning the adjectives to which it can apply, and it is thus one of the most productive of all English derivational affixes. In the case of derivation, then, type frequency will be closely tied to the extent to which there are restrictions on the application of an
affix to a base.

There is also layering in derivation: older affixes will be more tightly fused both semantically and phonologically to their bases, having undergone more sound change and semantic change; they may also be less productive and more irregular than newer ones. For derivation, however, the relation of productivity to transparency may work in two ways: new formations that are quite transparent may be highly restricted, but gaining in productivity; older formations that are fused may be losing productivity and thus gaining arbitrary lexical restrictions.

3.4 The Synchronic Relations Among Productivity, Regularity and Transparency

The preceding sections have demonstrated that there is a diachronic explanation for the fact that the most productive morphological patterns are also usually the most transparent and regular. The question to be discussed in this section is whether or not there is any synchronic psycholinguistic relation among these properties, such that e.g., patterns that are more transparent have a higher degree of productivity because they are somehow easier to process and acquire (Dressler 1985a:318). Berman & Clark (1987) have argued that Hebrew derivational patterns that exhibit the greatest degree of stem change are also the most difficult for children to acquire. However, it is also the case that such patterns have the lowest type frequency, and it could be their lower type frequency that makes them more difficult. Since it is generally true, for diachronic reasons, that greater fusion and low type frequency correspond, it is difficult to know which factor is impeding acquisition.

To review, we are considering the following three properties, one of which breaks down into two determining features:

- Productivity: (i) lack of motivated restrictions
- Regularity: (ii) high type frequency
- Transparency: (iii) lack of arbitrary lexical idiosyncrasies

There is evidence, in terms of synchronic distribution and experimental data, that some pairs of these properties are independent of one another.

First, consider a case which exhibits lack of fusion and lack of motivated restrictions, but has low type frequency and arbitrary lexical distribution: the German plural in -s. German has several pluralization patterns, as shown in Table 1. The one with -s has a very low type frequency, but could be regarded as the one with the least fusion (since it never has an effect on the stem) and the one most free of phonological restrictions. This pattern is only marginally productive: Köpcke’s (1988) nonce-probe task showed it to have a relatively low rate of use. It is, however, used on about half of recent loan words (Köpcke 1988) and in pluralizing proper names and words with very unGerman phonological structure (Marcus et al. 1995). This -s plural in fact was introduced into German through loan words. Janda (1990) predicts that it will become the most productive pattern for plural in German, taking over the territory of its competitors. This case shows that high type frequency and lack of restrictions do not always go together; however, type frequency is likely to increase in this case since the pattern does not have phonological or semantic restrictions.

Table 1: German noun plural formation in the 200 most frequent nouns

<table>
<thead>
<tr>
<th>Affix</th>
<th>type frequency</th>
<th>singular</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>-(e)n</td>
<td>42%</td>
<td>die Strasse</td>
<td>die Strassen</td>
<td>“the street”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>die Frau</td>
<td>die Frauen</td>
<td>“the woman”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>das Bett</td>
<td>die Betten</td>
<td>“the bed”</td>
</tr>
<tr>
<td>-e</td>
<td>35%</td>
<td>der Hund</td>
<td>die Hunde</td>
<td>“the dog”</td>
</tr>
<tr>
<td>(+ umlaut)</td>
<td></td>
<td>die Kuh</td>
<td>die Kühe</td>
<td>“the cow”</td>
</tr>
<tr>
<td>zero</td>
<td>12%</td>
<td>der Daumen</td>
<td>die Daumen</td>
<td>“the thumb”</td>
</tr>
<tr>
<td>(+ umlaut)</td>
<td></td>
<td>die Mutter</td>
<td>die Mütter</td>
<td>“the mother”</td>
</tr>
<tr>
<td>-er</td>
<td>10%</td>
<td>das Leben</td>
<td>die Leben</td>
<td>“the life”</td>
</tr>
<tr>
<td>(+ umlaut)</td>
<td></td>
<td>das Kind</td>
<td>die Kinder</td>
<td>“the child”</td>
</tr>
<tr>
<td>-s</td>
<td>1%</td>
<td>der Wald</td>
<td>die Walden</td>
<td>“the forest”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>das Auto</td>
<td>die Autos</td>
<td>“the car”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>der Park</td>
<td>die Parks</td>
<td>“the park”</td>
</tr>
</tbody>
</table>

(based on Janda 1990)

Thus (i) and (iv) are each independent of (ii) and (iii).

