Usage based models of language
Michael Barlow
Suzanne Kemmer

A Dynamic Usage-Based Model
Ronald W Langacker

Introduction: A Usage-Based Conception of Language

There are two usage based traditions: Firthian (emphasising the importance of context including social aspects) and Enuciativist linguistics (theories based on the speech act) but recently language has been treated as a more or less fixed system independent of context and cognition.

Usage based models share a number of characteristic assumptions:

The intimate relationship between linguistic structures and instances of use of language.

The importance of frequency. Frequency of a usage pattern is both a result and a shaping force of the system.

Comprehension and production as integral, rather than peiferal, to the linguistic system.

Focus on the role of learning and experience in language tradition.

Linguistic representations as emergent, rather than stored as fixed entities.

Importance of usage data in theory construction and description.

The intimate relation between usage, synchronic variation, and diachronic change.

The interconnectedness of the linguistic system with non-linguistic cognitive systems.

The crutial role of context in the operation of the linguistic system.

1. The Usage Based Conception

Generative theory trieds to minimise what a speaker has to learn and mentally represent. Cognitive Grammar that becoming a fluent speaker involves a prodigious amount of actual learning and tries to minimise the postulation of innate structures specific to language.

Rules and lists are not mutually exclusive (the rule/list falacy). An alternative is to include in grammar both rules and instanciating expressions.

Generative aGrammar is top-down. There are several respects in which Cognive Grammar is bottom up

2. Psycological Grammar

A number of essential basic and very general psychological phenomenoa are described

entrenchment is the occurance of psycological events that leave some kind of a trace that facilitates their reoccurence. A complex event attains *unit* status when it is sufficiently entrnched so as not to require conscious attention to its parts.

An *abstraction* is the emergence of a structure through reinforcement of the cominality inherrent in multiple experiences. *Schematization* is a special case involving the capacity to operate at varying levels of granularity. (or resolution). A *Schema* is the cominality the emerges from distinct structures when one abstracts away from their points of difference.

The ability to compare two structures and detect any discrepancy between them. One structure is the target and the other the standard. Categorisation is a special case. Where an act of categorisation registers some discrepancy then it is referred to as an extension.

Composition is the combination of simpler component structures to yeild a composite structure.

Association is when one well known experience evokes another. A special case is symbolisation, the association of conceptulisations with mental representisations of observable entities (sound, smell, guestures etc)

3. Processing interpretation

As a general orientation, the Author inclines to the

Connectionist style of computation.

Entrenchment is interpretted as an adjustment in connection weights.

A schema is immanent in their instantiations, rather than being seperately stored or accessible in general.

Categorisation is interpreted as capture by an attractor (in a topographical conception of state space, where attractors of different strength are thought of as wells and valleys of varying depth that a system settles into as it relaxes into a state of minimum energy). An experience is a point in state space or a trajectory through it.

Composition is plausibly then the activation of a number of components - ultimately reducing to an act of cetegorisation.

Symbolism is sreadily interpretableas one pattern of activation reliably serving to elicit another.

4. Basic Linguistic Problems.

"In Cognitive Grammar, a **language is described as a structured inventory of conventional linguistic units.** The units (cognitive "rou-tines") comprising a speaker's linguistic knowledge are limited to semantic, phonological, and symbolic structures which are either directly manifested as parts of actual expressions, or else emerge from such structures by the proc-esses of abstraction (schematization) and categorization (this restriction is **called the content requirement).** In describing these units as an inven-tory, I am indicating the non-generative and non-constructive nature of a linguistic system. Linguistic knowledge is not conceived or modeled as an algorithmic device enumerating a well-defined set of formal objects, but simply as an extensive collection of semantic, phonological, and symbolic resources which can be brought to bear in language processing. This inven-tory of resources is structured in the sense that some units are incorporated as parts of larger units, specifically in relationships of categorization, com-position, and symbolization" pg.8

There are several factors which determine which particular member of the activation set emerges as the active structure evoked to categorise the target:
- The level of entrenchment
- Contectural priming (which overides entrnechment)
- The amount of overlap between the target and a potential categorising structure.

The author discusses how language components are recognised, and constructed using this model. An example was the conversion of a verb to a noun by adding er. Print is converted to printer but this has captured its own meaning as an electronic computer device and for example excluded newsprint machines.

