

How structure acquires function: On the evolution of (un)useful case markers in Upper German

Guido Seiler

Universität Zürich, Deutsches Seminar (University of Zurich, German Department)

Schönberggasse 9, CH-8001 Zürich, Switzerland

gseiler@ds.unizh.ch

Every phenomenon or process in living organisms is the result of two separate causations, usually referred to as proximate (functional) causations and ultimate (evolutionary) causations. All the activities or processes involving instructions from a program are proximate causations. [...] They are answers to “How?” questions. Ultimate or evolutionary causations are those that lead to the origin of new genetic programs or to the modification of existing ones – in other words, all causes leading to the changes that occur during the process of evolution. [...] They are usually the answer to “Why?” questions.

It is nearly always possible to give both a proximate and an ultimate causation as the explanation for a given biological phenomenon. [...] Many famous controversies in the history of biology came about because one party considered only proximate causations and the other party considered only evolutionary ones. (Mayr 1997:67)

1. Introduction

Dialectal grammatical systems are a promising field for the empirical study of diachronic change in progress. In particular, the study of dialect variation uncovers the importance (and limitations) of functional factors in grammatical change, since dialects differ not only in their inventories of formal devices but also in the functional properties of these devices. One such case is prepositional dative marking (PDM) in Upper German (Bavarian and Alemannic). Upper German distinguishes three morphological cases: nominative, accusative, and dative. In several dialects dative DPs are preceded by a prepositional marker that is homophonous with the prepositions *an* (‘at, beside of’) and *in* (‘in, into’):

- (1) Bavarian, AN: *du muasst es a deinà frau vaschraibn lássn*
you must:2s it DM your:Dsf wife transfer let:Inf
‘you have to transfer it [=the money] to your wife’
(Malching; Ströbl 1970:66)
- (2) Bavarian, IN: *sàg’s in der frau*
say-it DM the:Dsf woman
‘say it to the woman’
(Upper Inn Valley; Schöpf 1866:286)
- (3) Alemannic, AN: *er git dr Öpfel a mir, statt a dir*
he gives the apple DM me:D instead DM you:D
‘he gives the apple to me, not to you’
(Glarus; Bähler 1949:31)
- (4) Alemannic, IN: *Di isch uf dā altā Eedāmburg ufpaū, di wòrschinlich*
this is on the:Dsf old:Dsf [a castle] built which:Nsf probably
i dā Edlā vo Jeeschtetā khöört hāt.
DM the:Dp noble:p from [a village] belonged has
‘This [chapel] is built on the old Edenburg which probably belonged to the nobles of Jestetten’
(Jestetten; Keller 1970:57)

I will call this construction ‘prepositional dative marking’ (PDM), the morphemes *an* or *in* ‘dative markers’ (DM). Used as dative markers, *in* and *an* are functionally equivalent; their distribution is geographically determined.

The dative marker is completely meaningless. The examples above have exactly the same meanings as their equivalents in other Upper German dialects or in Standard German, where no PDM occurs. Moreover, PDM is possible with all thematic roles that are encoded by dative DPs in Upper German.

Note that in PDM the prepositional marker is combined with dative case morphology. Thus, the dative case is expressed twice: by inflectional morphology and by the dative marker. Since in many Germanic and Romance languages the emergence of prepositional encodings of the indirect object co-occurs with the loss of dative case inflection (English *to my mother*, French *à ma mère*), one might think that the erosion of case morphology is a necessary condition for the emergence of prepositional strategies. In Upper German PDM, however, prepositional markers occur although dative case morphology is fairly intact. Thus, PDM is obviously not a compensatory strategy for eroded case morphology. The dative marker is redundant with respect to possible distinctions in case paradigms. Nevertheless, PDM provides some very interesting insights with regard to the leading questions of our Summer School here in Düsseldorf.

The present paper is organized as follows: In section 2, I will give a description of some structural and functional properties of PDM. First (section 2.1), I will show that the syntactic behaviour of the dative marker deviates from that of other Upper German prepositions in some respects. Referring to a realization-morphological approach to analytic constructions (as outlined by Spencer 2001), I will argue that dative marker insertion is due to a mapping rule that connects abstract morphosyntactic case features with overt morphological realizations. Interestingly, however, dative marker insertion is optional for most speakers. There are more and less preferred environments for PDM. The relevance of the constraining factors varies between different dialect areas. I will focus on three – geographically distinct – distributional patterns, involving morphological factors, information structure and metrical stress patterns (section 2.2). Optionality and ‘soft’ preferences must systematically be accounted for in our description. The framework of Stochastic Optimality Theory (Boersma 1997, Bresnan/Deo 2001) provides an elegant way to do that.

In section 3, I will attempt to give a historical explanation of the facts, and I will draw some conclusions with regard to evolutionary theories of grammatical change (Croft 2000, Haspelmath 1999). Croft’s and Haspelmath’s approaches differ in the role functional factors play in linguistic change: Croft claims that the selection of variants in speakers’ use is guided exclusively by social factors whereas functional factors are relevant only for the emergence of the new option as such. In Haspelmath’s approach, an option is favoured in selection due to its functionality. My data provide arguments for the latter approach.

