Analogy in Grammar: Form and Acquisition

Department of Linguistics
Max Planck Institute for Evolutionary Anthropology
Leipzig, 22–23 September 2006

Workshop programme

Friday, September 22

Welcome and Introduction

9:00–9:30 Jim Blevins (Cambridge) & Juliette Blevins (MPI-EVA)

Analogy in Language Change and Typological Variation

9:30–10:15 Andrew Garrett (Berkeley)
Paradigmatic heterogeneity

10:15–10:30 COFFEE BREAK

10:30–11:15 Andrew Wedel (Arizona)
Multi-level selection and the tension between phonological and morphological regularity

11:15–12:00 Greg Stump & Rafael Finkel (Kentucky)
Principal parts and degrees of paradigmatic transparency

12:00–12:30 Moderated discussion: Juliette Blevins (MPI-EVA)

12:30–14:00 LUNCH

Analogy in Synchronic Grammar

14:00–14:45 Keith Johnson (Berkeley)
Analogy as exemplar resonance: Extension of a view of sensory memory to higher linguistic categories

14:45–15:30 John Goldsmith (Chicago)
Learning morphological patterns in language

15:30–15:45 COFFEE BREAK

15:45–16:30 Susanne Gahl (Chicago)
The sound of syntax: Probabilities and structure in pronunciation variation

16:30–17:15 Farrell Ackerman (UCSD), Sharon Rose (UCSD) & Rob Malouf (SDSU)
Patterns of relatedness in complex morphological systems

17:15–18:00 Moderated discussion: Colin Bannard (MPI-EVA)

19:00– DINNER AT BAYERISCHER HOF

Saturday, September 23

Analogy in Language Acquisition

9:00–9:45	LouAnn Gerken (Arizona) *Linguistic generalization by human infants*
9:45–10:30	Andrea Krott (Birmingham) *Banana shoes and bear tables: Children's processing and interpretation of noun-noun compounds*
10:30–10:45	COFFEE BREAK
10:45–11:30	Mike Tomasello (MPI-EVA) *Analogy in the acquisition of constructions*
11:30–12:15	Rens Bod (St. Andrews) *Acquisition of syntax by analogy: Computation of new utterances out of previous utterances*
12:15–12:45	Moderated discussion: Elena Lieven (MPI-EVA)
12:45–14:00	LUNCH

Analogical Research and Modeling

14:00–14:45	Royal Skousen (Brigham Young) *Expanding analogical modeling into a general theory of language prediction*
14:45–15:30	Dedre Genter (Northwestern) *Analogical processes in learning grammar*
15:30–15:45	COFFEE BREAK
15:45–16:30	Adam Albright (MIT) *Modeling analogy as probabilistic grammar*
16:30–17:15	Harald Baayen (Radboud University & MPI Nijmegen) & Fermín Moscoso del Prado Martín (MRC-CBU Cambridge) *Bits and pieces of an information theoretical approach to inflectional paradigms*
17:15–18:00	Moderated discussion: Jim Blevins (Cambridge)
19:30–	CONFERENCE DINNER

Analogy in Grammar: Form and Acquisition

Department of Linguistics
Max Planck Institute for Evolutionary Anthropology
Leipzig, 22–23 September 2006

Workshop Abstracts

Farrell Ackerman (UCSD), **Sharon Rose** (UCSD) & **Rob Malouf** (SDSU)

Patterns of relatedness and based learning in complex morphological systems

FRIDAY 16:30–17:15 (ANALOGY IN SYNCHRONIC GRAMMAR)