There was an interesting concept in that the structure is bottom up rather than top down. The act of categorising a target is first captured by specific schemas (ree apples) before being captured by more general schemas (apples, fruit or food)

5 Structural Applications.

"The usage-based model described above is applicable to all domains of lan-guage structure: semantics, phonology, lexicon, morphology, syntax. A linguistic system comprises large numbers of conventional units in each domain, and a target expression is simultaneously categorized by numerous active units, each assessing a particular facet of its structure. A few basic psychological phenomena (listed in Section 2), applying repeatedly in all domains and at many levels of organization, give rise to structures of indefi-nite complexity, which categorizing relationships-each pertaining to a par-ticular structural dimension-link into cross-cutting networks. A description of this sort is further unified in that seemingly diverse phenomena are seen as residing in different aspects of the same or comparable structural assem-blies, or the same aspects "viewed" in alternate ways." Pg 32

This sections discusses the models applicability to sentence level constructions (as oposed to word level in the previous section.

There was a reminder of how quickly the brain converts a novel action requiring conscious effort into a automated skill.

Metaphor was briefly explored as "a vast subject".

"I therefore propose a synthesis along the following lines. In principle, the description of a metaphor always involves at least three structures: the source, the target, and a schema. The latter two structures do however show substantial variation in their nature and status, thus contributing to the im-pressive qualitative diversity of metaphors. The target domain ranges along the entire spectrum from being fully apprehended independently of the meta-phor, as when people are likened to animals, to being wholly constituted by the metaphor for all intents and purposes (e.g. a layperson's understanding of the Big Bang). The schematic commonality of the source and target do-mains likewise ranges from the blatantly obvious (people and animals have a lot in

common), through the tenuous and hard to articulate (e.g. an argu-ment as a building), to the limiting case where any commonality is negligi-ble. This would at least be so for instances where the target domain is wholly constituted by the metaphor.

Yet even three structures are not enough…………"
Pg.41

The Phonology of the Lexicon: Evidence From Lexical Diffusion.
Joan L. Bybee

"Two **important points for usage-based models are established** by the data presented here. It has been argued that sound change, or phonetic change, affects the lexicon gradually by spreading out across lexical items from the most frequent to the least frequent, and by producing small gradual changes in the phonetic representation of lexical items. One important consequence of this view is that sound change permanently affects the lexicon, and that lexical items are not represented in an abstract phonemic notation. A second consequence is the more general point that the phonetic properties of lexical items, like their morphosyntactic and semantic properties, are affected by language use." Pg.81

Bidirectional Processing in Language and Related Cognitive Systems

Just because toothpaste comes out of a tube in a single long cylinder does not mean that it is stored as a cylinder. The form that language takes is not necessarily indicative of the form that it has in the brain.

Three distinctions:
1. Linguistic Extension. The set of all texts, spoken or writen.
2. Linguistic Processes. The process by which linguistic extensions are produced and/or interpretted.
3. Linguistic System. A system capable of producing and/or receiving texts.

Q1. Why develop Usage Based Models?
A1. Language is ambiguous and really can only refer to the lingistic definitions above and Usage-based is similarly ambiguous.

Q1. (rephrased) Why develop a model which is compatible with what we know about linguistic processes?)

There have been various theories of Linguistic extension:

- The building-block approach
- Symbol and rules
- The approach that avoids formalism and just lists or describes patterns
- The network approach
- Mixed systems.

Q2.What about the "Gods Truth" approach? Might there be some reality behind the linguistic

extension?

We can construct a theory of the linguistic extention of an individual or a community
We can construct a theory of a linguistic system of an individual.

A2. No.

Q3. How can we investigate the Linguistic System?

A3. by studying the Linguistic Extension and trying to figure out what kind of a system is needed to produce and understand such material and to learn how to perform these processes.

Q4. Is a theory of the linguistic extension also a theory of the linguistic system?

A4. No! (various examples)

A3. Additional .. We must look beyond the Primary Data (linguistic extension) to Linguistic processes (competence to perform) and the biological basis of language.

A1. A cognitive model of language must necessarily be a usage-based model; it must represent a competence to perform.

Of the 5 approaches to modeling linguistic extention listed previously only the network model has any validity.

Q5. Is there a reasonable biological basis for a network model?
A5. Yes, but……

Many neural networks and the like have fundamental differences to the brains structure. They often have only one hidden layer while the simplest marine snail has two! Every neuron in layer one is connected to very neuron in layer two while in the brain a fraction of 1% is forward connected. The connection weights start random while people clearly do not start with random knowledge.   A applications of artificial networks are also very limited (ie adding ed to a very to make it past tense while a brain receives both as inputs)

The Author proposed the "Relational Network Model" from a language basis in 1971 and has refined from their. It is a functional network, not a neuron level network and any resemblance with the brain is a happy consequence.