2. Some properties of PDM

2.1 Dative marker insertion and realizational morphology

In general, the dative marker is very preposition-like. The most obvious similarity with prepositions lies in the fact that the dative marker and prepositions are in complementary distribution: the dative marker is never inserted if the dative DP is governed by another preposition:

- (5) *mit* (**i*) *de* *muetter*
 with DM the:Dsf mother
 (informant consultations)

Furthermore, many prepositions of Upper German surface as fusional morphs together with articles:

$$(6) \quad \begin{array}{l} bi \text{ 'at'} \\ \end{array} + \begin{array}{l} am \text{ 'the:Dsm'} \\ anərə \text{ 'a:Dsf'} \end{array} = \begin{array}{l} bim \text{ 'at_the:Dsm'} \\ binərə \text{ 'at_a:Dsf'} \end{array}$$

The same holds true for the dative marker as well:

$$(7) \quad \begin{array}{l} a / i \text{ (DM)} \\ \end{array} + \begin{array}{l} am \text{ 'the:Dsm'} \\ anərə \text{ 'a:Dsf'} \end{array} = \begin{array}{l} am / im \text{ 'the:Dsm'} \\ anərə / inərə \text{ 'a:Dsf'} \end{array}$$

So far, dative marker insertion seems to be triggered by an environment condition on dative DPs: the dative marker is inserted if there is no other preposition that governs the dative.

Thus, we can formulate dative marker insertion as follows:

(8) Dative marker insertion (version 1):

In Upper German, dative DPs must be governed by a preposition. If no other preposition is there, a semantically empty preposition (the dative marker) is inserted.

However, things are more complicated. There are further observations not predicted by (§§). These data suggest that PDM has also to do with the morphological realization of dative forms of determiners.

In the most typical case, prepositions governing the dative immediately precede dative forms of determiners, as in (5). However, there are mismatches between dominance and precedence. First, the sequence preposition–determiner can be interrupted, either by *nume* ‘only’ (9) or in coordination, where prepositions can have scope over two conjuncts (10):

(9) *i nume zwöi minuute*
 in only two minutes
 ‘in only two minutes (...it was ready)’

(10) *mit de Susi und de muetter*
 with:the:Dsf (name)and the:Dsf mother
 ‘with Susi and the mother’

However, a sequence dative marker–determiner cannot be interrupted by *nume* (11), nor has the dative marker scope over two conjuncts (12):

(11) *daas schicke mer (*i) nume (i) zwöi lüüt*
 that send:lp we DM only DM two persons
 ‘that we’ll send to only two persons’

(12) *i de Susi und *(i) de muetter*
 DM the:Dsf (name)and DM the:Dsf mother

Thus, the dative marker is, in contrast to other prepositions, required to precede a dative determiner immediately.

Second, the dative marker cannot be inserted when the determiner follows a preposition even if the preposition does not govern the dative of the respective DP. In Upper German (as in other non-standard varieties of German), a possessor can be expressed by an adnominal dative that precedes the 3rd person possessive pronoun:

- (13) *de muetter ires auto*
 the:Dsf mother her:Nsn car
 ‘mother’s car’

The adnominal dative can be introduced by the dative marker in PDM dialects:

- (14) *i de muetter ires auto*
 DM the:Dsf mother her:Nsn car

Surprisingly, however, the dative marker is not inserted when the entire complex is dominated by a preposition:

- (15) *mit (*i) de muetter ires auto*
 with DM the:Dsf mother her:Nsn car
 ‘with mother’s car’

This fact is surprising since it is not the preposition *mit* ‘with’ that governs the dative DP *de muetter* ‘the:Dsf mother’ immediately, as it was the case in (5). There is no general principle that prohibits sequences of two prepositions (*das isch für uf Düsseldorf* ‘that is for to Düsseldorf’ is perfectly acceptable). Thus, the incompatibility illustrated in (15) is a specific property of the dative marker.

We can summarize the observations as follows: (i) If a dative DP is governed by a preposition, no dative marker occurs. (ii) If a dative DP is not governed by a preposition, the dative marker is inserted only if the dative form of a determiner is not immediately preceded by a preposition. Here, interestingly, the only relevant criterion is the pure linear neighbourhood of word forms, without respect to structural configuration. As a consequence, the dative marker occurs only immediately before dative forms of determiners.

Since PDM is also a matter of possible environments of morphological dative forms of determiners, I suggest that PDM is part of the morphological realization of the dative case. In the following, I refer to a realization-morphological approach to analytic constructions proposed by Spencer (2001). This is a model of imperfect correspondence between feature content and means of expression. Spencer proposes a distinction between syntactic features (s-features) and morphological features (m-features). M-features specify only the shape of a word form, but they are inherently meaningless, i.e., they contain no information on the grammatical function of the respective form; this would be expressed by s-features. The morphological realization of an s-feature is due to a mapping rule that connects that feature with a morphological feature. In the most trivial cases, there is a one-to-one correspondence between s-features and m-features (if, for instance, one s-feature MOOD COND corresponds to an m-feature [mood:cond]). Frequently, however, one m-feature is mapped onto several s-features (see below), or it doesn’t correspond to any s-feature at all (this is, for instance, the case in inflectional class specifications).