Bybee & Newman (1994) have separated transparency from productivity experimentally and shown that there is no particular preference for morphotactically transparent formations, provided that the fused formations have sufficient type frequency. The experiment consisted of four artificial mini-languages in which sixteen nouns and their plurals were learned by English-speaking subjects. In each language, half the plurals were formed with suffixes and half with stem changes. In one language, both patterns were the same for all the nouns they applied to (i.e., they were both regular); in another condition, the
suffixed plural, but the stem changes exhibited four different patterns (i.e., they had a low type frequency and lexical idiosyncrasies); in the third language, the reverse was true, i.e., the stem changes were regular and the suffixes were irregular, exhibiting four allomorphs. In the final condition, both stem change and suffixes were irregular. After the subjects learned the sixteen nouns of the language, they were asked to supply plausible plurals for eight new nouns.

We hypothesized that the subjects would not necessarily generalize the suffixes, falsifying Dressler’s theory that suffixes (transparent formations) are more natural than stem changes (fused formations), but that subjects would use whichever pattern was regular in the language they had learned. The results showed no particular favoring of the suffixed plurals; subjects supplied stem changes and suffixes in approximately equal numbers (overall 48.75% suffixes and 50.25% stem changes). The subjects supplied new stem change or suffixed plurals in approximately the proportions they were present in the input data — about half and half. This result supports a usage-based model.

Moreover, we found that the regularity of the suffix affected productivity: in the conditions in which the suffix was regular (had no lexically arbitrary allomorphy), more suffixes were supplied for new forms; in the conditions in which the suffix was irregular, more stem changes were supplied for new forms. This result shows that productivity (based on type frequency and lack of restrictions) is independent of the degree of fusion of the pattern.

Are (i), lack of motivated restrictions, and (iv), lack of fusion, independent? There are formations that have motivated restrictions that are not fusional, but transparent, as the Kanuri “say” verb at an early stage, or any newly grammaticized construction. Patterns that are highly fusional but lacking in motivated restrictions appear in Semitic languages, as for instance, the jamic plural of Arabic (McCarthy & Prince 1990).

Are (ii), high type frequency, and (iii), arbitrary restrictions, independent? In a sense they are not, because any restriction reduces type frequency. Of course, there are systems where there are several patterns each having fairly high type frequency, but with each one applying to an arbitrary portion of the lexicon, for example, German plurals or Hausa plurals (Lobben 1991).

Synchronously, then, these properties appear to be independent of one another. In the Network model, this is just how they are treated, but in two other current models, the Dual-Processing model and the Level-Ordering model, these properties are treated as though they are structurally related to one another. The following two sections describe these models and how they view productivity, regularity and fusion.

4. The Dual-Processing Model

A proposal emerging from the child language and psycholinguistic fields sets up a strict distinction between regular and irregular morphology, treating them in two distinct modules of the grammar (Pinker 1991, Marcus et al. 1992, Marcus et al. 1995, Clahsen & Rothweiler 1992). Irregular morphology is treated as in the Network model with irregular forms listed in the lexicon and organized into patterns describable by emergent schemas. Regular forms are derived by a ‘symbol-concatenating rule’ that acts in a generative fashion on a base form to produce the regular derived form. While irregular patterns are highly affected by actual lexical distribution, the regular ones are not. Prasada & Pinker (1993) have shown experimentally that the lexical distribution of English irregular Past forms such as rung, strung, struck, etc. affect subject’s responses to nonce verbs, while the lexical distribution of the regular Past does not affect nonce form application. However, since regularity and type frequency are confounded in English, it is possible that this effect is due to the high type frequency of the regular Past, plus its lack of motivated restrictions.

Marcus et al. (1995) argue that the case of the German s-plural serves as an example in which regularity and type frequency are independent. However, in this case, despite the lack of phonological restrictions, a strong effect of lexical distribution can be observed. Köpcke (1988) found that the s-plural was much more likely to be applied to nonce words ending in full vowels than to nonce words ending in schwa. This result reflects the lexical distribution in which almost all words ending in full vowels have a plural in -s.

Clahsen & Rothweiler (1992) have argued that the case of the German Past Participle also separates type frequency from regularity. Here the productive affix is -t as opposed to -en. Clahsen and Rothweiler argue that both affixes have approximately equal type frequency. Their counting method takes verbs with separable prefixes to be distinct types, as though one were counting each of the following English verb-particle combinations as separate verbs: break, break up, break down, break out, break in. If the German verbs are counted as one would count English verbs, that is, counting break only once, then the distribution of regular and irregular types is similar to that of the English Past tense: out of 1258 base verbs listed in Ruoff (1981), there are 150 Past Participles in -en and the remainder use the suffix -t (Bybee 1994). In this case as well, type frequency predicts the productivity of -t.

