The internal structure of person portmanteaus

1 Allgemeine Angaben

1.1 Antragsteller

Name:	Dr. Jochen Trommer
Dienstadresse:	Universität Leipzig, Institut für Linguistik
	Beethovenstr. 15, 04107 Leipzig

Name:	Dr. Michael Cysouw
Dienstadresse:	Max Planck Institut für evolutionäre Anthropologie
	Deutscher Platz 6, 04103 Leipzig

1.2 Zusammenfassung

Person portmanteaus are markers which bundle the expression of person for subject and object of a verbal predicate into a single morphological unit. Although person portmanteaus implement a radically simple alternative to more familiar strategies of cross-referencing verbal arguments, they are cross-linguistically highly marked, and tend to be restricted to very specific combinations of subject and object. Person portmanteaus may therefore be regarded as a type of argument encoding sui generis as well as a litmus test for general properties of argument encoding in head-marking languages. An important though somewhat paradox property of person portmanteaus is that they are not necessarily unanalyzable. Many such markers can be understood synchronically and/or diachronically as non-portmanteaus, either because they are formed by concatenation from two different person formatives, or because they are identical in form to non-portmanteau person markers.

Based on these observations, this project addresses the following closely related questions: 1) Are there systematic restrictions on the location of portmanteaus in paradigms? 2) To which extent are portmanteaus analyzable? 3) What is the formal representation of portmanteaus? 4) What is the extent and distribution of portmanteau microvariation within groups of closely related languages? 5) How do portmanteaus develop historically (and how do they disappear)? To answer these questions, we combine methods from typology, formal morphology, and historical reconstruction. Apart from constructing a cross-linguistic typological database for person portmanteaus to extract general properties, we focus on the behavior of portmanteaus in three small-scale language families (Nilotic, Iroquoian, and Uralic). For these languages we will provide in-depth formal analyses of the relevant person portmanteaus, investigate the range of microvariation in the families, and attempt to reconstruct the diachronic development of person portmanteaus on the background of the overall pronominal cross-referencing systems (based on the available descriptive literature). From the innovative combination of different methodologies to microvariation data, we expect a better understanding of person portmanteaus, and more generally of syncretism in cross-referencing systems for verbal arguments.

2   Stand der Forschung, eigene Vorarbeiten

2.1   Stand der Forschung

Portmanteaus and Suppletion   Portmanteau morphemes have been defined in the structuralist literature as morphological units "which belong simultaneously to two (or, theoretically, more) morphemes, and have simultaneously the meanings of both" (Hockett, 1954:333). A classical example of a portmanteau morpheme is English **went** under the assumption that this is morphologically a single morpheme which otherwise functions for morphosyntax as a (hypothetical) combination of the two morphemes **go** and **-ed** would. The notion portmanteau must be carefully distinguished from the related concept of "suppletion" which originally covers cases of lexically unrelated stem alternants (Osthoff, 1899), but has been extended in the later literature to various degrees, up to the assertion that morphemes (or constructions) are suppletive if their semantic correlation is "maximally regular, while their formal correlation is maximally irregular" (Mel'čuk, 1994:358). Under this view, portmanteau morphemes are a subclass of suppletion (which Mel'čuk calls "radical megamorph suppletion"). For example, the irregular German comparative form **bess-er**, 'better' (cf. **gut**, 'good') is a case of suppletion (suppletive allomorphy), but not of a portmanteau since **bess-er** contains the regular comparative suffix **-er**, and **bes(s)** also occurs in the superlative form **bes-te**, 'best'. See Veselinova (2003, 2006) for an extensive summary of the literature on suppletion and http://www.surrey.ac.uk/LIS/SMG/Suppletion_BIB/WebBibliography.htm for a bibliography.

Whereas we are not primarily concerned with general aspects of suppletion, there are three influential hypotheses on the distribution of suppletion which are of crucial importance for the theory of portmanteaus. *First,* suppletion is often described as a maximally marked type of morphology (Dressler, 1985). *Second,* it has been claimed that suppletion is especially likely in frequent lexemes and constructions (Haspelmath, 2008; Ronnerberger-Sibold, 1980), and *third*, it has been argued that suppletion is particularly frequent in the "Nahbereich" (Dressler, 1985; Bittner, 1988, 1996), a conceptual space which is close to the speaker (e.g. there is gender suppletion for domestic animals such as **ox** and **cow**, but not for wild animals such as **lion** and **lion-ess**, cf. Veselinova, 2003:43). The notion "Nahbereich", however, is notoriously difficult to operationalize (see Fertig, 1998 for criticism).