Spencer’s approach is of particular interest for our purposes since it includes analytic constructions (as exponents of s-features) as well. Spencer derives his conclusions from Slavic auxiliary-participle constructions. In Russian, for instance, past tense is realized by a morphological verb form called ‘1-participle’:

- (16) *ja uš-l-a*
 I left-L-fs
 ‘I (f.) left.’
 (Spencer 2001:298)

One might think that the l-participle is a tense form. It occurs, however, also in conditional mood which is untensed in Russian:

- (17) *ja by uš-l-a*
 I cond left-l-fs
 'I (f.) would leave/have left.'
 (Spencer 2001:298)

Conditional mood is realized by a conditional marker *by* plus the l-participle. Spencer concludes that the l-participle neither 'means' past tense nor conditional mood by itself. Rather, s-features ('meanings') are mapped onto expressions made up from word forms that are specified by (inherently meaningless) m-features. Thus, the s-feature MOOD COND is linked to *by* plus a verb form specified by the m-feature [Vform:l-part], and TENSE PAST is linked to a verb form specified by the m-feature [Vform:l-part] alone. In Serbo-Croat, TENSE PAST is linked to an auxiliary form specified as [Tense:pres, Pers:α] plus a verb form specified as [Vform:l-part]:

- (18) *ja sam otiš-l-a*
 I AUX:PRS:1s leave-L-fs
 'I (f.) left.'
 (Spencer 2001:93)

Let us now apply this approach to PDM. In a first attempt, we link the s-feature CASE DAT of determiners with forms that are specified by the m-feature [case:dat] and preceded by the dative marker:

- (19) s-CASE DAT → *DM* [m-case:dat]

This rule, however, allows ungrammatical **mit i de muetter*, i.e. a sequence of preposition and dative marker. But if we accept that the dative marker is a prepositional dummy, the rule has only to specify that the morphological realization of CASE DAT requires the presence of a preposition; '(P)' means here an environment restriction on dative forms of determiners:

- (20) s-CASE DAT → (P) [m-case:dat]

In this notation it is not specified whether the preposition governs or simply precedes the dative form. In both cases, the environment restriction is satisfied. If there is no governing preposition nor a preposition immediately preceding the dative form, the environment restriction is not satisfied and thus the dative marker is inserted. In consequence, the dative marker is a semantically empty prepositional satellite of dative forms. Hence, we are able to revise our preliminary version of dative marker insertion (8):

- (21) Dative marker insertion (version 2):

Dative forms of determiners must co-occur with prepositions, i.e. they must either be governed or preceded by a preposition (or both). If they do not co-occur with a preposition, a semantically empty preposition (the dative marker) is inserted immediately before the determiner form.

Further evidence that a realizational-morphological account is right is given by the fact that in many dialects PDM is sensitive to different classes of determiners. Only in some PDM-dialects every determiner can be combined with the dative marker (for instance, in Central Switzerland or in Central Alsace). But in other dialects, PDM occurs only with articles, or with articles and demonstrative pronouns, etc.

2.2 Optionality and distributional asymmetries

2.2.1 *Optionality exists*

The properties of the dative marker as described above are valid over the entire PDM area. However, only in a small subset of the PDM dialects the dative marker is inserted consequently, i.e., unconstrained (this is the case in some regions in Central Switzerland, namely the Muotathal valley and the Lucerne district). In most dialects, however, PDM is optional. Furthermore, the factors favouring or constraining dative marker insertion vary across the different dialect areas.

For some speakers, PDM and bare datives are in completely free variation, i.e., no favouring or constraining factors can be identified: under all circumstances, the dative marker can be inserted or not, without any grammaticality contrasts or, at least, preferential asymmetries. Nevertheless, in most dialects PDM is preferred over bare datives (only) under certain conditions. The factors I will discuss in 2.2.2 define preference directions, that is, they define the circumstances under which the dative marker is more likely to appear. Only for some speakers the preference directions lead to grammaticality contrasts: they reject bare datives in a ‘PDM-environment’ (or they reject PDM in a ‘non-PDM-environment’, respectively). Therefore, we can predict that if there is an asymmetry in the occurrence of PDM vs. bare datives – with regard to frequency, preference, or grammaticality –, it follows the way the preference direction defines.

There are two ways to deal with optionality. Either we trace back the co-presence of two options to the co-presence of two grammars in the speaker’s mind (the ‘mixed-grammar’-approach). Or we take optionality for serious and write a grammar within which the co-presence of two options is systematically accounted for. There are (at least) three arguments for the latter approach:

The first argument comes from dialect geography. Suppose that in a dialect D two options, a1 and a2, coexist. a1 is attested in the neighbouring dialects in the North (N), a2 in the South (S). One might think that the properties of D are a mix between N and S. However, a1 vs. a2 is not the only optionality pattern of D. In D two options coexist for another variable, too, b1 and b2. Besides D, b1 is attested in the West (W), b2 in the East (E). Thus, with respect to b1 vs. b2 D is a mix between W and E. Consequently, with only two variable features our hypothetical dialect D is already made up from four distinct grammars. In reality, the number of variable features in one dialect is much higher than only two. Consequently, the number of co-present grammars in one speaker’s mind would increase dramatically (cf. Bresnan/Deo 2001:39).