Q6. Is a Realistic account of processing in a network possible?

Some of the brain is clearly a network.

The "Relational Network Model" (above) helps account for a number of real world linguistic phenomenon such as "slips of the tongue", fruedian slips, unintended puns, fruedian forgetting etc.

Also, the network model allows all linguistic information to be present at once, while also be organised in sub-systems.

A6. Yes.

Q7. Can we account for how al that structure gets learned?

DNA provides a basic structure with an overabundance of latent connections, from which certain ones are strengthened for the leaning process.

Evidence from language development studies suggests that development of the network is a largely a bottom up process.

Consider an example network where A, B & C are beginning to strengthen in their connections to K. K represents the composite of the three features A,B, & C. "The essential process hypothesized is that any connection will get strengthened (including a dedicated connection, which can get further strengthened) if it is activated at about the same time as the node to which it is connected receives enough activation to have its threshold satisfied; and that threshold is thereupon raised.

Q8. How can we account for bidirectional processing?

The signals from neurons and NECTIONS (of the Relational Network Model) only feed forward, not backward.

Q9. Is bidirectional processing unique to language?

A9. No.

Bidirectional processing is also apparent in visulisation and motor control.

Which direction is primary? Activation in the primary direction occurs more frequently, visulisations in the primary are more vivid and activation in the secondary direction can only occur after establishing the necessary connections in the forward direction.

Q10. Is bidirectional processing asymmetrical in language, as it is in perception?

A10. There is a degree of primacy to the reception side …

In language learning reception precedes production

People have larger vocabularies for understanding than production

The babbling phase of language acquisition correlates production to resulting sounds

When a person recognises a sentence as being ambiguous and starts over.

We usually get from a word to the meaning faster than from the meaning to the word.

Q11. How does bidirectional processing work?

Backwards activation might be achieved by not only sending activation from K forward to the next level but also back to all nodes in the previous level (including A,B&C). This downwards activation needs to be just strong enough to reactivate just those nodes of the next lower level which are currently active, while not activating the others. The diminishing strength of activation in this reinforcement cycle will cause it to become inactive after a few iterations leaving a strengthened connection and an elevated threshold.

Note that this model allows for a minor feature to be recognised (a pink nose) which then feeds back to activate the remainder of the pattern (a cat).

Q12. Is there a reasonable biological basis for this proposed answer to Q11?

This nection model is probably more like a group of neurons (a cortical column??) than single neurons.

A12. Yes!

Connectionism and Language Learning
Brian MacWhinney.

Common sense tells us that linguistic behavior is grounded in a ridged set of social conventions or rules. In the 60's and 70's this was reworked into a basis for AI by Chomskyan theoretical linguists and cog psy. But this approach became increasingly shaky as no child learns rules and there was no evidence of rules in the brain. The models developed by Rumelhart and McClelland in the 80 corrected the fundamental flaw inherent to symbolic models: the problem of excessive descriptive power. They were too strong in that they allowed the modeling of the learning of things that could never in reality be learnt. They were too weak in that they were not sufficiently general to cover language types and patterns.

Neural Nets impose two stringent limitations: a prohibition against **symbol passing** and an insistence on **self organisation**. A third constraint of **scalability** makes things very difficult. Performance degradations is particularly severe for language acquisition models.

The connectionist approach has grand pretensions but a modest reality. There are very successful models but in very limited domains.

When the models work simply with phonology then there are lexical issues such as
- Homophony – the maid wrung out the clothes, the soldiers ringed the city, or the choirboy *rang the bell*
- Compounds – irregular verbs such as *undergo, rethink* and *undo* with past tense *underwent*, rethought and *redo*.
- Derivational Status – we say *the groom ringed her finger* but never *the groom wrung her finger*
- Early irregulars – children often begin with *goed* and progressively transition to *went*.

These four problems can be linked to an absence of a way of representing lexical items. One way of addressing this problem is to mix sound features and meaning features without providing any explicit representation of lexical items. Models have been developed which work but break down with more than about 700 forms. However the author has overcome these problems by explicitly include lexical representation – lexical connectionism.

There is a discussion of various models pg 137-144 which is difficult to summarise but worthy of a re-read as it has good references to the detailed papers.