Portmanteaus in General   Although portmanteau morphemes are standardly treated in morphology textbooks, and many morphological descriptions and analyses invoke them implicitly, there has been little systematic theoretical research on the general formal and typological properties of portmanteaus. A notable exception is Stump (2001) who argues, based on Paradigm Function Morphology, a framework where morphological spellout happens in disjunctively ordered position classes, that there are specific morphemes (or morphemic rules) which block the morphemes of two otherwise independent position classes by one single morphological formative. To capture this, he introduces a specific type of "portmanteau position class" which is a formal implementation of the structuralist morphological notion of a portmanteau. What makes Stump's discussion rather unique in the literature is that it seems to be one of the few examples in which a detailed empirical argument is made to justify the existence of a specific class of portmanteau markers which has well-defined properties distinguishing it from "simple" markers (cf. also Trommer, 2003b).

A radically opposite stance to Stump's view is taken in Trommer (2003d) where it is argued for Hungarian that the evidence for portmanteau agreement in the language is actually elusive, and all alleged examples of agreement portmanteaus must be reanalyzed as non-portmanteaus once a wider array of data is taken into account. Based on this case study, Trommer hypothe-

sizes more generally that the descriptive notion of portmanteau might not have any theoretical significance and should be eliminated from the theory of morphology. Similar discussions can be found for various concrete examples. Thus Bruening (2001) analyzes the so-called "theme-markers" in Algonquian as being portmanteaus, whereas Brittain (2001) argues that they are "simple" object agreement markers. Since both positions are based on data from single languages, it is fair to say that they do not provide definite results, but rather offer a reference frame for a detailed theoretical and empirical investigation of portmanteaus, which is still to be accomplished.

Person Portmanteaus   Person portmanteaus are markers which encode features of two arguments of a verbal predicate (or appear to do so). They provide an ideal testing ground for the study of portmanteaus for at least three reasons: *First*, (verbal) cross-referencing morphology is a wide-spread feature in the languages of the world. This is an important precondition for a typological study. *Second*, cross-referencing morphology in general is a relatively well-understood area of typological and theoretical work on morphology (cf. e.g. Cysouw, 2003; Baerman, 2005; Stump, 2001 and papers in Müller and Trommer, 2006). In other words, there are well-developed tools for the identification, description and analysis of cross-referencing morphology which provides an ideal background to isolate the special features of person portmanteaus.

Third, person portmanteaus largely obviate what might be called the "chunking problem", which might be illustrated with nominal gender and number marking in Romance, where some languages mark nominal gender and number by different affixes (e.g. Spanish **manzan-a-s**, apple-FEM-PL, 'apples') while other languages express both categories systematically by single affixes (e.g. Italian **mel-e**, apple-FEM:PL, 'apples'). Analytically there are two possibilities: If gender and number are considered as "natural chunks" (similar to the co-occurrence of person and number in verbal agreement), Italian represents the standard case with the simple marker **-e**, while Spanish shows marked splitting of this unit into two markers ("fission" in the sense of Halle and Marantz, 1993 and Trommer, 2003c). On the other hand, gender and number might be taken as the standard units for morphological exponence resulting in an analysis where Spanish uses two simple markers while Italian has portmanteau affixes for gender and number. Obviously the criteria for chunking are non-trivial in general. However, there is broad agreement that in the cross-referencing systems of most languages person and number features for a specific argument are to be treated as one "chunk" (cf. Cysouw, 2003; Daniel, 2005 for typological arguments for this view).