Second, optionality patterns play a crucial role in language change. Innovations in grammar pass necessarily through phases of variation. Thus, optionality is a completely expected fact; see below.

The third argument follows from the observation that even in optionality areas dative marker insertion is more or less preferred under certain circumstances. If optionality were nothing more than the result of a random mixture of grammars we would not expect systematic distributional asymmetries; see below.

- (26) *ich ha die büecher (*i) allne ggä* x x
 I have these books DM all:Dp given x x x x x
 'I gave these books to all of them' *büe cher (i) all ne*

As mentioned above, these distributional patterns describe preference directions. Concrete informants' judgements vary. Some informants, for instance, reject PDM totally in an unpreferred environment, whereas for others it is acceptable but 'it sounds bad'. For some informants, PDM is required in preferred environments, for others it is acceptable only here, but not obligatory, etc. Nevertheless, informants' judgments are always in accordance with the preference directions as such: if there is a preferential asymmetry between bare datives and PDM, it follows the predicted direction.

2.2.3 Optionality and stochastic evaluation

I'd like close this section with a short remark on Optimality Theory and Stochastic Optimality Theory. Due to the lack of time, I cannot go into further details here.

Obviously, dative marker insertion is in conflict with preferences penalising it. Furthermore, we have seen that in different dialect areas different factors are relevant for PDM. These factors have in common that they are syntax-external (i.e., morphological, discourse-pragmatic, or phonological). I think that Optimality Theory provides a promising means of description for these facts. OT explicitly describes conflicting preferences. The ranking of OT constraints is language-specific, or, in our case, dialect-specific. Thus, diatopic contrasts in the relevance of the factors governing PDM can be expressed by re-ranking of constraints. Moreover, constraints belonging to different levels of grammar (phonology, morphology, etc.) can be involved in the evaluation of possible output candidates.

I will first exemplify 'classical' constraint ranking with the third distributional pattern above, i.e. the influence of metrical stress patterns. After that, I will propose that our data favour a constraint interaction approach with stochastic evaluation. The relevant sequence of constraints is:

- (27) NOCLASH: Avoid stress clashes. (Kager 1999:165)
 >>
 DON'T PROJECT: Avoid phrase structure. (Bresnan 2000:351)

NOCLASH favours dative marker insertion, if this avoids a sequence of two stressed syllables. However, DON'T PROJECT militates against dative marker insertion. Therefore, in a dialect where PDM occurs between otherwise clashing syllables, NOCLASH must be ranked above DON'T PROJECT:

(28)

			NOCLASH	DON'T PROJECT
	x	x		*
☞	x	x		
	σ_1	DM	σ_2	
	x	x	*!	
	x	x		
	σ_1	σ_2		

If, in contrast, NOCLASH is irrelevant (in cases where dative marker insertion does not contribute anything to stress clash avoidance), the candidate without dative marker wins, due to DON'T PROJECT. Thus, in this dialect PDM is grammatical if otherwise two stressed syllables would clash, and it is ungrammatical in all other cases. There is only one problem

with that: such a dialect does not exist. It is true that in some dialects PDM depends on metrical stress patterns, but there is only an asymmetry in preference between the two options PDM vs. bare dative, not a grammaticality contrast. In other words: What we need is a means of description which allows both candidates to win, but which expresses at the same time that PDM is preferred (not obligatory) under certain circumstances (i.e., under the metrical conditions introduced above).

The first we might do is leaving NOCLASH and DON'T PROJECT unranked:

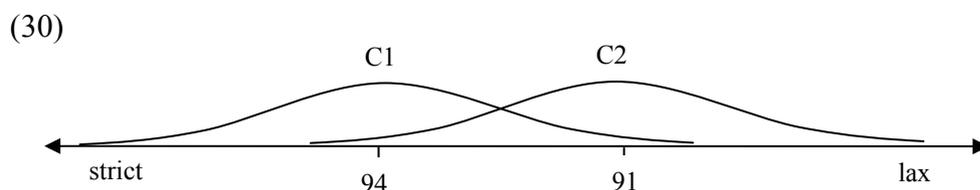
(29)

				NOCLASH	DON'T PROJECT
	x		x		*
☞	x	x	x		
	σ_1	<i>DM</i>	σ_2		
	x	x		*	
☞	x	x			
	σ_1	σ_2			

Here, none of the two violations is fatal, such that both candidates are equally optimal. The preference direction, however, cannot be expressed.

