Simulating Recruitment in Evolution

Henk Zeevat
ILLC, University of Amsterdam
Nieuwe Doelenstraat 15 1012CP Amsterdam
The Netherlands
henk.zeevat@uva.nl

Abstract

This paper documents an attempt to recreate the recruitment of functional items in an evolutionary simulation system. It discusses the motivation for such a system, describes the prerequisites for or recruitment to happen, documents some experiments, and draws some conclusions.

1 Introduction

Recruitment is the phenomenon whereby a language adopts a lexical word to fill in a place in its inventory of functional words or where a functional word acquires a new functional use. All functional words and morphs derive by this process. In this paper, I try to give an evolutionary account of this phenomenon.

My road to this enterprise was the study of a number of discourse particles (e.g., *too* and *doch*) as part of the more general study of presupposition. In earlier work (Zeevat (2002), I was trying to explain why these particles are different from other presupposition triggers and in particular what could be the explanation of on the one hand the absence of accommodation and on the other hand the impossibility of leaving them out in the natural environments where they occur. Especially the second phenomenon is very difficult to treat in a traditional grammar like Chomsky's or Montague's (the first property merely conflicts with the standard theory of presupposition), but can be dealt with in frameworks like optimality theory by postulating expressive constraints $max(F)$ that force the appearance of the particle when the input realises the feature F . One however feels that postulating these is a mere stipulation. The present study comes closer to an explanation of these $max(F)$-constraints. The explanation relies both on the proposed model and on typological facts. The fact that it is possible to use the semantic map method (Haspelmath (2003)) on modality (van der Auwera and Plungian (1998)) and on contrastive particles (Malchukov (2004)) shows that the same expressive needs exist in many if not all languages. In this method, one compares (functional) words and morphs in different languages, looking for translation equivalents. Translation equivalence is not a one to one relation, and it is also possible that languages lack equivalents for functions that can be expressed in another language. The method gives a range of functional primitives and a picture of their proximity, which can be related to directions of development.

Given that particles (or other members of the functional inventory) expressing the function occur in vast numbers of languages, these languages must all have gone through a recruitment process building the particle. That suggests that the absence of the means to express its function leads to communicative imperfection (without it no recruitment would happen). Communicative imperfection is the force that makes the particle obligatory in the sense I indicated. In principle, it should be possible to show the avoidance of the same imperfection with psycholinguistic methods, but to my knowledge this has so far not been attempted. The expressivity constraint would be just a problematic and strategic spot in the space of possible communications that attracts devices to express its characterisation.

A model of recruitment should however be general enough to cover all recruitment, not just the recruitment of additive and adversative particles. Making the empirical case for this is deferred to future work which I plan to do with my colleague, the typologist Andrej Malchukov. The model of recruitment is an important part of an account of linguistic evolution in a wider program. This program does not aim at the explanation of the origin of human languages but at giving a mathematical reconstruction of the forces behind language change. The next paragraphs give a brief discussion of what should be in a more complete model.

I take it that there are two forces at work in the history of human languages (Boersma (1998) is taking very much this line when discussing the functional foundation of phonological constraints in optimality theory.). The first is economy. Phonetic effort is reduced by words losing phonetic features and becoming less complex and by words merging. The other force is expressivity. Within the range of what is possible certain forms are selected for certain functions and thereby make those functions better expressible. One way in which expressivity is enhanced is by the invention of new words or new combinations of words for new concepts and their subsequent adoption by others. It has been pointed out that these processes are rather arbitrary and often involve fashion and conscious imitation. But in the case of the functional inventory, this cannot be the whole explanation. Conscious invention is difficult to imagine in this case, since the concepts involved are abstract and hard to delineate. An explicit proposal to use a new word for e.g. definiteness or adversativity is thereby hard to make. But sociolinguistic factors seem to play a role in the spread of new uses. For a full account, one has to allow for sociolinguistic factors: one of the functions of talk is to advertise one's personality and social standing. In selecting lexical words and pronunciation this may well be the decisive factor, especially in linguistically mixed situations. Broadening the concept of communicative success to the full effect that the speaker wants to make on the hearer, it turns out that social factors are not that different from the kind of factors I am considering here. I merely claim that some historical processes can be well understood by relating them to core communicative success: getting certain content reliably across to the hearer. This paper starts from an analysis of communicative success in section 2.

Here are some aspects of linguistic change that a full evolutionary theory of linguistic change needs to account for.

Phonetic reduction

These are processes that lead to disappearance of phonetic features from the realisation of words, or more generally, to reduction of phonetic markedness in any dimension one can think of. It is easy to model in a model of communicative success if one starts from the observation of variation in the realisation of certain features, both within language communities especially where the realisation of the feature is by a continuous parameter. If a reduced form is as good in communicating the content, there is no pressure to use the full form and the reduced form becomes a more easily pronounceable variant. This gives it an advantage especially with language learners and that may make it, over time, the dominant variant. If the feature is important for communication it will stay in place. Especially interesting here are global changes effecting many words. If they normally do not infringe on communicative success, they may well do so in special cases and create expressive deficiencies there. The ambiguity of bear in English comes from such processes, many grammaticalised items show signs of phonetic reduction, e.g. the English determiners with respect to their demonstrative (that > the) and cardinal (one > a(n)) sources.

Agglutination processes

These are the processes that merge separate words into single ones and are responsible for the emergence of inflectional and derivational morphology. Loss of word-stress and resulting clitic formation seem important phases. It is tempting to see these processes as naturally arising from phonological reduction, but there are important synchronic factors as well. The possibility of losing word accent and the integration with the host depend on the sentence accent and the rules for the word accent. The selection of the host must be dependent on syntactic factors. Whatever the right analysis here, it is a reduction process and thereby conditioned by communicative success: a process that destroys communicative success is not possible, global processes that only impair communicative success in special cases lead to expressive deficiencies.