Hypotheses from the literature   The most detailed discussion of formal aspects of person portmanteaus is provided in Wunderlich (2006). In a systematic survey of cross-linguistically attested argument linking types, Wunderlich identifies the "portmanteau type" (which is a term covering roughly what we call person portmanteaus here) as a special type of argument linking besides strategies such as case marking and fixed word order. He argues that the portmanteau type has two characteristic properties:

- **Person portmanteaus are the maximally marked linking type** because they require a large and hence uneconomical inventory of cross-referencing markers. This explains why person portmanteaus are rarer than simple markers and why there are very few languages where person portmanteaus are used systematically (Wunderlich cites only one possible such example, the Tanoan language Kiowa). This statement is in agreement with the Natural Morphology literature which characterizes portmanteaus (and suppletion more generally) as maximally marked morphological structure.

- **Person portmanteaus are especially suited for combinations of non-third person arguments**. Citing two languages where portmanteaus are used in $1 \to 2$ forms, Wunderlich states "To express the combination I $\to$ you is a special communicative task, so it does not wonder that a portmanteau morpheme adapted to this special task is found in several languages." (p.4). A similar observation was already made in earlier work by Heath (1991, 1997, 1998) who has investigated person portmanteaus from a cross-linguistic perspective. His claim is that person portmanteaus are particularly prominent in first person/second person situations (either subject/object or object/subject). Interestingly, this generalization seems to be incompatible with the claim that portmanteaus are especially likely in frequent constructions: At least for English, transitive clauses with 1st and 2nd person arguments only are much less frequent than transitive clauses with at least one 3rd person argument (Bresnan et al., 2001; Helmbrecht, 2004:462). However, the generalization that first/second person combinations are preferred for portmanteau marking might be compatible with an appropriate version of the "Nahbereich" hypothesis, which is actually the explanation favored by Heath.

An alternative hypothesis on the distribution of person portmanteauss is put forth in Trommer (2003a,b):

- **Person portmanteaus are especially suited for inverse combinations of arguments**. In other words, person portmanteaus favor constellations where the object is higher than the object for a specific prominence scale (e.g. 1st $\succ$ 2nd $\succ$ 3rd person). Trommer (2003a,b) bases this claim on an approach to direct/inverse (direction) markers, which analyses these as person portmanteaus. Since typologically overt direction markers for direct configurations (predications where the subject is higher for the given scale than the object) imply overt direction markers for inverse configurations in a given language (Aissen, 1999), this claim is well-established for languages which are traditionally treated as direct/inverse systems (and set apart from person portmanteaus), but it is largely an open empirical question whether there is independent evidence for it from languages with other types of portmanteaus. A possible example is a person portmanteau in Quechua for which Lakämper and Wunderlich (1998) argue that it implements a strategy to avoid violations of a general constraint penalizing separate expression of object features on the verb if the object is higher on a prominence scale than the subject. While the analysis of Lakämper and Wunderlich (1998) differs substantially from the one provided by Trommer, both converge in the prediction that person portmanteaus show a special affinity to inverse configurations. Crucially, this hypothesis links person portmanteaus in an interesting way to the general problem of scale effects in argument encoding. Finally, it is important to note that it is not logically incompatible with the claimed preference for person portmanteaus in combinations of local persons since the latter case might be regarded as a special case of inverseness under the assumption that local persons might establish different scale rankings in different languages (Silverstein, 1976) or different subsystems of the same language (Zúñiga, 2002; Macaulay, 2007).

Apart from the claims on the asymmetric distribution of person portmanteaus cross-linguistically and in single languages, Heath in the above-cited works makes two interesting and interrelated claims (cf. also Helmbrecht, 2004:462-463):

- **Portmanteauicity is gradual:** There are many cases of portmanteaus which still allow some level of analysis, but resist a fully compositional/transparent morphological treatment. In other words, markers can be gradually more or less portmanteau-like, or

have different degrees of "portmanteauicity". This insight plays an important role in our classification of portmanteaus.

- **Portmanteaus develop historically out of non-portmanteaus** either by fusion of adjacent subject and object markers or by zero marking of one of the arguments.

The work by Wunderlich, Lakämper, and Heath offers striking observations on the systematicity of person portmanteaus in single languages, and interesting hypotheses which invite further investigation. However the empirical evidence they provide is very limited and neither involves detailed analyses of single languages nor a systematic evaluation of cross-linguistic evidence, or a historical reconstruction of portmanteau genesis. Although Heath corroborates his hypothesis on the preference of person portmanteaus for local persons with examples from many Australian and Amerindian languages, any preference of portmanteaus for local persons is not systematically evaluated with respect to the frequency of portmanteaus within other person combinations. In general, it is striking that typologically the issue of person portmanteaus does not yet seem to have been investigated in depth (cf. the virtual absence of this topic in Siewierska, 2004).