If ranking the constraints and leaving the constraints unranked both fail, then the solution is perhaps ranking the constraints only slightly. If the constraints are categorically ranked, there is no overlap between them. If they are not ranked at all, there is a total overlap between them. What we need, however, is a partial overlap. We need to express that a candidate satisfying the higher ranked constraint is more preferred, but it does not rank out its competitor. Stochastic Optimality Theory exactly expresses that (Boersma 1997, Bresnan/Deo 2001).

In classical OT, constraints are ranked on a categorical scale. Suppose that positions on the scale are defined by numeric values. If the position of C1 is, say, 94, it is ranked above C2 on position 91. In classical OT, at each candidate evaluation C1 dominates C2. Stochastic OT, however, assumes that there is a certain amount of stochastic perturbation at each evaluation. That is, the position of a constraint on the numeric scale is variable. Thus, 94 is not the real position of C1 at each evaluation, but it is the peak of a probability curve of C1's concrete evaluations: the position of C1 is more probable the closer it is to 94. If the distance between two constraints is rather small, we expect single evaluations where their respective ranking is reversed. The closer the constraints are on the scale, the more the probability of reversed rankings increases. If the position of C1 and C2 is the same, C1 dominates C2 in 50% of the concrete evaluations:



(Adopted from Bresnan/Deo 2001:8)

Thus, to come back to our example, if C1 corresponds to NOCLASH and C2 to DON'T PROJECT, it is more probable that in a concrete evaluation NOCLASH dominates DON'T PROJECT, but the reverse ranking occurs as well:

More probable (more preferred): cf. (28) above.

>>

Less probable (less preferred):

(31)

				DON'T PROJECT	NOCLASH
	x		x	*!	
	x	x	x		
	σ_1	<i>DM</i>	σ_2		
☞		x	x		*
		x	x		
		σ_1	σ_2		

Stochastic evaluation thus gives us the right prediction: PDM occurs if the dative marker avoids a stress clash, but bare datives occur in this environment as well, though less frequently.

More generally speaking, in dialect geography we are frequently faced with patterns that can accurately be accounted for by Stochastic Optimality Theory: usually, the transition from one to another geographic variable is not abrupt, but the prominence of one option gradually decreases in the areal continuum, from obligatory via preferred and acceptable to inexistent. I insist that at each point in the areal continuum we find a possible grammar which is acquired by the local speakers. Therefore, we need a means of description that is able to simulate these grammars.

3. How structure acquires function: Consequences for a theory of language change

Why does variation within dialects and between dialects matter? Because of its crucial role in language change. I claim that the analysis of dialect variation contributes something to our understanding of how language change acts. In the following I will focus on evolutionary explanations of language change, because they give an appropriate account for the role variation plays in language change. In an evolutionary approach, variation and optionality is something entirely expectable. Furthermore, in different evolutionary approaches functional factors play a different role. The microscopic study of dialect variation provides an evaluatory tool for these approaches.

3.1 Reanalysis is the source of the dative marker

First, I will briefly sketch a hypothesis where the dative marker comes from. At the first glance, one might think that the dative marker *in* or *an* is grammaticalized from the Upper German true prepositions *in* 'in, into' or *an* 'at, beside of', respectively. In fact, there are cases in Upper German as well as in Modern Standard German where a recipient dative object is interchangeable with a directional PP introduced by *an*, compare (Standard German):

- (32) a. *er hat das Buch seinem Freund geschickt*
 he has the book his:Dsm friend sent
- b. *er hat das Buch an seinen Freund geschickt*
 he has the book to his:Asm friend sent

‘he sent the book to his friend’

However, in a directional PP *an* governs accusative, not dative. If PDM were grammaticalized from directional PPs, we would expect *an* plus accusative, not *an* plus dative. However, in PDM the dative marker is always combined with dative, never with accusative.

Instead, the dative marker emerged due to a historical accident, namely the phonological development and reanalysis of article forms which made it possible to generalize the (highly frequent) post-prepositional occurrence of the dative case over all instances. I suppose that three conditions are relevant for the emergence of PDM.

First condition: The Middle High German definite article dative singular masculine *dēme* lost its initial dental when it was cliticized to prepositions:

- (33) *obem* 1280, *uff(f)em* 1270, *am* 1277, *im* 1258, *underm* 1276, *us(s)em* 1409, *vom* 1277, *vorem* 1280, *hinderm* 1403, *bim* 1280, *zem* 1245
 (Idiotikon XIII, 1191f).

In Upper German, however, the form without initial dental has been generalized over all occurrences. Thus, except in extremely conservative dialects, the article form is *əm* (with some variation in the vocalism) over the entire Upper German dialect area, that is, even in dialects where no PDM exists. Nübling (1992:201) explains this process by the fact that in Upper German the dative is much more frequently governed by a preposition than it occurs as a dative object (without preposition).