There are some interesting comments on the way to manage short term memory for the processing of sentences. Grouping into conceptual clusters, holding the sentence in a phonological loop, or operate at a lexical items.

The Effect of an Interlocutor on Episodic Recall: An Experimental Study.
Connie Dickenson and T. Givon.

"Goals of the Research Project.

Our current study is but an early foray into a series of related areas, characterized initially by the question: Are the interactional and informational aspects of an ongoing communication-whether monologuic or dialoguic-processed and stored separately in episodic memory? Or conversely, is there an integrated system that is responsible for both aspects of communication?

That the question may not have a simple, clean answer is quite likely. Evidence from conversational analysis (cf. Goodwin 1982, 1988, 1984; interalia) suggests that many informational aspects, particularly relevance, topicality, or thematicity, are negotiated during communication. But the study of conversation also reveals that conversations can be either highly cooperative or extremely non-cooperative, and that non-cooperative conversations in a way resemble two parallel monologues. In cognitive terms, this may suggest that the mental representation of the communicative transaction in episodic memory could be of two extreme types. A cooperative interlocutor may create an integrated mental representation of the interactional and informational aspects of the communication. A non-cooperative interlocutor, on the other hand, may either ignore interactional clues, or may represent them separately, without integration into a unified episodic representation.

The present study is an attempt to gradually probe the effect of the interlocutor's contribution to the interaction on the speaker's episodic model of both the informational and the interactional aspects of communication. In the first five experiments, reported here, an information baseline is set up (Control Conditions I, II, and III), then the effects of two types of intervention by an interlocutor are assessed (Conditions IV and V). Two more conditions that aim at assessing the degree of integration of interactional information with episodic information are in the process of being investigated. The results we report here should be thus considered the first step in a complex, long-term study.

Methodology

Overall Design

Five experimental conditions were created in the phase of the study reported here. The first three (I, II, III) serve as controls for all subsequent conditions,

creating a base-level of episodic recall to which all subsequent conditions are then compared. The last two conditions (IV, V) assess the effect of two types of face-to-face interaction, taking place between episodic input and its recall, on subsequent episodic recall. In all conditions, subjects viewed the same short video film. The viewing was followed by various interventions, following which the subjects were asked to verbally recall the events of the film. Their verbal recollections, as well as the verbal interaction preceding it in Conditions IV and V, were recorded and transcribed. This section deals with the methodology common to all five experimental conditions. Procedures specific to the individual conditions will be discussed in subsequent sections.

Text Elicitation

A short, 6.25 minute-long video was used. This video had been developed previously for use in another experimental project (Givon 1991). The following is a brief summary of the story:

Synopsis of the Chicken Story

A man walks toward a tree, leans his farming implements on it and goes on to chop wood with an axe. A woman appears and walks to him. After some conversation, she takes the wood, moves aside and collects some more wood, then carries it all away. The man quits his chopping, collects his tools and walks off toward a grove. The scene shifts to the woman coming around a small shed. She unloads her wood, lights a fire, fetches water from a barrel and sets a pot of water to boil. She disappears behind the shed and comes back carrying a chicken. She tries to slaughter it with a knife, but clumsily bungles the job and the chicken escapes. After some perfunctory chasing, the woman goes back to the house, brings out some bread and cheese, makes a sandwich, wraps it up and leaves with it. The scene shifts to the man hoeing in the field. The woman arrives and offers him the package. They sit down, the man unwraps the food, rejects it, throws it back at the woman, then chases her around the tree with his hoe. (Givon 1991: 143-144)"

......

"Tentative Conclusions.

We have shown that interposing conversational interaction between the presentation of episodic information and its subsequent recall has two possible effects on the recalled information. When the interlocutor is uncooperative and introduces misleading, contradictory information (Condition IV), the amount of episodic information recalled later goes down. On the other hand, when the verbal interaction is cooperative, involving a greater number of coherent questions and responses (Condition V), the amount of episodic information recalled later goes up. This effect goes beyond the effect of mere prior verbalization (I), and probably reflects the salutary effect of the need to reorganize the information more coherently during cooperative interaction. The quality of the intervening verbal interaction (Condition V) during the consolidation of episodic information, rather than the mere quantity of rehearsed verbalization (Conditions I and IV), appears to make the real difference."

The Development of Person Agreement Markers: From Pronouns to Higher Accessability Markers Mira Ariel.