2.2 Eigene Vorarbeiten

Dr. Jochen Trommer Trommer has worked on the morphological analysis of person marking in various individual languages, recurrently discussing the problems raised by portmanteaus for the analysis of these languages. Trommer (2003b:95-96;ch.6.3) argues that person portmanteaus can be distinguished from non-portmanteaus since they have different inherent affix ordering preferences and behave differently with respect to blocking other affixes. Trommer (2003b:ch.7.4.1) identifies Arizona Tiwa as a language with pervasive person portmanteaus, where virtually all transitive person combinations are expressed by portmanteaus. Another argument in Trommer (2003b:ch.7.3) is that direct-inverse markers are coextensive with person portmanteaus.

Trommer (2003d) shows that apparent person portmanteaus in Hungarian transitive verb agreement must be reanalyzed partially by splitting the portmanteaus into two different markers, and partially by equating them with simple markers accompanied by zero agreement. Further, Trommer (2006) identifies the class of "ambiguous exponents": person-number affixes which occur as portmanteaus in specific contexts, and as simple markers in other contexts. These are analyzed as simple markers with highly underspecified context restrictions. Later, Trommer (2007) advocates the more radical hypothesis that all person portmanteaus cross-linguistically might be reanalyzed as simple markers (or combinations of simple markers). Finally, Trommer (2008) develops a late-insertion model of morphological spellout which, in contrast to Distributed Morphology allows portmanteau affixes and gives a detailed analysis of portmanteau marking in Turkana accounting for the syncretism of the person portmanteau use of the prefix **k-** with its use in intransitive 1pl forms.

Dr. Michael Cysouw Cysouw has worked extensively on the world-wide typology of various aspects of person marking. Cysouw (2003) focusses explicitly on monovalent marking of person to reach a workable delimitation for the cross-linguistic comparison. This extensive survey will be used as a basis for the survey of portmanteaus in this project. More specialized articles have focussed on the marking of clusivity, in which first and second person are combined into one argument (Cysouw, 2005a,c,d). Clusivity is in a sense the monovalent

counterpart of first/second bivalent person portmanteaus. Cysouw (2006, 2008) discusses aspect of the typology of linear order for bound person marking of both subject and object. There are rather strict typological preferences for the marking of both subject and object cross-references to a verb, probably influencing the development of portmanteau markers. Further, Cysouw (2005b) discusses some special effects in the typological development of person clitics into cross-referencing affixes, which is relevant for the development of the linear order of cross-referencing affixes on the verb. Finally, Cysouw (1997, 1998) are in-depth analyses of person markers in groups of closely related languages in which portmanteau markers play a central role.

3 Ziele und Arbeitsprogramm

3.1 Ziele

The main goal of the project is to answer the following five theoretical questions. Some subsidiary methodological questions are discussed afterwards.

3.1.1 Theoretical Goals

Question 1: Are there systematic restrictions on the location of portmanteaus in paradigms?
Based on our typological study, we will verify claims about the distribution of portmanteaus within paradigms. The central question to be answered in this part of the project is in which combinations of subject/object marking do we typically find portmanteaus, or portmanteau-like elements? As a point of departure we take the hypotheses by Heath, Trommer and Wunderlich that there is a special connection between local person (1st and 2nd person) and more generally the misalignment of subject and object with person scales. A central goal is to avoid the systematic methodological problems of earlier work. Especially we want to:

- take into account person portmanteau from different language families and linguistic macro-areas

- develop well-defined criteria for measuring frequency of portmanteaus in specific paradigmatic slots. It is obviously not sufficient to show that there are "many" portmanteaus for combinations of local persons to establish a preference for these combinations. The frequency of portmanteaus in local contexts must be systematically weighed against their frequency in other contexts.

- isolate independent factors which might account for the absence of portmanteaus in specific contexts. For example, in a language which systematically excludes overt marking of 3rd person objects, a portmanteau for a $1 \rightarrow 3$ form is independently excluded while a $1 \rightarrow 2$ portmanteau should be in principle possible (cf. Quechua). That a language of this type has no person portmanteaus in the first context but has them in the latter one, would then not reveal anything specific about person portmanteaus, but rather about the special connection of zero-marking and 3rd person.