Adoption of new lexicon

There are many sources for new lexicon, including free invention and borrowing from other languages. It may happen to gain expressive possibilities or the new word may have social advantages. From the perspective of this paper, the most interesting case is where the word comes from an existing word with a different meaning. I am interested in the case where new functional words arise from existing, possibly non-functional words. Words like television, radio or gas have been made up, Dutch apotheek (pharmacy) is Greek for a storage facility, and saddle for the seat on bicycles is metaphorical transfer. These are the cases this paper is not about.

Loss of lexicon

Lexicon may be lost by the adoption of new words which take over the expressive dimension of the old word or by historical changes eliminating the concept expressed from the communicative habits. It is straightforwardly incorporated in the model we adopt. In the first case a new word gains the upper hand in expressing a concept. Here evolution reinforces that tendency and – in the absence of contrary factors like social or geographical isolation – inexorably eliminates the old word. If the word becomes very rare by a change in culture, this will be reflected in it not being learnt by new generations.

Morphological reduction

Functional morphs can be lost like words by being pushed away by new morphs and also by phonological reduction. In creolisation, there is the extreme situation that morphological marking may not be a successful communication strategy and has to give way to more explicit marking strategies. A famous case is the loss of case marking in most West-European languages.

Recruitment

This is defined as the rise of a new functional use of an existing word or morphological category.

Recruitment is normally seen as just one aspect of grammaticalisation, the process by which languages acquire their functional inventory. Often it is expressed as a property of the recruited word. The word (paradigmatic form) undergoes the following processes in grammaticalisation (see e.g. Bybee *et al.* (1994)):

1. semantic bleaching (it gets a vaguer, more pragmatic meaning)
2. rise in frequency
3. possible phonological reduction
4. reduction in syntactic combination possibilities
5. loss of optionality

The structure of the paper is as follws. In section 1, successful communication will be analysed and related to the iterative model of evolution of Kirby and Hurford (2002). In section 2, the conditions of recruitment will be discussed. Section 3 discusses the experiments and their mathematical analysis. Section 4 tries to draw some conclusions.

2 The Gricean Theory of Evolution

Grice (1957) gives the following definition of non-natural meaning.

(1) S (an speaker) meant something (non-naturally) by A (an utterance or gesture) if and only if S intended the utterance or gesture A to produce some effect in an audience by means of the recognition of this intention

It does not just define non-natural meaning, it can also be read as a definition of communicative success.

(2) a communicative act A of speaker S is successful iff it has the intended effect on hearer H by means of A

In simple forms of communication like ordering a beer, requesting the salt or asking a question, the intended effect on H is easy to monitor. It is often

underestimated to what extent feedback is coming back in more complicated forms of communication. Clark (1996) is largely devoted to showing that dialogues are composed of joint actions with each move consisting of a speaker action and the hearer's uptake that are dovetailed to each other and in which feedback is continuously asked and provided. It is not problematic at all to ascribe to participants in simple dialogues a good deal of insight into the success of the ongoing communication, though there are genres (e.g. writing scientific paper or classical university lecturing) where the conditions are far less favourable.

So the probability of the success of an utterance U for an intention I can be described as the probability that the hearer recognises I from U. This can be estimated from a corpus of utterances with the interpretation as the probability that U is interpreted as I. (In principle, collections of observed utterances could be used for this purpose. Here, the corpus is just made up as an assignment of probabilities to pairs of utterances and interpretations.) The hearer has the situation of utterance, her model of the conversation so far, the speaker can employ facial expression, special intonation and of course the feedback mechanisms. Nonetheless, one obtains an approximation of the probability by looking at the corpus under the assumption that the corpus probability is one of the factors that determines actual error.

The success formula needs to be modified in one respect. It seems reasonable to assume that even in successful communication there are imperfections. In fact, it may seem unrealistic to assume that the full intention of the speaker is ever grasped. It is comparable to the speaker having an attitude: the speaker wants something, believes something, is afraid of something. Following Haas-Spohn (1995), the somethings here need to be individuated not as propositions in the tradition of Kripke and Kaplan, but as thoughts, the subjective meanings of the words by which one would express the attitudes (if one would choose words to which the subject could assent). Without this assumption, many problems in the area of attitudes are untreatable. It follows that communication can never be literally successful. The subjective meaning of the speaker has to be compared with the subjective meaning the hearer arrives at. They are at best similar in various ways. Our languages have evolved for optimising the similarities. But also for tolerating discrepancies.

So degrees of success are needed. It seems reasonable to have a distance function $d(I, J)$ that measures the distance between the speaker intention and the hearer interpretation. It will be zero if the similarities are maximal and higher if it is less than optimal. Its main function is to make it possible to give more weight to some errors than to others. Pure identity of intention and interpretation does not allow that.

The clever thing for the speaker is to express her intention I with an utterance U such that the probability that the hearer will recognise it as expressing J with $d(I, J)$ minimal is maximal. This is sophisticated and maybe this is what speakers do. Pickering and Garrod (2004) make it clear that participants in a dialogue go to considerable lengths in adapting to each other. Our assumption is that speakers trust in the wisdom of the generations: they select the utterance according to probabilities in the corpus of choosing utterances for the given intention. It is not incompatible with adaptive strategies: adaptation would be reflected to some extent in

the corpus. A reason for not choosing to model an adaptive strategy is that it would just speed up evolution: the speaker would avoid low success strategies even if they are frequent in the corpus and the low success strategies would hardly reproduce. We get the same effect by making the new corpus dependent on success of communication.

There is a similar dilemma for the hearer strategy. The best thing that the hearer could do is to select an interpretation that maximises the chance that it is minimally different from what the speaker intended. So she could study the probabilities of using utterances for intentions. Again we let her follow the wisdom of the generations and just choose an interpretation according to the chances in the corpus of interpretations for the given utterance.