"2.2. A Superfluous NP Subject Appears: Accessibility and Inflection

In effect in line with Sanford and Garrod (1981), Givon (1983) and Chafe (1987), I have argued (Ariel 1985, 1988, 1990, 1991, 1996) that when reference to mental entities is made, the speaker chooses her referring expression according to how she assesses the accessibility of the specific entity for her addressee at the current stage of the discourse. The higher the mental accessibility, the higher the accessibility marker chosen. Accessibility markers are specialized for a variety of degrees of accessibility, and non-arbitrarily so. Lower accessibility markers are more informative (calculated according to amount of lexical material), more rigid (i.e., identify a mental entity relatively uniquely) and less attenuated (lengthier or accented, regardless of informativeness). Higher accessibility markers are on the whole less informative, less rigid and more attenuated (regardless of informativeness). Such a form-function correlation seems only reasonable, given that mental entities currently not highly accessible are better retrieved by supplying more information, for the addressee has to choose among very many mental entities he stores in his long-term memory. Highly accessible entities, on the other hand, cannot be numerous, because we can only keep a very limited number of mental entities highly accessible at any given moment.

However, there are indications that accessible information is deaccentuated, even pronounced less intelligibly as an unambiguous signal that the information concerned is Given rather than New.' In other words, it is not only that the speaker invests more energy in coding New information, because it is harder to process. She also intentionally aims at reducing the form expressing accessible information, since it is a useful linguistic cue for the addressee to search for an already available entity. Now, the form function correlation between accessibility markers and degree of mental accessibility is not perfectly transparent, but the following scale of accessibility markers is quite representative of many languages (though not all languages have all the options listed, see Ariel (1990) for discussion of accessibility and universality). The differences in the degrees of accessibility coded by (any) full NPs, free pronouns, bound pronouns, agreement and true zeroes prove to be crucial in the development of verbal person agreement markers:

(4) The accessibility marking scale

zero < reflexives < poor agreement markers < rich agreement markers < reduced/cliticized pronouns < unstressed pronouns < stressed pronouns < stressed pronouns + gesture < proximal demonstrative (+NP) < distal demonstrative (+NP) < proximal demonstrative (+NP) + modifier < distal demonstrative (+NP) + modifier < first name < last name < short definite description < long definite description < full name< full name+ modifier.

I have focused on two types of considerations taken into account by the speaker when assessing the degree of accessibility associated with specific mental entities in the addressee's memory (although other factors are no doubt involved): entity salience and unity. The former criterion refers to the degree of salience of the potential antecedent (linguistic or non-linguistic); the latter refers to the strength of the connection between the referring expression/accessibility marker and the potential antecedent. The Unity criterion is mainly relevant for anaphoric references, and pertains to the distance and the degree of cohesion between the units (e.g., clauses) containing the two expressions. It is the first criterion, antecedent salience, which will ultimately prove crucial for the formation of agreement inflections. All things being equal, the entities mentioned on the left in 5 are more salient than the ones on their right:

(5) Antecedent salience

a) Speaker > addressee > non-participant (third person)
b) High physical salience > low physical salience
c) Topic > nontopic
d) Grammatical subject > nonsubject
e) Human > animate > inanimate
f) Repeated references > few previous references > first mention
g) No intervening/competing referents > many intervening/ competing referents

"

......

"Note that for Accessibility Theory, it is not really inter-person frequency comparisons that matter. It is the frequency of high accessibility coding per person which is crucial. In other words, we expect many of the first/second person referents, but few of the third person referents, to be extremely accessible, regardless of their absolute frequencies in discourse."

Further Reading

Deane, paul. 1992. Grammar in Mind and Brain: Explorations in Cognitive Syntax. (Cognitive Linguistics Research 2.) Berlin: Mouton de Gruyter.

David E. Rumelhart and McClelland, James. L. (eds.) 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol 1: Foundations. Cambridge, MA/London: MIT Press/Bradford.

McClelland, James. L. and David E. Rumelhart (eds.) 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol 2: Psycological and Biological Models. Cambridge, MA/London: MIT Press/Bradford.

Copeland, James E and Philip W. Davis (eds) 1980. Papers in Cognitive Stratification Linguistics. Rice University Studies 66(2)

Burnod, Yves. 1990. An adaptive Neural Network: The cerebral Cortex. Paris:Masson; London: Prentice Hill

Lamb, S.M. (1999). *PATHWAYS OF THE BRAIN: The Neurocognitive Basis*

of Language Amsterdam & Philadelphia: John Benjamins Publishing Co.,

Worth reading some of the detail for various models noted in pages 137-144