This model provides no link between productivity and language use. The best it can offer as to why a certain pattern is treated differently from others (handled as a symbolic rule) is that human beings are endowed with an innate neural architecture that makes two types of processing possible. This model
treats the English s-plural and the German s-plural both as symbolic rules and provides no means of accounting for the difference in their productivity.

As I mentioned above, the Network model also is 'Dual-processing' in a sense: a regular form may either be accessed whole from the lexicon or derived by applying a schema to a base, depending on the token frequency of the form. The Network model does not base the difference in derivational mode on a structural difference, but rather on a usage difference - high frequency forms are accessed whole and low frequency are not (see Losiewicz 1992 for experimental evidence that this distinction is based on word frequency).

5. **Level-Ordered Morphology**

Level-Ordered morphology is based on the insight that both morphological and phonological patterns have different degrees of involvement with the lexicon (Kiparsky 1982a). We have already seen that for morphological patterns this is a result of the way grammaticization proceeds, creating layers of new morphology on top of old morphology. For phonological rules there is a diachronic explanation as well. Phonological rules begin as phonetically-motivated processes and gradually become more and more involved in the morphology and lexicon. They tend to lose their phonetic motivation and remain only as fossilized alternations in assorted morphological environments. Thus phonological processes also create layers of newer patterns on top of older ones.

Level-Ordering models this diachronic layering by recreating it in a synchronic grammar. This works fairly well for English where the deepest level is largely comprised of morphological patterns that entered English through the borrowing of French words, many of which were already morphologically complex and thus carried with them fossilized phonological processes that had occurred in French. Thus certain old phonological processes of French (like Velar Softening) as well as the older sound changes of English (like the Great Vowel Shift) affect words at this level and not at later levels, where the more productive, largely Germanic patterns are described.

However, in other languages, which may not have this bifurcation in the lexicon, it is not so clear that rules are positioning so neatly on distinct levels. (See Kaisse & Hargus 1993a for an overview of the recent literature, which seems to show that there is no general agreement about how phonological and morphological rules interact.) Even in English there are problems, especially with affixes that have properties of both levels, such as -ity (Aronoff & Sridhar 1987). Still, given the way language change proceeds and the way new phonological processes and grammatical affixes are added, any theory that recapitulates diachronic development will need to show some type of interleaving of morphological and phonological rules.

In addition to the fact that there is no consensus on the internal structure of the theory, there exist a set of problems that affect all theories of this type. As in other versions of generative grammar, Level-Ordered morphology makes the highly improbable claim that underlying forms of high frequency items, such as damn, are affected by low frequency items such as damnation. The underlying representation of damn with a final /n/ is set up to accommodate the derivation of damnation, just as an underlying representation for sign as /sign/ is set up to accommodate the derivation of signature (Borowsky 1993). I consider it very unlikely that a learner would modify her/his underlying representation for an established and highly frequent word simply because s/he has now acquired a low frequency word which might be related to the established word.

Productivity is modeled by the Elsewhere Condition, whose premise is that more specific rules apply before, and block, more general ones. The Elsewhere Condition has an effect similar to the criterion of the openness of the schema, described above for the Network model. But the effects of type frequency, the most significant factor in determining productivity, cannot be represented in a theory in which rules exist independently of the forms to which they apply and which does not take into account the way language is used.

Another serious flaw in the theory is the lack of a way to account for the semantic composition of morphologically complex words. It is often observed that words formed with affixes at the deeper levels do not have compositional meaning. This should be taken as sufficient evidence that such words are listed in the lexicon as units and not decomposed morphologically. However, since practitioners of this theory still want to do morphological decomposition, then they are obliged to come up with a semantic theory that can predict the results of morpheme combination.

In my view, Level-Ordering reconstructs diachronic development of form, while neglecting meaning, and produces analyses that have no synchronic reality and no prospect of synchronic verification. This is precisely the reason that it is so difficult to find a model of level-ordering that works consistently. A usage-based model, such as the Network model, takes into account the way forms are used synchronically and is constructed to bear a direct relation to the surface forms of the language. In the following I compare the Network model to Level-Ordering models.