A communication event is a triple IUJ and its success is $1 - d(I, J)$. Evolution reproduces communication events to the degree of their success. We model this by adding IU $d(I, J)$ often to the new corpus. We could also add JU $d(I, J)$ often, but it does not make a difference. The adaptation to the degree of understanding and the perspective of the other participant is automatic.

Language evolution starts from an initial corpus of forms and intentions with a frequency. Communication takes place by selecting an intention from the corpus according to the frequency of the intentions expressed in it, then a form for the intention according to the corpus frequencies, then an interpretation for the form according to the corpus frequencies.

The initial corpus will be taken as fixing the frequency of the intentions for the whole course of evolution. This is a necessary and substantial idealisation. That it is substantial can be seen from inspecting oneself as a second language speaker: one says different things in different languages. That it is necessary follows from the difference between what one wants to say and how one says it. In language evolution the forms should change, but one will want to say roughly the same things. From a more technical perspective the idealisation is crucial. The model works because one wants to say things that one cannot say with a reasonable chance of being understood. Without the idealisation, it is what speakers want to say that will adapt to communicative success and not the relation between forms and intentions.

3 Conditions for Recruitment

Recruitment happens under five conditions.

1. *the new communicative function is unoccupied*

This may seem to require least motivation, but it is perhaps the most problematic when one looks at historical examples of recruitment, since it has been observed that often other devices expressing the function are around when recruitment happens Bybee *et al.* (1994). There is maybe a way out. In the next section, one experiment will show that it is possible to have forms that only apparently have a function: they are systematically misunderstood and the form contributes only very little to proper recognition. The existence of a bona fide form for the new function will however block recruitment systematically in the model. It does so, because a bona fide form has a higher success rate than the candidate recruit before it has

changed its meaning. So it is not possible to give up on this assumption. It is modelled by having no proper expression for the new function.

2. *without the new use of the recruited word, the new function is systematically missed, because all appropriate forms have a dominant other interpretation that omits the new function.*

This is a more general formulation of (1). In the model, (2) is modelled in a different way. There is an unmarked form that the new function shares with its negation and the new function is less frequent than its negation.

Negation is not correct terminology: it is only with truth-conditional meanings. In other cases, it is better to think of the function as a feature and of the negation as the absence of that feature.

3. *the old function of the form weakly entails the new function*

This assumption uses the weakest possible entailment relation: If A holds B is more likely than its negation. It is modelled by initially setting the combination of the old function with the negation of the new function at a lower frequency than the combination of the old function and the new function. In models of semantic change, this aspect is sometimes neglected. The old meaning must make the new meaning plausible.

There is a large class of relations that lead to weak entailment importantly including next to normal entailment, pragmatic implicatures and natural probabilistic dependencies. But see the discussion of metaphor – the apparent exception – in the last section.

4. *the new function is not less frequent than the old function*

If it is, the old use will suppress the new use. It will just not make headway.

5. *misunderstanding the form with the new function as having the old function is less problematic than not recognising the new function*

Without this assumption the penalty of overmarking the old meaning (as the old meaning weakly entailing the new meaning when the old meaning itself is not intended) outweighs the gains that are made by (weakly) marking the new function. It is here modelled by letting d(Old & New, ¬ Old & New) ¡ d(New, ¬ New). (For a different view, see again the conclusion about metaphor and semantic epenthesis.)

It would be desirable to eliminate some of these conditions. There is however not a substantial conflict between the conditions and what is understood about historical cases of recruitment. My simulations seem to show that nothing happens when one of these conditions does not hold.

4 Experiments and Analysis

The evolutionary model is implemented in Prolog as an update function on the corpus, here taken to be a function from form-meaning pairs to probabilities. The original corpus is C_0, functions C_n are obtained by iterating the update function.

The update function is defined as follows.

$f_{n+1}(I,U) = \Sigma_{V \in UTT} C_0(I,V) * \frac{C_n(I,U)}{\Sigma_{V \in UTT} C_n(I,V)} * \frac{\Sigma_{J \in INT} C_n(J,U) * d(I,J)}{\Sigma_{J \in INT} C_n(J,U)}$

This is the product of the frequency of I with the chance that U is chosen for I with the correctness of the understanding J (its similarity to I). It

is computed for all form-meaning pairs and then normalised to 1:

$C_n(I,U) = \frac{f_n(I,U)}{\Sigma_{J \in INT, V \in UTT} f_n(J,V)}$

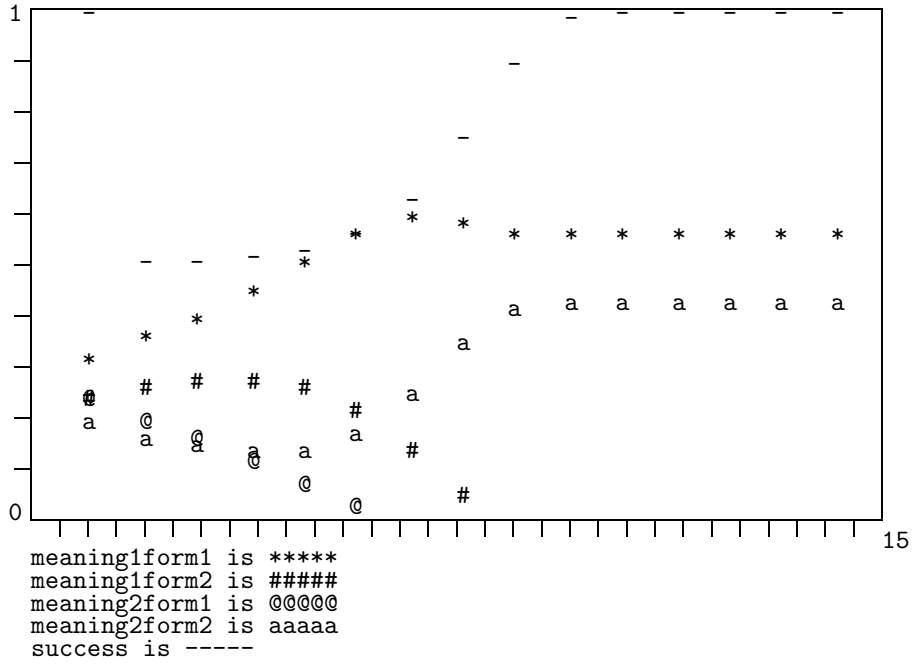
Most experiments stabilise after 50 generations or so, often even quicker.