There is a fundamental problem confronting native speakers of highly inflected languages with numerous declension classes for nouns or conjugation classes for verbs, namely, paradigm completion. Paradigm completion concerns the licensing of reliable inferences about the surface wordforms for the inflectional (and derivational) families of wordforms associated with (classes of) lexemes, i.e., given a novel inflected word form, what are all the other wordforms in its inflectional (and derivational) families? Thus, given an inflected form of a new lexeme what permits correct inferences to the rest of the inflectional forms for this lexeme? When phonological relations are transparent, allomorphy is limited, and the morphosyntactic feature combinations are few the problems of mapping meanings to surface forms are often trivial. So, in English a child's exposure to a singular form of a new lexeme aardvark permits paradigm completion for number to infer aardvarks. Many extraordinary claims have been developed about the structure of the human language faculty and the nature of language learning with respect to morphology on the basis of simple systems such as English and German (Pinker 1999, Clahsen & Temple 2004). However, languages ordinarily depart from such simple mappings, some quite dramatically. So, several fundamental questions arise: (i) how are such complex systems organized, (ii) what role does this organization play with respect to licensing inferences concerning paradigm completion, and (iii) what relation does this organization and the possibility for inferences have concerning the learnability of such complex systems? The examination and modeling of complex synchronic morphological systems offers insights not only into the patterned organization of morphology but also provides an empirical basis for better grounded speculation on learnability. After all, it makes little sense to speculate about the nature of some language module or of how it may be learned in default of a solid grasp of the target.

With these larger issues in mind, we will identify the patterned nature of paradigm organization for (subparadigms of) the Tundra Nenets (Samoyed branch of Uralic) nominal declension system. In the inflectional paradigms of this language complex (pieces of) wordforms recur in cells: there are more morphosyntactic feature combinations than there are distinct forms to encode them, i,e., there is considerable homonymy/syncretism. As a consequence, following Trosterud (2004:54): "wordforms are signs, parts-of-wordforms are not." That is, surface wordforms are construable as recombinant gestalts, not simple (or even complex) combinations of morphemes: surface words, stems, and affixes (or more generally, operations) are best viewed in their paradigmatic contexts of expression. This essentially follows the Saussurean perspective on paradigmatic relations: words, stems, and affixes (or

operations) are best understood in terms of the networks of related forms in which they play a role, rather than in terms of single meanings or function ascribed to them as in morpheme-based proposals. Learning, accordingly, is based on assimilating new forms within previously encountered patterns, partly, on the basis of analogy.
In this exploratory exercise we will adopt a word-based perspective on morphological analysis (Matthews 1972/1991, Bochner 1993, Blevins 2005, to appear, Kirby 2005, Stump & Finkel 2006, Bonami and Boyer 2006, among others). In line with Stump & Finkel's taxonomic approach to pattern detection, we establish the basic organization of the Tundra Nenets and identify their status with respect to implicational relations. We will explore two approaches to paradigms organization, namely, a mutual implication local alliance hypothesis (Bochner 1993) versus a single surface base hypothesis (Albright 2002). Given the implausibility of assuming that speakers store entire whole paradigms in morphologically complex languages, we examine a way in which the postulation of fully specified paradigm representations can be avoided while still accounting fully for synchronic representation as well as for their learnability. For this, time permitting, we focus on an instructive and representative subset of the our data and explore a connectionist model of it following the work of Thyme (1993), Thyme, Ackerman & Elman (1994) (both based on the insights of Paunonen 1976 and Skousen 1975, 1989 for similar speculations) as well as recent work by Goldsmith and O'brien (2006).

Adam Albright (MIT)

Modeling analogy as probabilistic grammar

SATURDAY 15:45–16:30 (ANALOGICAL RESEARCH & MODELING)

Formal models of analogy face two opposing challenges: on the one hand, they must be non-deterministic enough to handle gradient, variable data and to accommodate a wide range of statistical tendencies. At the same time, they must account for the fact that analogy obeys some striking restrictions: only a tiny subset of the logically possible analogical inferences are actually attested in language change, child errors, etc. In this talk, I discuss several such restrictions, using data from acquisition, psycholinguistic experiments, and historical change, and consider their implications for formal models of analogy. I argue that both the probabilistic nature of analogy and also the restrictions on possible inferences can be modeled using a grammar of stochastic rules, which can encode fine-grained knowledge of statistical tendencies, but at the same time, is constrained by the formal limitations of the grammatical system.