Second condition: There is a whole paradigm of fusional morphs where prepositions and article forms are merged in all Upper German dialects:

(34) Alemannic

	+ <i>əm</i> ‘the:Dsm’	+ <i>əmənə</i> ‘a:Dsm’	+ <i>ənərə</i> ‘a:Dsf’
<i>a(n)</i> ‘at’	<i>am</i>	<i>əmənə</i>	<i>ənərə</i>
<i>i(n)</i> ‘in’	<i>im</i>	<i>imənə</i>	<i>inərə</i>
<i>bi</i> ‘by, next to’	<i>bim</i>	<i>bimənə</i>	<i>binərə</i>

Depending on sentence stress and speech rate, the phonetic quality of the first vowel varies between [ɑ] and [ə] (in the amalgams of *a(n)* ‘at’), or between [ɪ] and [ə] (in the amalgams of *i(n)* ‘in’), respectively. Thus, the amalgams *am* <at_the:Dsm>, *im* <in_the:Dsm> and the bare article form *əm* <the:Dsm> are potentially homophonous. Again, this situation is common to almost all Upper German dialects, even those where PDM does not occur. Compare (Zurich German, without PDM):

- (35) *hüt əm fōifi*
 today at_the five’
 ‘today at five’

In some Bavarian dialects, *əm*, *am* and *im* fell together phonetically in *in* (cf. WBÖ I:209a).

The third condition is common to all Upper German dialects, too: the most frequent and thus prototypical occurrence of the dative case is post-prepositional. Following Nübling (1992:221), more than 90% of the occurrences of datives are governed by prepositions.

In this context, an innovation took place in a subset of the Upper German dialects. These dialects reanalysed the forms of the dative articles +DEF:Dsm, –DEF:Dsm and –DEF:Dsf as prepositional amalgams. Thus, they interpreted the – not very frequent – bare article forms as belonging to the highly frequent pattern of fusional morphs <preposition_article>. Once the (etymologically) bare article forms are reanalysed as prepositional amalgams, this pattern can analogically be extended to other determiners where no fusional morphs with prepositions exist. Consequently, the dative in Upper German PDM dialects is realized ‘in analogy’ to post-prepositional datives. The result of the reanalysis is a new morpheme, our dative marker.

3.2 Two approaches towards linguistic evolution

I will focus on two evolutionary approaches, Haspelmath (1999) and Croft (2000). Let us start with what these approaches have in common.

Evolutionists claim that the guiding mechanisms in language change are variation and selection. If a new pattern comes into play, it does not abruptly replace the inherited pattern, but it is in competition with it. The emergence of a new variant thus just adds a new option. It is possible, then, that one of the available options is selected more and more frequently in speakers’ activity. If selection frequency further increases, the pattern can become obligatory, i.e., its competitors possibly die out. As a consequence, language change goes necessarily through phases of variation. In this context, optionality is nothing exotic, but it is completely expected: it is a necessary stage in the development of an innovation.

However, Although Croft’s and Haspelmath’s approaches are compatible in principle they differ in an aspect that is very central for the subject of this Summer School. When the variant has emerged, selection in speakers’ use can take place. But if there are two (equivalent) options, what principles, then, drive selection? Why do speakers choose one variant more frequently?

Croft says that selection is exclusively guided by social factors. Croft distinguishes ‘innovation’ (actuation) from ‘propagation’ (transmission). ‘Innovation’ means the change as such. Propagation is a matter of selection: it means the spread of the change across a speakers’ population. Given a range of alternative variants, it is a matter of social significance only whether an option is selected by the speakers or not. Functional factors, however, play a part only in innovation but not in selection (propagation):

- (36)
- a. «Functional factors – the phonetic and conceptual factors appealed to by functionalist linguists – are responsible only for innovation, and social factors provide a selection mechanism for propagation» (Croft 2000:38).
 - b. «The empirical evidence indicates that linguistic selection is governed largely if not exclusively by social factors that have little or nothing to do with functional adaptiveness for communication» (Croft 2000:39).
 - c. «The propagation of a linguistic variant is a selection process: one variant is selected over another one. Selection is a process that takes place among interactors, however. A speaker does not produce one linguistic variant in preference to another in an utterance because of its linguistic properties. A speaker identifies herself with a community or a subset of a community and that causes her to produce one linguistic variant in preference to another» (Croft 2000:178).

Thus, the circumstances under which a speaker selects a variant are – from a purely linguistic point of view – totally arbitrary. We might expect that occurrence asymmetries in optionality patterns correlate with sociolinguistic factors, and not with functional factors. As a consequence, optional variants are sociolinguistic variables.

In contrast, in Haspelmath's approach functional principles are relevant for selection, too. Haspelmath claims that via variation and selection a causal link can be made between functional adaptivity and linguistic structure. This presupposes, of course, that such a link is needed, i.e. that structural properties of language reflect functional requirements in communication (for some critical remarks, see below, section 3.4). Haspelmath argues that the constraints Optimality Theory assumes are intuitively plausible and seem to be functionally motivated, although in standard OT there is no place for an explanation of this fact. Haspelmath's explanation is evolutionary: structural patterns are 'good' for the speakers' needs because such patterns are favoured in selection. Thus, selection is driven by functional principles. Consequently, an option can become obligatory at least in some contexts. (Remark: One might say that language change is seen as variation between competing candidates and evaluation of these candidates by universally available preferences. The similarities between this approach and grammatical description in an optimality-theoretic fashion are obvious.) If Haspelmath's assumptions are valid we expect that occurrence asymmetries in optionality patterns correlate with functional factors.

