Chapter 3

Modeling Morphology

3.1 Linguistic Theories

As has already been discussed in Ramaswamy, V. (2000), theories of morphology greatly differ in their view with respect to the morpheme, as a building block or as a basic unit of morphological analysis. The relevant issues in the pursuit of research in morphology therefore remain in focus with respect to the identification and the realization, of the unique and unified functional aspects of the basic morphological unit.

Hockett’s, two models of morphology viz, the Item and Arrangement (1958), and Item and Process (1958) embody Structuralism in morphology. A morpheme, in that sense, was 'the smallest individually meaningful element in the utterance of a language' (Hockett, 1958: 123).

The Item and Arrangement (IA) assumes a concatenative approach wherein morphemes are lexical units, morphology is an agglutination of such units, and words are seen as linear sequences of morphs. The central observation of the model is that certain forms reflect a partial phonetic-semantic resemblance to other forms. Describing the morphological patterning of a language would in consequence involve — a specification, of the inventory of morphemes (items), of the possible sequences in which these morphemes can occur (arrangement), and defining those morphs through which every morpheme of the language is realized (a morphological link). A word is analyzed as involving no process at all, but only a permissive sequence of morphemes. The IA model is thus essentially morphemic, but non-processual.

The Item and Process (IP) model works on the conception that word-forms constitute paired features of structure and function, or rather form and meaning, and that elements undergo a variety of modifications, through transformations, to result in newer forms. Morphology is regarded as a set of derivational processes, acting on morphemes or words, producing new word-forms. Hence in IP, there is only one recognized underlying basic allomorph that contextually takes the form of an allomorph.

In contrast, the Word and Paradigm (WP) (Robbins 1963), model of morphology takes the *word* operating within a paradigm of variables, as its basic unit. The WP model is distinct in that it focuses primarily on each word as a whole, and on its overall effect to the unique set of properties that it expresses. The model evades complexities since it divides the word into two parallel non-overlapping units: one for formal elements, and the other for functional properties. In the WP framework, it is a morphophonemic (MP) rule, a conglomerated set of morphosyntactic and morphosemantic operations that align with a word, and not a phonological (P) rule. Parallel to instances in analytical phonology, there
exist no absolute constraints that might limit the nature or number of correspondences between morphological properties and morphological forms.

Stump (1998) describes the word and paradigm approach to inflection as follows: "a rule's applicability to a stem X is conditioned by the set of morphosyntactic properties associated with X, by X's phonological form, by X's membership in a particular morphological class, or by some combination of such factors." Rules of inflection here are specified into sets each of which applies to only one of its kind. With regard to the role of paradigm, the supposition is like this: "a is a cell in the paradigms of lexemes belonging to some class C, and the paradigm function for cell a is that function f, such that for each L∈C, f, applies to the root of L to yield the word form occupying a; one can then say that the sequence of rule blocks in a language may vary according to the definition of its individual paradigm functions."

Generative grammar developed a remarkably powerful device in grammatical theory, i.e., transformations, a computational tool connecting different linguistic levels. Morphological processes were rather derivational mechanisms than transformational operations. Leading to the proposal of the Lexicalist Hypothesis (Chomsky 1970), which expressed the need for an autonomous component of morphology in Generative grammar. The theory addressed two central claims about the organization of the grammar of a language: (1) that all morphological operations are governed by the lexical component; (2) that all morphological facts are accounted for, by specific rules called Word Formation Rules (WFRs).

Halle's (1973) proposition was the first model consistent with the Lexicalist Hypothesis. As to the association between morphology and native intuition, Halle developed a model constituting an inventory of morphemes, a set of WFRs, a Filter to rule out possible but nonexistent words analyzed by WFRs, and a Dictionary containing all and only the actual words of the language, in all their paradigmatic forms, ready to undergo lexical insertion. A major point of disagreement, against this however culminated in what was called later, the Word-Based model, as proposed by Aronoff (1976). Aronoff's hypothesis has the following claims to make: (1) All regular word-formation processes are word-based; (2) A new word is formed by applying a regular rule to a single already existing word; (3) Both the new word and the existing one are members of major lexical categories.

The Word-Based approach distinguishes two different kinds of rules, operating at two levels: Word Formation Rules (WFRs), and Readjustment Rules (RRs). According to Aronoff (1976), WFRs function not merely in deriving new words, they also account for their internal structure. Readjustment rules are those of the lexical component, which operate after the WFRs. Readjustment rules are optional in the sense that they need to operate only if the morphological environment demands so.

The Word Network model (Singh and Ford 1984) following a relational approach to morphology, establishes firmly the dependency of morphology on the lexicon. The approach defines morphology on two primary ideals: (1) the basis of morphological analysis is not the morpheme, but the word; (2) words, as units relate to each other through complex networks.