The first restriction concerns what type of pattern can be analogically extended. I argue that analogical inference must be supported by words with a particular element in common: containing a certain type of segment, belonging to a certain inflectional class, etc. In order to capture this, a model must use structured representations, allowing patterns to be expressed as relations over corresponding elements. This requirement is often taken for granted by the use of variables in, e.g., proportional notation (X:Y :: A:B), Bochner's lexical relations (X -> XY), or standard rules (A -> B/C_D). Similarity-based exemplar models are more powerful, however, and can capture analogies supported by sets of words with no single common property. Using experimental and historical data from Spanish, I show that a popular exemplar model (Nosofsky 1986) does not perform as well as a rule-based model, because it predicts unattested analogical inferences, based on support from diverse types of words.
The second restriction concerns what it means for a pattern to be well supported. In principle,

patterns could gain strength through either high type frequency or high token frequency. Based on Spanish data, I argue (following Bybee 1995 and Pierrehumbert et al. 2001), that strength of morphophonological patterns is based solely on type frequency. Although this restriction could be built in to most formal models of analogy, I argue that it follows most naturally from a model that abstracts rules from a lexicon of words, rather than a stockpile of exemplars.

The last restriction concerns directional asymmetries. It has long been observed that analogy characteristically affects some forms in the paradigm but not others (basic -> derived), and in fact, these asymmetries are seen even more strongly in child errors. Based on acquisition data from Spanish, German, and Korean, I show that only a small fraction of statistically well-supported patterns actually lead to analogical errors. This directionality is correctly predicted by a model of grammar in which some forms are derived from (or depend on) other forms in the paradigm.

Harald Baayen (Radboud University & MPI Nijmegen)

Fermín Moscoso del Prado Martín (MRC-CBU Cambridge)

Bits and pieces of an information theoretical approach to inflectional paradigms

SATURDAY 16:30–17:15 (ANALOGICAL RESEARCH AND MODELING)

Blevins (2003) argues for a word and paradigm approach to inflectional morphology. In word and paradigm morphology, generalization is based on analogical inference across inflectional exemplars. We have been using information theory to come to a better understanding of analogical inference in morphology.

In our presentation, we review a series of recent experimental studies that demonstrate the importance of paradigmatic structure for lexical processing in comprehension and production. We demonstrate that both word-specific and paradigm-specific probability distributions jointly drive analogical generalization.

We introduce a set of information-theoretical measures that gauge, across a variety of experimental tasks, different aspects of paradigmatic generalization. These measures also allow us to understand how token-based experience results in what seems to be type-based generalization. We conclude with a first sketch of a model that grounds these measures in the neuro-biology of the brain.

Considered jointly, our findings bear witness to the pervasive and extensive sensitivity of the brain to the combinatorial probabilities of inflectional variants within their paradigms.

References

Baayen, R. H., Feldman, L. & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language* 55, 290-313.

Blevins, J. P. (2003). Stems and paradigms, *Language* 79, 737–767.

Kostić, A., Marković, T. & Baucal, A. (2003). Inflectional morphology and word meaning: orthogonal or co-implicative domains? In Baayen & R. H., Schreuder, R. (Eds.), *Morphological structure in language processing*. Mouton de Gruyter, Berlin, pp. 1–44.

Moscoso del Prado Martín, F., Kostić, A., Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94, 1–18.

Rens Bod (University of St. Andrews)

Acquisition of syntax by analogy: Computation of new utterances out of previous utterances

SATURDAY 11:30–12:15 (ANALOGY IN LANGUAGE ACQUISITION)

One of the most important challenges for analogical models of language consists in describing the processes that deal with the acquisition of syntactic structure. That is, how can we learn to produce and comprehend an unlimited number of new utterances by constructing analogies with previously perceived utterances? In this talk we will treat this question from a computational-linguistic perspective.

In previous work (Scha 1990, Bod 1998), we proposed that human language production and comprehension works with concrete past language experiences rather than with abstract linguistic rules. We developed a model, known as Data-Oriented Parsing (DOP), which uses a corpus of previously perceived utterances with their appropriate analyses as a representation of a person's past language experience. New utterances are produced and analyzed by combining fragments from analyses that occur in the corpus. These fragments can be of arbitrary size or shape, ranging from single words to entire utterance-representations (e.g. semantically enriched tree structures), thus allowing for constructions and prefabs of any complexity. By taking into account the occurrence-frequencies of the fragments it can be decided which is the most probable utterance for an intended meaning (or the most probable meaning for a perceived utterance). By allowing for both lexicalized and unlexicalized fragments, DOP captures the effects of both token and type frequency. Recency effects are incorporated by a monotonously decreasing frequency-adjustment function. DOP thus proposes a probabilistic account of syntactic analogy which maximizes the similarity between a new utterance and utterance-representations in the corpus. This probabilistic notion of analogy correlates with the number of corpus representations that share fragments with the sentence, and also with the size and recency of these shared fragments.