Experiment 1: Form-Meaning Iconicity

Two meanings M_1 and M_2 are competing for two forms F_1 and F_2. The initial assignments of probabilities are random.

Development

Normally a quick convergence to a strict (100%) alignment of the most frequent meaning with the most frequent form and less frequent from with the less frequent meaning. Exceptions arise if the most frequent form is not also most frequent in the more frequent meaning in the initial assignment. This can lead to the inverted pattern.



Discussion

The experiment rebuilds the formation of a meaning opposition between two expressions and embodies some aspects of the iconicity thesis: the most frequent form normally goes with the the most frequent meaning. If we combine it with Zipf's law which expresses the tendency of short forms to go together with frequent words, it explains why frequent meanings go together with short forms. The law can also be underpinned by assuming that phonological reduction is speeded up by frequency which gives that, eventually, frequent meanings are expressed by phonologically simple words.

But it is not so clear that results of this kind go much further than that.

In particular, it starts from an arbitrary association of both words with both meanings and so is not a model of recruitment of either word. It explains why the forms become obligatory for expressing their meaning, but that is just an artefact of having just two forms and two meanings.

So the experiment turns out as it does, but it seems unrelated to occurrences in the history of human languages: initial settings need to be natural.

In particular, it does not seem to have bearing on the obligatory occurrence of particles or articles. Here the opposition is typically between a form with the particle or article and a form without it. The obligatory use of particle or article can be explained from an earlier optional use, but not the recruitment of the particle or article in that new role. The next two experiments discuss two exceptions to the pattern established in experiment 1.

Experiment 2: Donkeys and she-donkeys.

The word "donkey" is ambiguous between generic donkey and male donkey, she-donkey is unambiguous for female donkey.

It may be argued that it is not always vital to be able to express the male/female distinction, i.e. that there are three possible inputs: generic, male and female. Overinterpreting donkey for generic as male or female is not a total failure, and neither is underinterpreting she-donkey as generic donkey. So $d(generic, male) = d(generic, female) = 0.5, d(male, female) = 1$ and $d(X, X) = 0$.

This gives two words "donkey" and "she-donkey" and three meanings generic donkey (D), female donkey (DF) and male-donkey (DM) and one regularity only:

she-donkey means DF

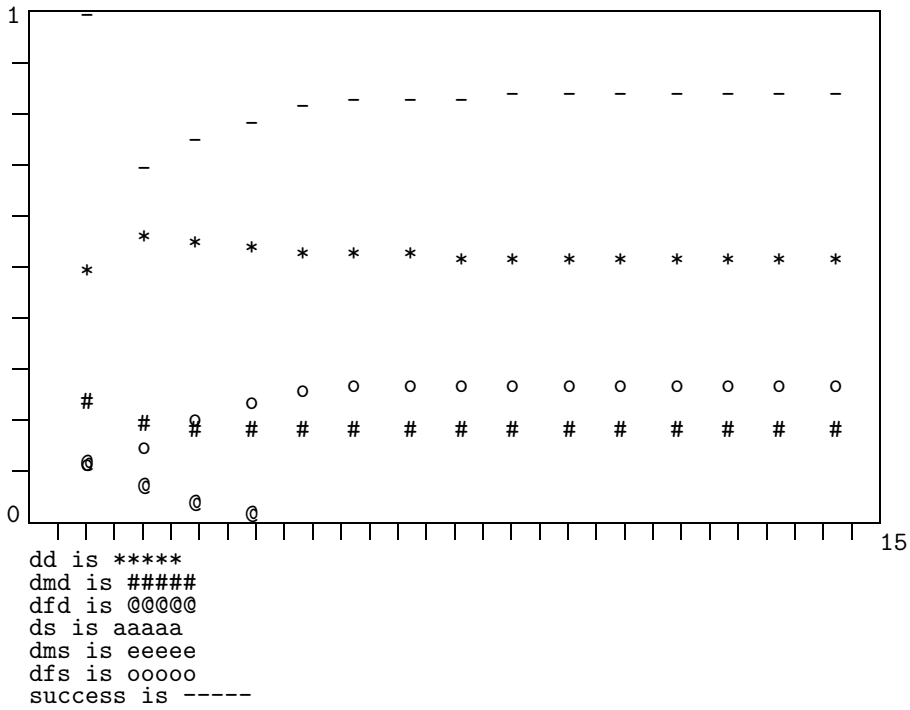
I make the assumption that in 50% of the cases the input is D and that the other 50% is equally divided over DF and DM, with DF half of the time expressed by "she-donkey".

(3)
D-donkey,50
DM-donkey,24
DF-d,12,
D-she-donkey,1
DM-she-donkey,1
DF-she-donkey,12

In addition, semi-success is allocated to the over- and underinterpretations that were indicated.

The result is a system in which DF-d is blocked and DM-d is somewhat depressed: "donkey" is obligatory for DM but donkey is normally interpreted as D. The system has no ambition to reach the pattern of experiment 1. The system reflects a normal analysis of the opposition here: she-donkey is obligatory if it is really intended to mark the female gender of the donkey, but not otherwise. Also male-donkey is the marked meaning of "donkey", that need reinforcement by contrast to be expressed as in (4).

(4) That is not a donkey, it is a she-donkey.