The Word Network model has certain fundamental assumptions as to what the nature of morphology is: (1) Rejection of 'multiple morphologies. No distinction is made between binary types such as inflectional vs. derivational, concatenative vs. non-concatenative, or non-linear etc. (2) Determining the 'word' as the basic unit upon which morphology operates. No morphological status is given to concepts like 'root', 'stem', 'lexeme', etc.
3.2 Psycholinguistic Theories

3.2.1 Storage and Retrieval of Morphological Structure from The Mental Lexicon

One general view of morphology describes it as dealing fundamentally ‘with’ the internal structure of the potentially complex words of a language. All of these words might not be actually existent, but all of these are governed by the set pattern of morphological operations. Different views of the ‘lexicon’ are held. Contrary to the general conception of the lexicon of a language, the lexicon has been treated as exclusively an individualistic, independent, internalized linguistic component. Either way, the lexicon can be best defined as ‘a list of existing items in the language, those that a speaker has to know because they are arbitrary signs: unpredictable in some way’ (Aronoff & Anshen 1998)

A priori, one might be led to infer that, since morphology trades in potential words, and the lexicon, only in actually existing ones, the two might not have anything substantially common enough to relate to each other. As a matter of fact, this could be proved a misconception, for two valid reasons: (1) both morphology and the lexicon, dealing primarily with words, serve the same linguistic function. (2) morphology and the lexicon, share a fundamentally interdependent relation.

Quoting Aronoff & Anshen (1998), “The difference between which words exist and which are potential is defined solely in terms of the individual’s lexicon and morphology.” That morphology and the lexicon proceed from each other is implied by the production of morphologically complex words. This is largely through the application of morphological rules, to actually occurring base words that are stored in a speaker’s mental lexicon. Another instance of lexical words being governed by rules of morphology is in the ‘inheritance of irregularity’ wherein complex words having conventional semantics exhibit a slight variation from their predicted sense. The morphology depends on the lexicon, however, inasmuch as the bases of morphologically complex words are normally lexical entries.

Several views of word recognition processes in morphological theory have been proposed as possible solutions to the question of how morphological structures are represented in the central lexicon. Various queries arise, in this regard.
I quote (McQueen & Cutler 1998), "Is the mental lexicon organized in a way which
codes morphological relationships? For example, is the fact that two words share the same
stem, or the same affix, coded in the lexicon? This is primarily a question of representation:
what information about the internal structure of words is stored in long-term memory and
how? But it is also a question of processing. When a word and its morphology are
recognized, does this involve contact with other entries (words and/or morphemes) in the
lexicon? What role does morphological structure play in the process of mapping perceptual
information, from spoken or written input, on to the mental lexicon? Again there are
questions of both processing and representation: what type of morphological parsing must
take place, and what form of access representations might be the product of such a process?"

Do humans actually parse words while they access it? And if they do, what theoretical,
logical or, statistical assumptions underlie these analyses? These remain another set of
questions. Regardless of what might determinedly answer these, or what can be decidedly
established about the question, whether there are as many morphologies as there are
speakers, is that, morphology provides a common universal framework across a number of
lexica, since rule patternings are unexceptional. Though a lexicon may be regarded as an
individualized conception, morphological processes capture generalizations that surface
evenly across different lexica. Differences in the lexicon need not imply differences in
morphology.

To study the differing views about the role of morphology in lexical access, the following
psychological models of morphology may be considered.

3.22 Kinds Of Models

A volley of questions can be raised with respect to the psycholinguistic modeling of
morphology. Does perception of morphological structure play a role in access, or arise from
it? Are stems and affixes stored and retrieved in the same way, in conjunction with each
other, in the same lexicon, or are different records and processes involved? As to the nature
of the lexicon and the representation of word forms therein, might the storage be so abstract
as to obliterate differences between written, spoken, read and heard forms of words? Or,
might it be considered as an assemblage of separate, but interlinked, specialist lexicons?

Such are the critical questions that relate to the status of morphological structure in
lexical access. The question of how words of a language are accessed from the lexicon is
rather decisive to how the human mind parses complex words. In an attempt to examine
mental morphology, a few vital psycholinguistic models of word-recognition were explored.
Different theories of lexical access have been proposed, a discussion of which follow suit.

3.2.2.1 STRING-SCAN MODELS

The Transition Network Model

Forster (1976) suggests a direct model of lexical access that tries to retrieve a word from the
mental lexicon based on the principle of manual dictionary search for a required string. The
mechanism is that of a transition network, which scans for the string from left to right,
sequentially progressing with each letter at a time, until the string is exhausted. At each phase,
the transition progresses if a match is found, with backtracking occurring wherever
necessary. Search for a string is according to that of a dictionary search, with phonological conditioning filtering an improbable output. A finite state network for the word determines if the progress is on correct lines.