I think that the question whether Haspelmath's or Croft's approach is more appropriate is not only matter of philosophical speculation. It is simply an empirical matter, too. Both approaches make predictions about variation. Why don't we just look at how variation actually works?

3.3 And what about PDM?

Let us now have a look at PDM from an evolutionary perspective. If this perspective is right, we expect that the first result of an innovation is the coexistence of two equivalent options. Obviously, this is really the case: Through reanalysis of article forms and analogical extension, PDM emerged as an additional and thus optional realization pattern for dative objects. As expected, we find PDM and bare datives in completely free variation in many dialects. Some dialects, however, go one step further, and this step is crucial for our analysis. Here, the variation between the two options is not completely random: there are more and less preferred environments for PDM. In other words, the initially purely formal contrast PDM vs. bare dative evolves functional contours. In different dialects, different systematic arrangements between the options PDM vs. bare dative appear independently from each other: in some dialects, the dative marker is inserted predominantly if the dative case is morphologically underspecified (South Bavarian), in others the dative marker is used due to stress clash avoidance (Lucerne) or if the dative DP is focused (widespread in Switzerland, for instance Schaffhausen). The asymmetries can easily be motivated by more general functional patterns: morphological distinctness, information structure or preferred stress patterns. But it is certainly odd to say that the dative marker emerged 'in order to' mark focused dative objects, for instance. Instead, PDM is 'just there', and the functional asymmetries PDM acquires in different dialect areas are secondary effects of speakers' activity. Therefore, they are due to selection. In other words, selection in speakers' use may result in distributional asymmetries between two almost equivalent options: speakers choose one option more frequently under certain systematic circumstances, not only under certain social

circumstances. These functional asymmetries in different dialects are independent from who is speaking to whom.

I conclude from these facts that functional factors *are* relevant for selection. Evidence for this claim can be found in two observations: first, the appearance of functional contours between PDM and bare datives are optional, i.e., they not necessarily need to arise. If functional factors were relevant only for the innovation properly but not for selection, we would not expect that. And second, the systematic arrangements developed in different dialects are independent from each other. If they were not secondary results from selection, we would expect a more uniform pattern cross-dialectally.

Therefore, I assume – in contrast to Croft – that the change proper is something very mechanic, it is merely structural variation. Functional factors come into play only after, as unintended results of selection in speakers' activity. (Remark: Note that nothing is said here for or against social factors. Of course social factors are crucial for selection, too, as we know since decades. My claim, however, is that there are selection asymmetries that cannot be explained without reference to functional factors. In other words: Not all variation patterns can be explained as sociolinguistically significant variables.)

As a consequence, it is necessary to introduce a terminological adjustment in order to distinguish the variation patterns I have in mind from those traditionally referred to by historical linguists and sociolinguists. Correctly, evolutionary theory predicts that there is a phase of variation between the first occurrence of an innovation and the completion of the change. Therefore, language change is necessarily gradual. I propose the term *emergence* for the genesis of a grammatical pattern as such. Emergence just adds new options. In contrast, *implementation* involves the status this option takes in the respective system of grammar (a matter of *valeur linguistique*). An innovation first *emerges* (via purely structural variation) as an additional option and can then pass through phases of *implementation*. *Implementation* leads to specific, but possibly cross-dialectally distinct systematic arrangements between the two options. Thus, it is a matter of *emergence* that PDM as a technical pattern has become available in large parts of Upper German. But it is a matter of *implementation*, what status the innovation acquires in the respective grammars: whether it remains completely optional, whether PDM is preferred under certain circumstances, whether it becomes obligatory in certain cases, etc. Only if we divide a change in *emergence* vs. *implementation* we can speak about the stepwise integration and of an innovation in a – gradually changing – system of grammar.

The much more common pair *actuation* vs. *transmission* is not sufficient for these purposes. Transmission refers to the status of an innovation in a speakers' population. The change itself, however, is not gradual, although it gradually spreads across society.

3.4 Towards obligatorization

In some dialects of Central Switzerland (Muotathal valley and parts of Lucerne), PDM has become obligatory. Thus, the PDM option has been chosen by speakers increasingly often such that it totally replaced its competitor, bare datives. This development is relatively recent (Haas 1989:570). In consequence, dative marker insertion is not linked to any syntax-external factor at all; its only function is to guarantee morphosyntactic well-formedness. Of course, one might motivate even this process in terms of speakers' needs: first, if speakers have the choice, they tend to choose more explicit patterns in order to attract attention, call it

extravagance (Hapelmath 1999a) or expressivity (Lehmann 1985). Second, it is iconic that similar categories (datives) are mapped onto similar realizations (PDM). But these factors only explain why an option is chosen increasingly frequently during implementation. They are, however, completely irrelevant for the result of that process as a synchronic state. A speaker of the Muotathal dialect does not always focus datives, nor does she choose PDM due to expressivity. Instead, PDM is automatic, i.e. purely structurally triggered. Neither the occurrence nor the design of the PDM pattern is determined by any extra-syntactic factor here. For that reason, a functional motivation for dative marker insertion is not available, and, more importantly, it is not necessary. Functionalist approaches to morphosyntax sometimes fail in not recognizing that (synchronically) many regularities are due to well-formedness only, even if their diachronic origin can be explained by functional motivations (i.e., functional factors favouring one option during implementation).