A major question for the DOP approach is where the initial representations come from. There is an increasing body of evidence that linguistic structure is learned entirely in a statistical, item-based way. The key idea in statistical models of language acquisition is that word sequences surrounded by equal or similar contexts are likely to form a certain constituent. The probability that a substring of a sentence constitutes a certain constituent is computed from the substring's frequency and the frequencies of its contexts (Van Zaanen 2001, Clark 2002, Klein and Manning 2005). Unfortunately, all known statistical learning models a priori restrict the lexical relations that are taken into account in learning syntax. For example, the model by Klein and Manning (2005) only uses statistics of contiguous substrings while it is well known that many lexical dependencies are non-contiguous or structural (i.e. they can be separated by other words or constituents). What would be needed is a model that (initially) takes into account all contiguous as well as non-contiguous substrings of sentences in learning syntax.

To this end, we proposed in Bod (2006) a model which assigns all possible analyses (i.e. all tree structures) to initial sentences and next uses all fragments (i.e. subtrees) from these analyses to compute the most probable analysis-trees for new sentences (there are efficient algorithms to do this). The learned analyses are added to the corpus and can in turn be used to analyze and produce new sentences. The underlying idea of this new DOP model is that if we do not know which analyses should be assigned to first utterances, we may just as well assume that initially all analyses are possible and let the statistics decide which analyses -- and fragments thereof – are most useful to understand and produce new utterances. Such an

approach to language acquisition is just another application of the general DOP idea: rather than operating with already learned (sub)trees, we start out to operate with arbitrarily assigned (sub)trees which are next selected on their usefulness in processing new sentences. This DOP model is congenial to the item-based or usage-based approach to language acquisition (e.g. Tomasello 2003, Bybee 2006), but extends it in an important way by providing a formal model that computes fresh utterances out of previous utterances, maximizing the analogy between new and old language experiences. In my talk I will go into some experiments with this DOP model using corpora of English, German and Chinese (Mandarin) language data.

Susanne Gahl (University of Chicago)

The sound of syntax: Probabilities and structure in pronunciation variation

FRIDAY 15:45–16:30 (ANALOGY IN SYNCHRONIC GRAMMAR)

There is a remarkable degree of consensus on one claim in the study of language: Sentences have a hierarchical internal structure, and they cannot be adequately described in terms of Markov chains of words, i.e. transitions from one word to the next. Yet, Markov models ("word-to-word" grammar models) continue to be the dominant model in speech recognition, and consequently in research on pronunciation variation. Unfortunately, the predominance of word-to-word models in speech research contributes to a lingering misperception that probabilistic linguistics is incompatible with sophisticated grammar models. This chapter lays out research showing that even in pronunciation variation, structure permeates language. Pronunciation reflects probabilities of syntactic structure, not just of word-to-word transitions. Evidence for this claim comes from phenomena that have been central to pronunciation variation: the duration of words and pauses, and segment deletion.

Andrew Garrett (University of California, Berkeley)

Paradigmatic heterogeneity

FRIDAY 9:30–10:15 (ANALOGY IN LANGUAGE CHANGE AND TYPOLOGICAL VARIATION)

Three main problems organize the analysis of analogical change in morphology: (1) the relation between leveling and extension; (2) directionality; and (3) selectivity. The selectivity problem arises in cases where an analogical change such as paradigm leveling affects some but not other seemingly equally eligible items: What factors determine this analogical selectivity? More specifically, what sorts of items resist otherwise expected analogical change? At least three specific types of analogy-resistant item have been identified in the literature: relatively frequent items; items whose alternations are in stressed syllables; and items with multiple morphophonological alternations. The last pattern has been called "Paul's Principle" (Paul 1880); all three patterns concern properties of the analogy-resistant item itself. Based on evidence from English and Latin, I will identify a new and somewhat different pattern of resistance to analogy, which I call "paradigmatic heterogeneity": items in morpho- phonologically more heterogeneous paradigms may resist analogical change, even if the locus of heterogeneity lies outside the area targeted by the change. I will suggest that Paul's Principle may be a special case of this more general pattern, that it may provide evidence for paradigms as objects in morphological analysis, and that it may cast light on the overall cause of analogical change in morphology.