In the Network model it is use rather than structural criteria that determines representation. A lot of what is studied as morphophonology is diachronic
for English by the Level-Ordering model is not compositional in the Network model at all; rather, the ‘derived’ words are listed in the lexicon. The evidence in favor of this position is the simple fact that most of these derivational formations have unpredictable meaning. Many also have unpredictable form (including the irregular inflection), because alternations that are fossilized have many exceptions and idiosyncrasies of application. To the extent that there are valid morphological relations among such lexical forms, these can be captured in the lexical connections, as shown in Figures 1-3.

One insight of Level-Ordering is that more peripheral phonological and morphological rules do not have access to the internal structure. This is modeled in Level-Ordering by erasing the morphological bracketing at the end of each level. Thus each new morphological or phonological application treats the item it is applying to as an unanalyzable whole. This is precisely what would be expected if the input to new morphological formations were items that were already stored in the lexicon. In fact, only because generative models insist upon maximal decomposition of words does it come as a surprise that outer, newer patterns do not have access to older, more internal composition. In the Network model where words are entered into the lexicon as they are formed, there would be no reason to expect that the diachronic composition of a form would affect its synchronic behavior.

Kiparsky (1982a) has pointed out that even when irregular forms are the heads of compounds the compound as a whole is often inflected as though it were regular. Thus in the Toronto Maple Leafs the plural is regular rather than the irregular *Maple Leaves. In the Network model, when a compound is formed that is a noun, it is associated with the schema for plural nouns, and treated like an unanalyzable noun. The irregular plural [livz] is listed in the lexicon along with [lif], but it is not accessed for the new compound plural.

A verbalized noun is associated with the verb schemas, and the strongest schema for Past tense is the -ed pattern. For instance, the verb to ring formed from the noun ring has no access to rung, but will rather participate in the regular Past schema. Thus the phenomena that Level-Ordering accounts for are also accounted for if words are stored in the lexicon as they are created and associated with appropriate morphological schemas. (See my discussion of Heather Goad’s contribution later in this volume for a further explanation of how these facts are accounted for in the Network model.)

Another fact easily accounted for by the lexical storage of derived words is the minor tendency for irregular plural nouns to appear in English compounds while regular plural nouns usually do not. Thus mice-infested is possible because mice is a lexical item, but rats-infested is not acceptable because it contains the regular plural marker. Note that the irregular plurals of English all designate objects that tend to occur in groups of more than one: mice, teeth, feet, geese, oxen, or they are plurals of extremely high frequency nouns: children, women, men. Thus the irregular plurals are not only listed in the lexicon but they tend to be highly accessible due to their token frequency, either in absolute terms or in relation to their singulars. It should be added, however, that compounds with plural nouns in them, even irregular plurals, are extremely rare in English.

In addition the Network model provides a way of accounting for two usage-based phenomena in morphology not accounted for in generative models. One is productivity, which we have already discussed: productivity is directly determined by type frequency, and type frequency affects the lexicon by strengthening the schemas that are used with greater numbers of distinct items. The other usage-based phenomenon is the maintenance of irregularity in forms with high token frequency. Because the use of items affects the lexicon, items that are frequently used have strengthened representations and forms that are less often used have weaker representations and may even fade from memory. Irregular inflections that are highly available in the input will be easily accessible and will not be regularized. Weaker, less accessible irregular forms are more likely to be replaced by regular formations (Bybee 1985).

6. Language Acquisition

An interesting controversy that distinguishes the structural or generative models from the associative or usage-based models has arisen in the field of child language development. The linguistic behavior of little children seems to point to the existence of structure of certain types, but since structural theorists do not believe that young children have enough exposure to data to acquire the structure in question, they propose that these structures are innate. The usage-based approach would claim that children have ample exposure to store certain items in their lexicon, or more generally have representations in memory of words and phrases they have heard and used. The ‘structures’ in question do not have to be innate or acquired by the child, because in fact they do not exist.

Gordon (1985), following Kiparsky (1982a), observed that even though English compounds can contain irregular plural nouns (mice-infested), such compounds are extremely rare and the use of the singular form in the

5 The same account is given in Marcus et al. 1992.
compound is much more common. This means that young children will have been exposed to many compounds with singular nouns, but few, probably none, with irregular plural nouns. Gordon hypothesized that if children correctly assigned irregular pluralization to Level 1, then irregular plurals could be used in children's compounds, since compounding is assigned to Level 2. Gordon created an experiment in which children were exposed to plural nouns in the context of multiple objects and asked to label a puppet as an X-eater. The results showed that children could use irregular plurals in the compound, e.g., teeth-eater, but not regular plurals; that is, they produced bead-eater not beads-eater.