What do we have to conclude from these facts? Due to structural variation, various options are allowed by morphosyntax. These options can be arranged according to functionally motivated preferences. However, in the course of further implementation, a pattern *can* be fully automatized (though needn't to), i.e. it is purely structurally triggered and merely due to morphosyntactic well-formedness. Consequently, its use need not necessarily to be motivated by the factors guiding variation during implementation. Therefore, two conclusions have to be drawn: first, functional motivations are most relevant for the periods of the gradual diachronic implementation of a pattern (i.e. the phases of variation) but not necessarily so for the results of the implementation process. And second, one has to accept that functional requirements and structural requirements both exist and play their part. Both are explanatory, but they explain different things. It is something different whether you model the design of grammar or whether you ask for the function of that design. Design is not determined by function because functional contours appear as a secondary effect of selection asymmetries in speakers' use, and because the influence of functional factors can collapse in the course of further implementation. However, since functional and structural requirements explain different things, they are ultimately compatible.

4. References

- Bäbler, Heinrich. 1949. *Glarner Sprachschuel*. Glarus: Verlag der Erziehungsdirektion.
- Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Institute of Phonetic Sciences, University of Amsterdam, Proceedings* 21:43-58.
- Bresnan, Joan. 2000. Optimal syntax. In: Dekkers, Joost *et al.*: *Optimality theory: Phonology, syntax and acquisition*. Oxford: University Press. 334-385.
- Bresnan, Joan / Deo, Ashwini. 2001. Grammatical constraints on variation: 'Be' in the *Survey of English Dialects* and (Stochastic) Optimality Theory. Ms. Stanford University. <http://www-lfg.stanford.edu/bresnan/be-find.pdf>
- Croft, William. 2000. *Explaining language change. An evolutionary approach*. Harlow (etc.): Longman.
- Haas, Walter. 1989. Appendix to: Fischer, Ludwig: *Luzerndeutsche Grammatik. Ein Wegweiser zur guten Mundart. Mit einem Anhang zur Neuausgabe 1989 von Walter Haas*. Hitzkirch: Comenius. (=Reprint of 1st edition 1960).
- Haspelmath, Martin. 1999. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18:180-205.

- Haspelmath, Martin. 1999a. Why is Grammaticalization irreversible? *Linguistics* 37:1043-1068.
- Idiotikon. *Schweizerisches Idiotikon. Wörterbuch der schweizerdeutschen Sprache*. Frauenfeld: Huber. 1881ff.
- Kager, René. 1999. *Optimality theory*. Cambridge: University Press.
- Keller, Rudolf Ernst. 1970. Jestetten Kr. Waldshut. In *Phonai* 7, 7-89. Tübingen: Niemeyer.
- Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20: 303-318.
- Mair, Raimund. 1938. *Die Flexion der Stubaiier Mundart, historisch dargestellt*. Doctoral dissertation (ms.) University of Innsbruck.
- Mayr, Ernst. 1997. *This is biology. The science of the living world*. Cambridge (Mass.) / London: Belknap Press of Harvard University Press.
- Nübling, Damaris. 1992. *Klitika im Deutschen. Schriftsprache, Umgangssprache, alemannische Dialekte*. Doctoral dissertation Albrecht-Ludwigs-Universität Freiburg. Tübingen: Narr.
- Rader, Wolfram. 1966. *Sprachbiologische Untersuchungen an der Mundart von Feldkirchen in Kärnten*. Doctoral dissertation (ms.) University of Vienna.
- Schöpf, Johann B. 1866. *Tirolisches Idiotikon*. Innsbruck: Wagner.
- Seiler, Guido. Forthcoming. *Präpositionale Dativmarkierung im Oberdeutschen*. Doctoral dissertation University of Zurich, German department.
- Spencer, Andrew. 2001. The paradigm-based model of morphosyntax. *Transactions of the Philological Society* 99-2:279-313.
- Stranzinger, Oswald. 1951. *Laut- und Flexionsverhältnisse der Innsbrucker Umgangssprache*. Doctoral dissertation (ms.) University of Innsbruck.
- Ströbl, Alex. 1970. *Grundlinien einer formalistischen Syntax der Mundart von Malching*. Doctoral dissertation Ludwig-Maximilians-Universität München. München: Fink.
- WBÖ: *Wörterbuch der bairischen Mundarten Österreichs*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 1963ff.

5. Abbreviations

A	accusative
D	dative
DM	dative marker
DP	determiner phrase
f	feminine
m	masculine
p	plural
PDM	prepositional dative marking
PP	prepositional phrase
s	singular
UG	Upper German ;-)