Dedre Genter (Northwestern University)

Analogical processes in learning grammar

SATURDAY 14:45–15:30 (ANALOGICAL RESEARCH AND MODELING)

The acquisition of grammar has long stood as a challenge to learning accounts, leading many theorists to propose domain-specific knowledge and processes to explain language acquisition. But the discussion of general learning processes has largely overlooked the most likely candidate mechanism, namely, analogical learning. In this paper I review what is known about analogical learning and discuss evidence that these processes may be central in the acquisition of grammar.

LouAnn Gerken (University of Rochester)

Linguistic generalization by human infants

SATURDAY 9:00–9:45 (ANALOGY IN LANGUAGE ACQUISITION)

My work is directed at understanding the types of language-like generalizations that infants and young children most readily make, and the learning conditions under which they are most likely to generalize. At the workshop, I will present data from several types of studies suggesting that (1) Learners are fairly conservative in their generalizations; (2) Some amount of variability in the input is necessary for generalization; (3) Learners entertain multiple bases of generalization and rule out incorrect bases in a near absolute fashion; (4) Some dimensions of generalization are more readily learnable than others that appear equally complex; and (5) The types of generalization that are easily learned differ by domain, perhaps as a function of experience with different domains.

John Goldsmith (University of Chicago)

Learning morphological patterns in language

FRIDAY 14:45–15:30 (ANALOGY IN SYNCHRONIC GRAMMAR)

The classical definition of an analogy is a representation in which A:B :: C:D – that is, A is to B as C is to D. An elementary example would be: *coughs* is to *coughed* as *jumps* is to *jumped*, or (1):

$$(1) \quad \frac{coughs}{coughed} = \frac{jumps}{jumped}$$

This way of writing suggests the utility of a notation in which we could explicitly characterize something like (1): it is a stem followed by a choice of *s* or *ed*, as in (2), and what both sides of (1) have in common is (3), which is a pattern of a particular sort. The pattern can be used, in turn, to analyze these four words, as in (4)

(2) $\{cough\}\begin{Bmatrix} s \\ ed \end{Bmatrix}$

(3) $\{stem\}\begin{Bmatrix} s \\ ed \end{Bmatrix}$

(4) $\begin{Bmatrix} cough \\ jump \end{Bmatrix}\begin{Bmatrix} s \\ ed \end{Bmatrix}$

In the first part of this presentation, I give an overview of *Linguistica*, a system that rapidly finds patterns of this sort, with the goal of inducing the morphology of a natural language with no prior knowledge. An application of this idea to a language such as English or French finds the following:

• many spurious patterns of the form in (4) are found;
• not all correct morphological patterns in a sample can be inferred from patterns such as (4);
• too many patterns like those in (4) are found, in the sense that a linguist would say that generalizations are being missed if both (5)and (6) are postulated in an analysis of English.

(5) $\begin{Bmatrix} represent \\ assault \\ attack \end{Bmatrix}\begin{Bmatrix} ed \\ ing \\ s \end{Bmatrix}$

(6) $\begin{Bmatrix} add \\ ask \\ \dots \end{Bmatrix}\begin{Bmatrix} NULL \\ ed \\ ing \\ s \end{Bmatrix}$

• the method does not generalize easily to languages with more complex morphology.

We present a method, based on Bayesian principles, to overcome most of these problems. The software can be downloaded from <http://linguistica.uchicago.edu>.

In the second part of the presentation, I discuss a generalization of this method, based on string-matching, that allows patterns to be discovered in languages like Swahili. I will present some quantitative measures of its success, and discuss the principle remaining challenge (for both approaches), which is the necessity of forcing the system to generalize more than it is currently doing.