Since there is little chance of exposure to compounds such as teeth-eater, and since children produced these forms (given a bias in the experiment towards plurals) but did not produce regular plurals in compounds, Gordon reasons that they have correctly placed irregulars at Level 1, and that they also correctly use Level 1 as input to compounding. Given the lack of appropriate input to the child and the fact that children between the ages of 3 and 5 gave consistent responses, with no apparent learning taking place, Gordon suggests that the Level-Ordered structure might be an innate property of the lexicon (p. 87).

A much simpler explanation for the facts that does not require so much innate architecture is that children have stored the irregular plurals in their mental lexicon and can access them for compound formation. This is the account given in the Network model and also in the Dual-Processing model (Marcus et al. 1992:142). Since the experimental situation primed the children with plural forms rather than with singulars, and since for English irregular plurals there is some phonological distance between singular and plural, the plural was more available in the situation and was thus used.

Several factors could explain why these same children did not use regular plurals in their compounds even though in the situation regular plurals were primed also. First, the children probably know and use a number of compounds of the relevant type, none of which includes an internal regular plural marker. The formation of new compounds will be based on existing compounds; in the experiment the children accessed their compound-forming schema, which does not include any plural inflectional affixes. Second, these children were successful at forming compounds, which suggests that they understand the basics of compound formation, which is that the component words of compounds, especially the non-head element, is decategorialized (Hopper & Thompson 1984). In the case of nouns this means that the noun is unable to refer, and thus does not carry the inflection that a referring noun would. The lexical noun in a compound represents the semantic essence of a class of entities, but without referring to a particular entity.

After being primed with regular plural nouns, we might think of what the children do as removing the inflectional affix. After all, they will not have been exposed to compounds with plural -s in them, e.g., it is Cookie Monster, not Cookies Monster, despite the fact that he eats lots of cookies. Evidence for a strategy such as this is the fact that 69% of the responses for scissors singularized it to scissor to put it in the compound. If there is a prohibition against a plural affix (rather than plural meaning) in the compound, that would explain why irregulars with affixes such as children and oxen do not sound acceptable in compounds: child-eater is better than children-eater, or, to compare across items, mice-eater is better than children-eater. Unfortunately, Gordon did not include the items children and oxen in his experimental material.6

Gordon also included the pluralia tantum nouns, clothes, pants, glasses and scissors in his experiment. Level-Ordering predicts that these nouns will retain their lexical plural marker, since it is derived at Level 1. The Network model would predict some variation with these items, since they do not have singular forms, and yet they include the regular plural marker which does not appear in compounds. If the children are using the strategy of removing the inflection before forming the compound, then they will run into trouble with clothes which would be [klaubs] with the inflection removed, and with pants and scissors, which do not have singular bases, and even with glasses, whose singular base has a different meaning. In the Level-Ordered account, there is no conflict: these plurals are derived at Level 1 and should fit nicely into the compound, just as mice and teeth do.

Gordon's results on the pluralia tantum do not have the uniform and categorical nature that one would expect if Level-Ordering were innate. Instead, clothes and pants tend to go into compounds as plurals, while glasses and scissors tend to have the 'inflection' removed. The Level-Ordered account fails for scissors, since there is no such singular. One would have to posit an inflection-stripping strategy, as I have for the Network account, but if an inflection-stripping strategy is to be invoked, then it might as well be invoked to explain the difference between regular and irregular nouns, and level-ordering is not necessary.

Another problem for the Level-Ordering account (acknowledged by

---

6. Sternberger 1994 argues that regulars and irregulars in English are not parallel phonologically, and Gordon's attempts to separate them using real English words fails because the morphological is confounded with the phonological.
Gordon) is the fact that Dutch and German do allow some plural inflections in compounds. This can be viewed as a simple difference between English and her sister languages in terms of the patterns for compound formation; in Dutch and German there is no restriction against a plural affix in a compound. For Dutch, the most common and productive plural affix is -en, and this affix is found in compounds such as paardenhief "horse-PLURAL#thief"; and tandenborstel "tooth-PLURAL#brush". Such compounds can be produced if all Dutch plurals are lexical, as the Network model would predict, and Dutch does not restrict plural affixes inside compounds.

German also allows plural affixes in compounds, with the exception of -s, which never occurs in compounds (Autobahn, not *Autosbahn). Clahsen et al. (1992) argue that the facts of German can be described by positioning the different plural affixes of German at different levels: -er and -e are at Level 1, -en at Level 2 and -s at Level 3. Compounding takes place at Level 2, so all affixes except -s can appear in compounds.