Experiment 3. Optional Discourse Marking

Another case of optional marking is the optional insertion of "because" in a sentence that expresses the cause of the last sentence before it. "Then" marks it -also optionally- as the next event. It is different from the optional gender marking of donkeys: "because" means cause, but the strategy of leaving out the marker is misdescribed as unmarked, it uses the fact that the second sentence expresses a normal cause of the first sentence.

(5) John fell. Bill pushed him.
(?) John fell. Because Bill pushed him.
John fell. Then Bill pushed him.
John fell. Because Mary smiled at him.
John fell. Mary smiled at him.
(?) John fell. Then Mary smiled at him.

There are really four marking strategies, with zero falling apart into the case where the second sentence is a "cause default" and the case where that is not so. If we disregard the possibility of marking by then, we have three "forms".

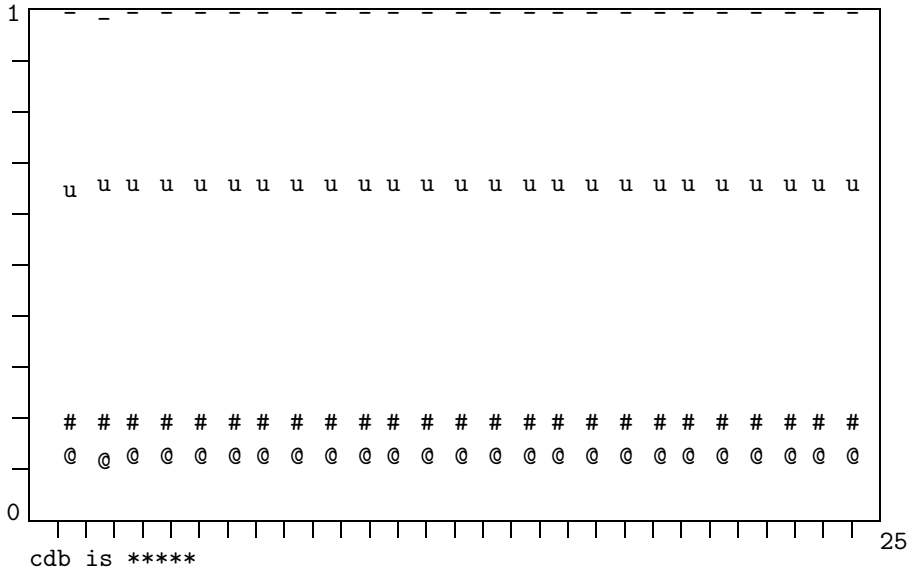
(6) zero expressiona and cause default
zero expression and not a cause default
because

The semantic outcomes seem to be: *cause* and *nextevent*, but only the normal causes can mark themselves as causes. So we have to distinguish

the inputs. The success rates can be adjusted to equate normal and abnormal causes, and next events that are also normal causes with plain next events.

If "because" is never used with normal causes, and always implies "cause", the system is stable. If some error is allowed where the default strategy and "because" are combined, it easily happens that "because" becomes obligatory.

	normal cause	because	nothing
normal causal	0.25	0.01	0.01
causal	0.01	0.05	0.01
temporal	0.01	0.01	0.65



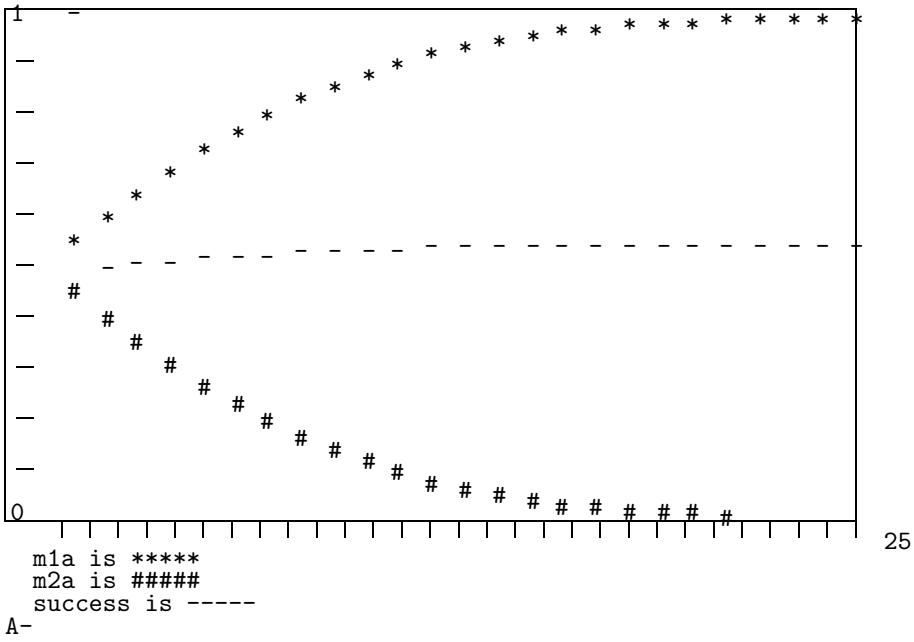
Experiment 4: Moot Expressions

If two meanings are competing for the same form, the marked meaning will eventually disappear. There are two variables, FM_1 and FM_2, with a different probability. If $C_0(F, M_2) < C_0(F, r_1)$, $C_n(F, M_2)$ will converge to 0 and $C_n(F, M_1)$ to 1.

In our model, ambiguity can only survive when the two meanings are equally strong[1] The reason is simple. The minority meaning is not well expressed by the form, the odds are that it is misunderstood. The odds determine its survival rate, even in a situation where there is no alterna-

[1]This is an artefact of using precise probabilities and may seem odd. The system can be set up so that probabilities are normalised in categories like very large, large, larger, roughly half, smaller, small, very small. This would allow more stability.

tive means of expression available. It is a paradoxical result. The model predicts that F will still be used just as often for M_2 as at the outset: there is no alternative. It will just never be interpreted as M_2.