Keith Johnson (University of California, Berkeley)

Analogy as exemplar resonance: Extension of a view of sensory memory to higher linguistic categories

FRIDAY 14:00–14:45 (ANALOGY IN SYNCHRONIC GRAMMAR)

Early in the 20th century, Richard Semon proposed that memory for sensory experience is encoded in the nervous system with separate "engrams" for each experience. Though the engram view of memory has been abandoned in favor of a network view, exemplar-based accounts of sensory memory have appeared periodically in the memory literature ever sense then because people seem to accumulate sensory experience in a way that entails retention of detail together with abstraction over separate instances, and exemplar-based models provide a mechanism that captures both of these aspects of memory for perceptual categories. Words are interesting because they are perceptual objects in the sense that in speech communication the listener must recognize the words produced by the talker, yet they are also actions (or sequences of actions) produced by the talker, and also are symbolic objects entailing an

arbitrary relation between abstract mental categories and perceptual and action categories. In this talk I will review some research on exemplar-based modeling of the sensory perception of speech and outline an extension of this research (using ideas of resonance – Grossberg, and reentrant mapping – Edelman) to account for analogical phenomena in the synchronic mental representation of language.

Andrea Krott (University of Birmingham)

Banana shoes and bear tables: children's processing and interpretation of noun-noun compounds

SATURDAY 9:45–10:30 (ANALOGY IN LANGUAGE ACQUISITION)

I will present research showing how analogy plays a role in children's processing and interpretation of familiar and novel compound words (e.g. chocolate cake or banana shoes). The results of the first two experiments suggest that pre-school children are more likely to segment and to recognize the morphological structure of a noun-noun compound (e.g. chocolate cake) when they know other compounds with the same constituents (other chocolate words such as chocolate milk, chocolate bar). This is evidenced in children's explanations of familiar compounds and true for both English and French compounds. Two further experiments show that children's knowledge of compounds with the same head (cake) or modifier (modifier) does not only affect their understanding of familiar compounds, but also their interpretation of novel compounds. We asked 2-5 year-old children to either explain the meaning of novel noun-noun compounds or to choose among a set of pictures the one that represents the novel compound. Results suggest that already 2-3 year-olds are affected by their knowledge of similar compounds, despite their very limited vocabulary. While this effect of similar compounds on compound interpretation resembles that for adults, children differ from adults in that their interpretations are more strongly affected by their knowledge of similar compounds with the same head noun as the target compound, while adults are more strongly affected by their knowledge of compounds with the same modifier as the target compound. Thus, children focus more on how the head is used in other compounds, while adults focus on how the modifier is used. Taking all experiments together, it appears that the compound vocabulary of a child appears to facilitate his/her understanding of particular compounds rather than of compounds as a whole. It therefore supports the idea of a development from a word-based analysis to a more abstract category-based analysis of linguistic structure.

Royal Skousen (Brigham Young University)

Expanding analogical modeling into a general theory of language prediction

SATURDAY 14:00–14:45 (ANALOGICAL RESEARCH AND MODELING)

In the introduction to this paper, I will review some of the unique properties of Analogical Modeling (AM), including those that derive from the quantum computing of AM, especially the result that the heterogeneity of a given (supra)context can be determined independently of any other (supra)context. An important advantage of AM is there is no training stage beyond the need to collect exemplars; for instance, there is no need to determine in advance the significance of the variables in predicting outcomes. Predictions in AM are also made on the fly and without any reference to previous analogical sets that were earlier derived and

used. And one important distinction is that AM is not equivalent to nearest-neighbor approaches since there are cases where more distant gangs of exemplars are preferred over nearest neighbors.