Evidence for Level-Ordering can be found, they claim, in data on compounds from 19 dysphasic children they studied. These children never put -s in compounds, as would be predicted since adults never do either, but eight of them left -n out of a compound requiring it at least once. Clahsen et al. argue that the omission of -n in the compound can be accounted for by assuming that these eight children had placed -n on Level 3 rather than Level 2. Six of these eight children overregularized with -n more often than with -s, a finding which supports a Level 3 assignment of -n. None of these children omitted the other plural markers -e and -er from compounds, as predicted by the assignment of these affixes to Level 1. Thus Clahsen et al. argue that children, even dysphasic ones, have access to Level-Ordering, which, then, must be innate.

One problem with this account is that none of the children always omitted -n from compounds. Assignment of a morphological rule to a level is a categorical decision, not a probabilistic one. Thus these children would have to have been moving -n pluralization from one level to another during the period of the study. Another problem is that the one nonimpaired child studied in Clahsen et al. did in fact produce a compound that omitted the plural -er. Simone (Miller 1976) produced both the correct form Bilderbuch "picture book" and the incorrect Bildbuch. A general problem with applying Level-Ordering to child language is that Level-Ordering predicts discrete, categorical behavior, which is rarely found in children, or adults for that matter.

Constructing a full Network account would require more information about the frequency of occurrence of compounds of different types in German than is available to me at the moment. However, it is possible, even in the absence of that data, to outline a description of the acquisition of compounding.

In the Network account, existing compounds have lexical representations, in keeping with their unpredictable semantics and their propensity to undergo further semantic and phonological change as a unit. Children acquiring German compounds will have some with singular and some with plural nouns in them. If these are not yet all strongly represented, there could be retrieval errors involving omission of inflections internal to the compound. Children will also be producing new compounds using existing patterns as schemas. This means that some of their formations will have singular nouns in them and some will have plural.

Why would -n be commonly omitted from compounds but not -er and -e? Here is the point at which the token frequency of the nouns and compounds in question becomes important. First, a highly frequent compound is easy to retrieve and is not likely to be produced incorrectly. It might turn out that the compounds with -er and -e used by children are of high token frequency. Second, a highly frequent plural, even inside a compound, might be less likely to lose its plural marker. Again, the relevant vocabulary statistics are not available, but they are likely to show that the more lexically restricted -er and -e tend to occur on words of high token frequency, while the more productive -n occurs on words of a wide range of token frequencies.

Thus, the Level-Ordering account, which describes rules as more or less deeply embedded in the lexicon, can be matched by an account without ordered rules, which represents the lexical involvement of morphological patterns as a matter of direct lexical representation.

7. Conclusion
In contrast to structuralist and generativist theories, the more modern usage-based theories of language find significance in the fact that linguistic behavior is continuous and not discrete, probabilistic and not categorical, and dynamic rather than static. Language does not exist in a mental prison, insulated from real-world factors of meaning and use. Language is a social instrument that is in constant use, and the varying details of this use have an effect on storage and processing, creating and recreating the mental associations that we study as grammar. Morphophonology is no different from morphosyntax in this regard, and its study in this framework could lead to answers to some of the persistent questions of the field — what determines productivity, what is its relation to regularity, and how are processing and storage affected or not affected by the degree of fusion among morphemes.
References

---. 1994a. “Le mécanisme des stratégies de réparation en phonotactique


--------. 1996b. “A Functionalist Semiotic Model of Morphonology”. This volume.


-------- & --------. 1996. “Quelques avantages d’une linguistique débarassée de la morpho(phono)logie”. This volume.
Gvozdev, Alexandre N. 1949. Formirovanie u rebenka grammatičeskogo stroja russkogo jazyka [The building-up of grammatical structure by a Russian child]. Moskva: Izdatel'stvo Akademii Pedagogičeskogo Nauk RSFSR.


--------. 1983b. “Morphemes Aren’t Something that Grows on Trees: Morphology as more the phonology than the syntax of words”. Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax ed. by John Richardson et al., 79-95. Chicago: Chicago Linguistic Society.


Mundartforschung, suchungen an der Gegenwartssprache. Frauenfeld: Huber & Co.
Munda, D. dissertation, Austin: Univ. of Texas.

REFERENCES

--------. 1995b. "Hierarchies All the Way Down". Ms., Univ. of Chicago.
References

Wonderly, William. 1951. "Zoque II: Phonemes and morphophonemes". IJAL 17:105-123.