Experiment 5: Recruitment

Experiment 4 is the basis for example 5, which is the model of recruitment. It contains a form with an old meaning, the old meaning weakly entails a new meaning, and the new meaning is expressed by a form that is overwhelmingly used for the default meaning.

Weak entailment between meanings A and B is the circumstance that if A, B is more likely than $not-B$. The default meaning is $not-A\&notB$, the new meaning is $B\&not-A$, the old meaning is A, falling apart into $A\&B$ and $A\&not-B$. There is a marker for A that invariably appears on the two A-meanings. It does not make a difference whether the default meaning and $B\&not-A$ share a different marker or are unmarked. We further have some tolerance for overmarking: extra A or B is only half-failure. This gives the following assumptions for the experiment.

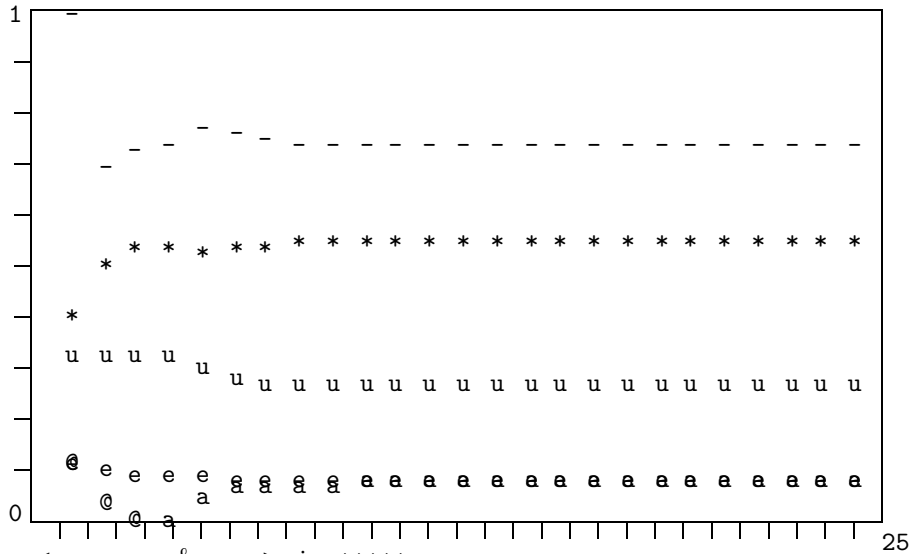
$notA\&notB > notA\&B$(dominant default meaning)
$notA\&notB > A\&B$ (A&B is the marked meaning)
$A\&B > A\&notB$(weak entailment)
F iff A (F means A)
$d(notA\&notB, A\&notB) = d(notA\&notB, notA\&B) = d(notA\&B, A\&B) = d(A\&notB, A\&B) = 0.5$(tolerance for overmarking)

Result:

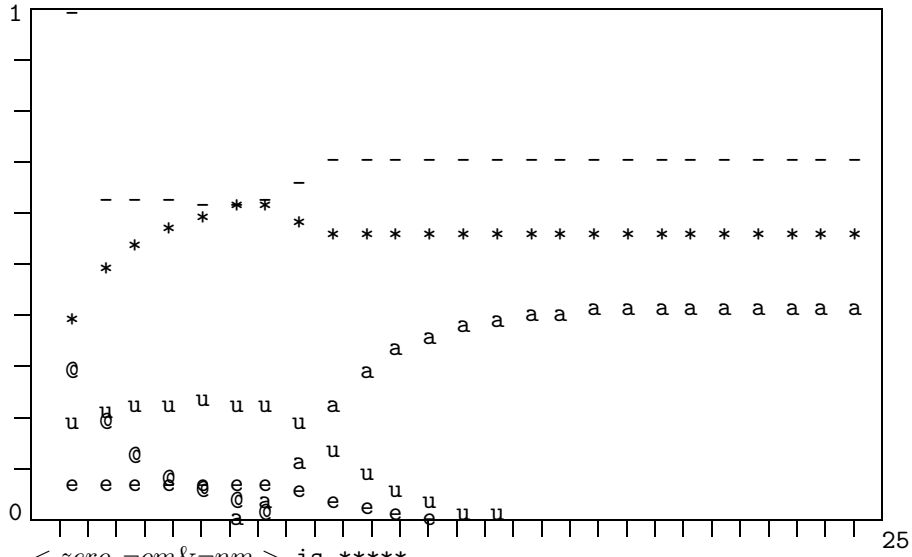
$notA - B - zero$ is depressed by $notA - notB - zero$ as in example 4

converging to 0. $notA-B-F$ is made stronger by overmarking tolerance and acquires some status. If $notA-B-F > A-B-F$ then $notA-B-F$ usurps F. Otherwise $A-B-F$ (and possibly $A-notB-F$) keep a positive but depressed value. The latter case is an example of how another meaning ($notA-notB$ for zero) can keep an ambiguity alive: $A-B-F$ depresses $notA-B-F$ but the alternative $notA-B-zero$ is even more strongly depressed by $notA-notB-zero$.

In simulation experiments, one can show all the conditions to be essential. Example 4 can be proved by an analytical argument (see appendix) and it can be used as a lemma to prove example 5. It follows from the lemma that $notA-B-zero$ converges to 0. By overmarking, $notA-B-F$ has a small positive value deriving from interpreting F as A-B. It follows that $notA-B-zero$ will become smaller than $notA-B-F$ and will be employed more often. This advantage will rise until all $not-A-B$ is expressed by F. If F is more frequently successfully used for $notA-B$ than for $A-B$ the lemma applies again and $A-B-F$ becomes moot.



Spreading pattern



Usurpation Pattern

Discussion.

Two kinds of recruitment are modelled. The source item can acquire the target meaning and lose the source meaning. The new use usurpates it. In the other case, the source item acquires an extra meaning without losing its old meaning. After usurpation, the distinction $A/notA$ is no longer expressible. The need for this expression can be responsible for a new recruitment of some marker G that weakly entails A.