In the main part of the paper, I will consider a multitude of issues that have arisen in applying AM to specific language problems. Traditionally, variables in AM have been treated in terms of completely unrelated variables (that is, variables were defined as if they were independent of each other). So one task in AM is to develop a systematic way of dealing directly with strings and trees (without trying to remake them as a vector of unrelated variables). A second issue has been how to treat scalar variables. Thus far, variables in AM have been treated as discrete categories, which has led to considerable artificiality in dealing with scalars, both discrete and nondiscrete. A third general problem has been applying AM to certain types of language variables (such as semantic variables). Related to this is the question of how variables of different types should be weighted with respect to each other; for instance, should a semantic variable count the same as a phonetic variable in predicting outcomes? In addition, treating morphological alternations as outcomes has led to problems in applicability; for instance, should we allow the verb alternation between /see-saw/ in English to apply when trying to predict the past tense for the verb /be/? Additional problems arise in dealing with multiple outcomes; for instance, in German plural formation, should the specific plural endings and the related umlauting of the vowel be treated together as a single outcome or as separate outcomes? Finally, there is the general issue of sequential applications of AM that seem quite necessary in predicting language use in time. In other words, to what degree should we allow AM predictions to be dependent on the results of previous AM predictions. In this paper I will attempt to outline some solutions to these various problems.

Gregory Stump (University of Kentucky) & **Rafael Finkel** (University of Kentucky)

Principal parts and degrees of paradigmatic transparency

FRIDAY 11:15–12:00 (ANALOGY IN LANGUAGE CHANGE AND TYPOLOGICAL VARIATION)

A lexeme's PRINCIPAL PARTS are a subset of the cells in its paradigm from which all of the other cells in the paradigm can be deduced. There are at least two ways of conceiving of principal parts: under the STATIC conception, the same cells in the paradigm of every lexeme are that lexeme's principal parts; under the DYNAMIC conception, the cells constituting a lexeme's principal parts may vary from one lexeme to another.

We define a lexeme's paradigm as MAXIMALLY TRANSPARENT if any cell in that paradigm could function as that lexeme's sole dynamic principal part; if a language learner has learned a language's inflection classes, then s/he need only learn one form in a maximally transparent paradigm in order to deduce all of the other forms in that paradigm. Inflection classes can be distinguished according to the kind and extent of their deviation from the ideal of maximal transparency.

An inflection class C may deviate from this ideal because its members require multiple dynamic principal parts: In particular, (i) different cells in the paradigm of a lexeme L belonging to C may have to be deduced from different ones of L's principal parts; and (ii) certain cells in the paradigm of a lexeme L belonging to C may have to be deduced from a combination of L's principal parts. Moreover, an inflection class C may deviate from the ideal of maximal transparency because its members have paradigms some of whose cells are less than fully informative: in particular, (iii) certain cells in the paradigm of a lexeme L belonging to C may be uninformative in the sense that they fail to determine the realization of any

other cell in L's paradigm; and (iv) certain cells in the paradigm of a lexeme L belonging to C may be only partially informative in the sense that as dynamic principal parts, they only determine the realization of other cells in L's paradigm in combination with each other.

We illustrate these distinctions with evidence from Comaltepec Chinantec (Oto-Manguean; Mexico); drawing on a computational analysis of dynamic principal parts in this language, we demonstrate that its conjugation classes are in some instances maximally transparent and in other instances embody one of more of deviations (i)-(iv).

We conclude by discussing the implications of these findings for the so-called No Blur Principle (Cameron-Faulkner & Carstairs-McCarthy 2000), the hypothesis that among the rival affixes for any inflectional cell, at most one affix may fail to be a class identifier, in which case that one affix is the class default for that cell (Cameron Faulkner & Carstairs McCarthy 2000: 816). This hypothesis presupposes that only affixal differences are relevant for distinguishing inflection classes and imposes a substantial constraint on deviations from maximal transparency, entailing that of all the inflection classes for lexemes of a given category, only one should require more than one principal part. Drawing on affixal evidence from Fur (Nilo-Saharan; Sudan), we demonstrate that this constraint is simply too strong, hence that the No Blur Principle cannot be maintained.

References

Cameron-Faulkner, T. & Carstairs-McCarthy, A. (2000). Stem alternants as morphological signata: Evidence from blur avoidance in Polish nouns. *Natural Language and Linguistic Theory* 18.813-835.

Jakobi, A. (1990). *A Fur grammar*. Hamburg: Helmut Buske Verlag.