Although economy has been noted in:

- **Merging of Pathways**: A merging of pathways is achieved by merely stating a single instance of a letter for each position in a word. A structural positioning of letter-elements would enable search paths to define the existence of particular entries.
- **Number of Pathways**: The instances of specified pathways through the network are less than the number of all possible transitions.
- **Structure of Pathways**: The structure of a pathway may often correctly determine the word form, even before the completion of the scan.

The disadvantages the model however faces, as Forster observes are:

- Non-words are more easily detectible
- Longer words consume a longer time to be accessed.

The model has been observed to fail in accounting for an aspect of lexical access that takes into account the word-frequency factor. The model in fact seems to be more of a phoneme-detection model, rather than a word-detecting one. Moreover, Garman's (1990:262) observation cannot be overlooked in what he points out as a shortcoming in the mechanism: "We know that the normal sort of dictionary is in the form of a book, in which all the separately listed items are spelled out fully; but this involves considerable redundancy in the interests of ease of use in turning over the pages and scanning their contents; and it is possible that this sort of organization is not required and actually too costly, as far as the mental lexicon is concerned."

**The Autonomous Search Model**

After the letdown of the direct method, Forster's (1979) alternative was a two-stage processing model with a serial mode of operation where the first stage ends before the second begins. In the first stage, the search is ordered by frequency of the lexical items. Unlike the earlier model, focus is on frequency-rank ordering rather than physical structure. Rather than a scanning of the input string, an abstract location marker in the second phase of operation helps gain access to the lexical item, based on a hint obtained in the first phase. The mechanism is similar to that of cataloguing. A master file holds the catalogued information, which aids further access and retrieval.

The Autonomous Search model has the following implications:

- High frequency words are accessed faster than low-frequency words.
- Rejection of non-words takes a longer time than the acceptance of real words.

Experiments however observe that the theoretical and logical tenets of the two-stage model aren't really promising.

**The Logogen Model**

Morton's (1969) analyzer claims prospective assumptions about the lexical identity of the input string apparent at each stage, with an aided correlation between phonological analysis and lexical access.
Words are realized on the basis of selectively tuned response characteristics. The system does not merely rely on the specific linguistic properties, but is also able to pick up signals to the presence of a word from extra-linguistic signals as well. These are the Logogens, the tuned perceptual devices that respond to sensory and semantic input. The sensory and contextual inputs interact, and give rise to outputs, to the cognitive system.

The Logogen model implies that high-frequency words have lower thresholds associated with them, and hence require less processing to yield access. Forster questions the efficiency of the filtering mechanism in this model. In other words high frequency items are more available both as correct and incorrect responses. Although this has been noted as truly a word-detection model that focuses on the auditory/visual properties of words, their contexts, and frequency ranking of words, the incapacity to reject incorrect analyses makes the model not very popular.

The Cohort Model

Designed by Marslen-Wilson & Welsh (1978), this has been an improvement upon the transition network, and the logogen models in which activation levels can be precisely stated. Instead of a partial activation of logogens in some unknown proportions, a preference of only two degrees of activation has been suggested, - zero or, full.

A complete activation of all members of a cohort occurs on the basis of the first item of the input sequence. Thereafter, instead of the summation of partial activations to individual threshold values, there would be a progressive elimination of fully activated items, as the left-to-right processing of the sequence continues. The point of recognition would be determined, not by the end of the string, but by the elimination of the last alternative letter from the cohort.

Hybrid features of the earlier models effect vital implications here:

- It embodies the sequential nature of the auditory signal directly into the nature of word-recognition.
- There is a ‘uniqueness’ point at which the item diverges in linear structures from all other words in the language (and the cohort).

The cohort model follows the letter-tree theory as discussed in Knuth (1973). Theoretically though, the mechanism and predictions of the model could not be confirmed by theorists like Katz et al. (1987).

3.2.2.2 CATEGORY-SPLIT MODELS

Henderson’s review of the Taft & Forster Model

This was the first time a psycholinguistic model makes explicit the role of morphological analysis in word recognition, specifically considering the model for written word recognition by Taft and Forster (1975). The process involves a stripping of affixes from stems ‘prior to’ lexical access. (See Fig. 3.1) Crucial to its operation is the identification of word boundaries. The following is a flow chart that illustrates the working of this model.

1 In Forster’s sense.
The model is a serial one that operates on a set of stages in an ordered fashion. Such a decomposition model, argue Taft and Forster, allows for economy of stored representations, since a stem need be represented, by merely a single instance of itself. Stripping operates at each level, and lexical search continues on the basis of the remnant stem. As conceived by Henderson, the model requires a lexicon with a set of stem-morpheme entries, and a set of composition rules activated by the stripped affixes, and which in turn determine the characteristics of particular word formations.