5 Some Conclusions

Recruitment and Grammaticalisation

The following list of properties are standardly associated with grammaticalisation. I will go through them trying to relate them to the model.

1. semantic bleaching (it gets a vaguer, more pragmatic meaning)

2. rise in frequency

3. possible phonological reduction

4. reduction in syntactic combination possibilities

5. loss of optionality

Semantic bleaching is captured by means of weak entailment. The core of the new meaning is contained in (or just activated by) the old meaning. The old meaning (or that part of it that is not contained in the new meaning) is lost. A later strengthening is not excluded. In fact, this is one of the processes that follows from experiment four. Not just that marked minority meaning becomes moot, but also stereotypical majority meaning becomes dominant.

The new meaning is more pragmatic, because it belongs to the functional inventory. It is not just the case that recruitment processes by definition build functional items. Lexical deficiencies are conscious and can be filled in by syntactic combination and by invention or borrowing of new words. The proper recruitment process seems superfluous for lexical items.

Functional items are by their very nature more frequent than lexical items. This is enforced by their more pragmatic meanings and the lack of optionality. Pragmatic meanings correspond to dimensions that are there in every utterance. It would seem that even marked positions on these dimensions are more frequent than proper lexical material. If that is so, the rise in frequency follws from non-optionality.

Loss of phonetic properties can be connected to higher frequencies but also to loss of accent typical for the functional inventory. This aspect is not modelled.

Loss of optionality results from the extreme unattractiveness of the unmarked expression in which another meaning is dominant.

Push and pull

Some authors have taken the view that gramaticalisation is something that just happens to an item. Others have emphasized the importance of expressing certain functons (modal meanings, tense, aspect, etc.) It is clear that the model in this paper assumes both a pull (an unexpressed distinction that leads to defective communicative success) and a push (lexical or functional items that express something containing the target meaning, but also material that is unintended in the new use). If the source does not weakly entail the new meaning nothing happens. If the new meaning is not of strategic importance for communicative suceess nothing happens either. The typological data on functional items strongly suggest that it is possible to lack means of expression. The explanation can be that other means are used for expressing the distinction (e.g. aspect expressing definiteness in the Slavic languages) but a more banale explanation is lack of appropriate sources.

Weak entailment, metaphor and overmarking

The strongest reason for doubting the current model is recruitment that seems based on metaphor (e.g power to possibility, movement in space to movement in time). Metaphors do not straightforwardly give rise to weak entailment, but as a communication strategy seem based on the reliability of the inference that the literal meaning cannot be the intended one. Perhaps this is also important in the other cases.

The overmarking mechanism seems to play a crucial part in the model. If there are two features A and B with A entailing B to some extent, one can

make the assumption that a marker for A is defective for A & not B (a marked case) and also for the other marked case B & not A, while being good in A & B.

One way of motivating that is by saying that there is no harm in the hearer assuming that A is intended if it is not (it would be just a complication), while it would be problematic that B would not be recognised. Or harmless to assume B, while harmful not to recognise A.

In interpreting (7)

(7) (∅/How about that baby?) It is cold.

one should realise that "it" does not have a contextual referent (the baby) before concluding that the sentence is about the weather. A fruitless search for the contextual referent before concluding to generality is not a communication error: it is wasted energy.

Similarly, in entertaining the possibility that the speaker is surprised by the early timing of the interlocutor's meal before concluding that she is just doing a colloquial perfective question, no communicative harm is done.

(8) You eat already?

The important factor in these cases may well be the availability to the hearer of a mechanism for inferring that there is no antecedent for it or that the speaker is not surprised, an availability exploited by the speaker. That would make it indeed very close to metaphor. It would be the semantic epenthesis model from Fong (2003) ((8) is also taken from her paper).

It is a bit surprising that there is a sense in which the suppression mechanism is incorporated in our model. Given that there is a chance of $not-A$, and an intensional correlation between $F\&not-A\&B$ (the hearer is predisposed to infer B if $F\&not-A$) getting some of the points for an interpretation $A\&B$ when $not-A\&B$ is intended captures the effect of the viable communication strategy of saying F when $not-A$ is common ground while intending the hearer to infer that B (by metaphoric processes or by weak entailment) because F cannot mean A when $not-A$ is common ground.

It is hard to make an educated guess for the power of the alternative communication strategy I discussed. But the same holds for the degree to which overmarking strategies are successful. Estimating the degree of success is however not important: the processes will happen for any estimate. Weak entailment can still not be eliminated from the model. Without it, $A\&not-B$ will usurp, before recruitment can happen.

Lexicographical nightmares

One of the surprising results of the model is the existence of two kinds of recruitment, usurpation and spread. Spread can be held responsible for creating lexicographical nightmares: modal verbs, prepositions, certain particles (*doch*). And case markings, articles and aspectual marking should be counted in with this category, even though they ares rarely considered the business of a lexicographer. Here one finds lots of seemingly unorganised meanings. Though even here, some have argued for a core meaning and a derivation process, it seems far more fruitful to think of the intensional relationships between the different uses as the stuff that

makes spread possible rather than as the cognitive glue that keeps the different uses together.

If the model is correct, there should be nightmares of this kind. This is not to say that usability of an item should not meet a criterion of efficiency: it should be possible to decide in very finite time which use one is dealing with in a particular situation. There may well be a processing restriction of this kind, but I am not aware of any strong argument in favour of that and neither of arguments that show that quick elimination of the other uses is impossible. The mechanism of overmarking (the item is expressing a set of semantic features and the concrete use selects a subset) as proposed by Fong (2003) seems prima facie a quite reasonable alternative to logically inspired models. (In the view of this paper, such logical models do a kind of historical reconstruction.)