Pace, W. J. (1990). Comaltepec Chinantec verb inflection. In W. R. Merrifield & C. R. Rensch (eds.), *Syllables, tone, and verb paradigms* (Studies in Chinantec Languages 4). Summer Institute of Linguistics & The University of Texas at Arlington, 21–62.

Andrew Wedel (University of Arizona)

Multi-level selection and the tension between phonological and morphological regularity

FRIDAY 10:30–11:15 (ANALOGY IN LANGUAGE CHANGE AND TYPOLOGICAL VARIATION)

Sound change often produces irregularity within morphological paradigms, while restoration of morphological regularity comes at the price of increasing phonological irregularity (Sturtevant 1947). Here, I will argue that we can use the tools of evolutionary theory to illuminate this conflict between phonological and morphological regularity. It has been proposed (e.g., Sturtevant 1917, Wang 1969, Labov 1972, Hock 2003) that both the spread of a sound change throughout the lexicon and local morphological analogies are subtypes of analogy, in that both involve the extension of existing patterns. Rich-memory models of lexical storage and production, such as exemplar models, provide the beginnings of a mechanistic account for phonological and morphological change as products of such a putative general pattern-extension process.

For example, in some exemplar models of language all levels of linguistic organization are represented by inter-connected exemplar-based categories (Pierrehumbert 2003, Bybee 2002). The spread of sound change is based in occasional pattern-extension at the level of sound categories in production, while morphological analogy is initiated in occasional

pattern- extension at the level of sound-meaning categories in production. Positive feedback between production and perception allows occasional entrenchment and spread of a new pattern (Pierrehumbert 2003, Bybee 2002, Wedel 2004, in press). However, because sound- and sound-meaning categories represent distinct levels of organization in the larger linguistic system, pattern extension at one level can result in pattern-breaking at the other.

This tension between structure at distinct levels of organization is characteristic of multi-level evolutionary systems, and in this paper, I will argue that a rich-memory model of language production and perception fulfills the criteria for multi-level selection. I have argued elsewhere that under an exemplar approach, some aspects of language production and perception can be properly modeled as an evolutionary system (Wedel in press).

Any set of reproducing entities will evolve if they vary in transmissible traits that affect future reproduction. In a hierarchical system such as language (as well as biological systems), reproducing categories are often themselves made up of collections of smaller reproducing categories. As a result, selection can act simultaneously at more than one level of organization. In biology, hierarchically nested, reproducing entities that can evolve adaptations to promote their own reproduction include molecules, organelles, cells, organisms, and groups of individuals (see e.g., Keller 1999).

Within an exemplar model of language production the current state of the system influences the likelihood that a particular pattern will be extended in production, which in turn influences the composition of the set of variants perceived by others. This biasing of the set of variants produced at a given time by internal properties of the system is a form of selection, just as sexual selection in biological species proceeds as a self-referential bias in the set of offspring produced in a given generation. Although selection at distinct levels can be abstracted as separate processes, they can interact in complex ways because change at one level alters the context for selection at the other. To illustrate this proposed mechanism, I will show that in a simulation of a two-level, exemplar-based categorial system, error-biases toward similarity within each level results in a continually shifting equilibrium between regularities in both levels.

Selected References

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14, 261–290.

Hock, H. H. (2003). Analogical Change. *The Handbook of Historical Linguistics*. Janda, Ri. and Joseph, B. (eds.), p. 452–457. Oxford: Blackwell.

Keller, L. (ed.) (1999). *Levels of Selection in Evolution*. Princeton: Princeton University Press.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia, University of Pennsylvania Press.

Pierrehumbert, J. (2003). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. *Language and Speech* 46, 115–154.

Sturtevant, E. H. (1917). *Linguistic Change*. Chicago: University of Chicago Press.

— (1947). *An Introduction to Linguistic Science*. New Haven: Yale University Press.

Wang, W. S.-Y. (1969). Competing changes as cause of residue. *Language* 45(1), 9–25.

Wedel, A. B. (2004). *Self-organization and categorical behavior in phonology*. Unpublished Ph.D dissertation, University of California, Santa Cruz.

— (in press). Exemplar models, evolution and language change. *The Linguistic Review*.