**Augmented Addressed Morphology**

This is in opposition to the Taft & Forster model, which believes in lexical access through morphological decomposition. Carama2za et al (1988) defend the theory of 'augmented addressed morphology', which supports the view that morphologically complex words simultaneously trigger both the whole word, and its constituent morphemes, in the mental lexicon. Lexical access via achieved through whichever form is encountered first, by the subject. This is in opposition to the Taft & Forster model, which believes in lexical access through morphological decomposition. Experiments have however shown that the predictions made are not in accord with results.
**Satellite Entries Model**

Lukatela et al (1988) hold the claim to the satellite entries model. The hypothesis states a basic allomorph that holds together all other allomorphs in a relational manner. This slightly differs from the IP theory of morphology, in the sense that instead of derivation of allomorphs from a base form, they are linked to each other relationally, in the WN fashion.

Sproat (1992) is of the view that the Lukatela model can be accommodated within a model of morphological parsing. Emphasis is on the base allomorph, through which all other allomorphs relate.

**Separate-Entries Model**

Another 'separate-entries' model (S. Andrews 1986; Fowler et al. 1985; Grainger et al. 1991; Schreuder et al. 1990; Schriefers et al. 1991, 1992) believes that the various forms of a word ought to be accessed via their surface form. This has its basis on the frequency of the various surface forms.

**Naive Decompositional Model**

The naive 'decompositional' model is on the lines of the Taft & Forster model that maintains lexical access through stems and affixes.

### 3.3 Implications Of Psycholinguistic Findings

The fact that native speakers have the internal knowledge of the morphological structure of their language and that they access the lexicon via morphological decomposition, has been established through psychological evidence (Caramazza et al. 1988; Taft & Forster 1975; Tyler & Nagy 1990; Holmes & O'Regan 1992; Laudanna et al. 1992; Marslen Wilson 1994). Undoubtedly therefore a preference for category-split models, over string-scan models appears favourable.

Forster's model can be computationally viewed as a hash-table representation of the lexical database. Hash codes generated on the basis of the first few letters would lead on to the main lexicon. Forster's prediction however, about the transition network model that predicts a word before its realization, appears to go wrong in respect of the reaction time the concerned word, fails to evoke.

That, words are ordered by frequency-rank in the mental lexicon is almost instinctive. Nevertheless, designing a computational model might not be plausible, since such a concept cannot be wholly accommodative as it were. On these lines, the autonomous search model too does not appear to be popular.
3.4 A Fitting Model For Computational Purposes

The questions posed by Sproat (1992) are:

- Does the model of lexical retrieval used in the system resemble a good deal of what humans actually do?
- A majority of computational models of considerable interest involve greater or lesser amounts of morphological decomposition; are the algorithms and representations employed reasonable models of the way humans do things?

Psychological reality in morphological analysis is a prerequisite to a system that believes in natural language parsing in an intuitive way. Accessing a word from the mental lexicon can be either pre-lexical (indirect) or post-lexical (direct). A pre-lexical access would involve processing of the input up to the point of entry into the lexicon. In other words, words are split up into their constituent morphemes before they are identified or retrieved from the mental lexicon. Post-lexical processing occurs after a lexical entry has been accessed. Its phonological and graphological form specifications become simultaneously available for analysis.

Although studies on the Taft & Forster model had been viewed as unsuccessful by Emmorey (1998), experiments have indeed revealed that real stems and real affixes are easily rejected as non-words, than pseudo-stems and pseudo-affixes. Whatever be the inferences, the claim that inflectional class vitally determines the process of morphological decomposition, is what is of consequence here. McQueen & Cutler (1998) opine: "Access to morphologically structured lexical representations, ... need not involve morphological structure overtly, decomposition of derived forms may be an optional procedure, available when the whole-word access procedure fails."

Experiments have established that for English, analysis of words began with stripping of items that resembled prefixes, proceeded to stems and then to suffixes. Relative access time for an entry is a controversial issue in such models.

Any computational representation of morphology needs to ingrain certain aspects of psychological evidence to explicate the nature of natural language processing in the human mind. Bearing in mind the pros and cons of the models considered, it can be reasoned out that any category-split model that accounts for a reasonable level of morphological analysis that bases itself either on the IA, IP, or on a model that features a blend of both, could be regarded as feasible for application.

AMPLE as a case in point, would characteristically involve a listing of stems and affixes, along with constraints that define which affix could combine with which stem. Those, in the case of the KIMMO model would be used as 'continuation patterns'. This would presuppose storage of the basic morphological forms, alongside the essential rules of phonology and morphophonology. The analyzer developed in PERL, chooses, for reasons of simplicity, a model that ingrain effects of both the IA and the IP model, with a parsing mechanism based on morphological paradigms.