Cycles

The model predicts cycles of recruitment, especially around usurping recruitment of functional items (though even iun spread, the old meaning can get moot). If an item is recruited, it leaves behind an expressive disaster area in its old meaning of the just the same kind that causes recruitment. If another functional items is recruited, this need for recruitment is further promulgated. It would seem a slow wave pattern.

If language is stable, it has also finished recruiting. But phonological decay (e.g. the disappearance of a morphological distinction) is a guarantee that this cannot be a permanent state. So cycles are the rule and not the exception.

Weaknesses of the model

That both overmarking and weak entailment is needed makes the system harder to interpret than is pleasant. More concrete historical case studies may help to gain a clearer picture here.

An unpleasant idealisation is that new meanings are categorised as distinct meanings with a label. This may be wrong, since often meanings seem continuous parameters that drift in time. It also suggests that not marking the new meaning is something bad by itself: it is misunderstanding 'the meaning'. The better interpretation is in terms of the processing of the hearer. If the new meaning is unmarked, the hearer will be likely to misprocess the utterance to some harmful degree. This suggests a formulation where pairs of utterances and positions in intention space are weighted by a continuous function that gets reset by communicative experience. It is not clear though that a model of this kind will be better in doing the modelling that we did here. But in interpreting the model, these idealisations need to be borne in mind.

Intelligence

People consciously use metaphors, irony and overstatement. Our language users are in comparison just zombies, trying to model their own behaviour on perceived habits. It may be that the model can be improved by making the users more intelligent and creative.

The model suggests that intelligence is important. Success rates can be dismally low and do not improve by evolution: sometimes they even drop. Good speakers must do better than our users to guarantee that where it is important communicative success will be reached. Good hearers that can empathise with the speakers will do significantly better. Language by

itself is a very imperfect method for communication, but good speakers can do miracles.

References

Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Ph.D. thesis, University of Amsterdam.

Bybee, J., Perkins, R., and Pagliuca, W. (1994). *The evolution of grammar. Tense, aspect, modality in the languages of the world*. University of Chicago Press.

Clark, H. (1996). *Using language*. CUP, Cambridge.

Fong, V. (2003). Unmarked 'already': Aspectual expressions in two varieties of english.

Grice, H. (1957). Meaning. *Philosophical Review*, **67**, 377–388.

Haas-Spohn, U. (1995). *Indexikalitt und subjektive Bedeutung*. Berlin.

Haspelmath, M. (2003). The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In M. Tomasello, editor, *The new psychology of language*, pages 211–243.

Kirby, S. and Hurford, J. (2002). The emergence of linguistic structure: an overview of the iterated learning model. In A. Cangelosi and D. Parisi, editors, *Simulating the Evolution of Language*, pages 121–148. Springer.

Malchukov, A. L. (2004). Towards a semantic typology of adversative and contrast marking. *Journal of Semantics*, **21**, 177–198.

Pickering, M. and Garrod, S. (2004). Towards a mechanistic pyschology of dialogue. *Brain and behavioral science*, **27**.

van der Auwera, J. and Plungian, V. A. (1998). Modality's semantic map. *Linguistic typology*, **2**, 79–124.

Zeevat, H. (2002). Explaining presupposition triggers. In K. van Deemter and R. Kibble, editors, *Information Sharing*, pages 61–87. CSLI Publications, Stanford.

Appendix

Experiment 4 (moot expressions) allows of an underpinning by proof. Let d be the initial strength of the majority meaning and x_n its strength in the n-th generation.

Lemma: Let $0.5 < d < 1$. Let $x_n, n > 0$ be given by:

$x_0 = d$
$x_{n+1} = \frac{dx_n}{dx_n + (1-d) * (1-x_n)}$
Then $x_{n+1} > x_n$.

Proof:
$\frac{dx_n}{2dx_n - x_n - d + 1} > x_n \Leftrightarrow$

$\frac{d}{2xd_n - x_n - d + 1} < 1 \Leftrightarrow$

$d < 2dx_n - x_n - d + 1 \Leftrightarrow$

$2d - 2dx_n < 1 - x_n \Leftrightarrow$

$2d(1 - x_n) < 1 - x_n \Leftrightarrow$

$2d > 1$ which is given

So we have shown that $x_{n+1} > x_n$

Lemma: x_n converge to 1.

Proof: Let's have a closer look at

$fx = \frac{dx}{2dx - d - x + 1}$

$\frac{dx}{2dx - d - x + 1} = \frac{\frac{d}{2d-1}x}{x + \frac{1-d}{2d-1}}$. Let's call $\frac{d}{2d-1}$ a and $\frac{1-d}{2d-1}$ b. We can now write fx as $\frac{ax}{x+b}$

$f(x) - x$ can be written as $\frac{ax}{x+b} - x = \frac{ax}{x+b} - \frac{x(x+b)}{x+b} =$
$\frac{x(a-b)-x^2}{x+b}$

$a - b = \frac{d}{2d-1} - \frac{1-d}{2d-1} = \frac{2d-1}{2d-1} = 1$

So our difference reduces to $\frac{x-x^2}{x+b}$. This is larger than $\frac{x-x^2}{b+1}$ since $0.5 < x < 1$.

Now assume $l < 1$ is a limit for x_n and consider $\epsilon = \frac{l-l^2}{b+1}$.

We show that $l - \epsilon + \frac{(l-\epsilon)-(l-\epsilon)^2}{b+1} > l$ and a fortiori that $f(l - \epsilon) > l$ and for any x_n such that $l - \epsilon < x_n < l$ $f(x_n) > l$ and so that there can be no such l.

$l - \epsilon + \frac{l-\epsilon-(l-\epsilon)^2}{b+1} > l$ iff

$\frac{l-\epsilon-(l-\epsilon)^2}{b+1} > \epsilon = \frac{l-l^2}{b+1}$

This is true because $f(x) = x - x^2$ is monotone descending in (0,1).

So our sequence x_n converges to 1. (And 1 is a fix point for $f(x)$)