- avoid analytical bias. For example, markers which are special to transitive configurations are often called direct-inverse affixes if they mark non-local configurations. Such transitive markers should of course also be considered as possible portmanteaus.

- consistently use a gradient notion of portmanteauicity, instead of an all-or-nothing approach to being analyzable or not. This aspect is detailed in the next research question.

6

Question 2: To which extent are portmanteaus analyzable? The central goal of the project is to find out to which degree portmanteaus are analyzable into smaller pieces or can be reduced as a whole to simple markers. Based on similar ideas in Heath (1991) at least the following subtypes of portmanteaus can be distinguished:

- **Simplex Portmanteaus:** these are portmanteaus which are actually simplex morphemes, though they appear to be portmanteau because they express a composite meaning. For example, in Maricopa, a Yuman language from North America, the transitive person prefix for the combination 2nd person subject/3rd person object is **m-** (Gordon, 1986). Though this might look like a clear case of a portmanteau (having a complex meaning, but just a simplex morphological expression), it seems better to analyze this prefix as a simplex form, because the intransitive prefix for a 2nd person subject is also **m-**. From a morphological perspective, one might simply analyze the transitive prefix as having a zero-element, or being underspecified for the marking of the object.

- **Transparent Composite Portmanteaus:** these are (claimed) portmanteaus which turn out on closer inspection to be actually a transparent concatenation of simple morphemes.

- **Blurred Composite Portmanteaus:** these are portmanteaus which are concatenations of simple morphemes obscured by (morpho)phonology (e.g. German *zu + dem → zum*). The "degree of blurring" is relatively low if the involved phonology is general and only partially morphologically conditioned, and higher if it is sensitive to morphological factors or specific items.

- **Cranberry Portmanteaus:** Portmanteaus which consist of one portion which corresponds to a marker used elsewhere in the paradigm for the person category expressed by the portmanteau and a second part which occurs only in the portmanteau.

- **Opaque Portmanteaus:** portmanteaus which cannot even partially be decomposed into smaller formatives which are independently motivated person markers.

The first two types are what one might call descriptively "fake portmanteaus", whereas the last type might be called "true portmanteau". The third and fourth types of portmanteaus are intermediate stages which defy a clear dichotomy between portmanteaus and simple markers.

This situation leaves us two analytical options: to develop finer-grained descriptive scales of portmanteauicity which capture differences as well as the transitionality between single types, or to explore formal representations of different portmanteau types based on independently motivated theoretical assumptions on morphological structure. We will pursue both approaches. *First*, we will introduce a quantitative scale of portmanteauicity rating individual morphemes as to the extent to which they can meaningfully be analyzed within the structure of the language. The question as to the existence of portmanteaus then becomes a question how the cross-linguistic distribution of this scale looks like, and where on the scale individual morphemes from specific languages are found. Morphemes on the high extreme of such a scale then are the more typical portmanteaus. Following observations in Heath (1991, 1998) we hypothesize that typical portmanteaus are typologically very rare. *Second*, we will explore formal representations for the different kinds of portmanteaus. Our approach to these formal representation will be detailed in the next research question.

Question 3: What is the formal representation of portmanteaus? Formal approaches to morphology can be separated into formalisms which provide explicit means for representing portmanteaus as a special type of marker (e.g. as markers corresponding to two position

classes as in Stump, 2001, or as markers specifying two distinct feature structures, cf. Wunderlich, 1996; Lakämper and Wunderlich, 1998; Trommer, 2003b), and formalisms which implicitly exclude such representations because every morphological formative corresponds to a morphosyntactic unit of a fixed size which is usually assumed to be a syntactic head as in Distributed Morphology (Halle and Marantz, 1993).

The working hypothesis of the project is that at the level of formal analysis there are no portmanteaus, i.e. portmanteaus are either formatives corresponding to a single unit of morphosyntactic features or to a combination of different formatives, each of which is a single unit of morphosyntactic features, where the factors which conspire to produce portmanteaus at the level of descriptive analysis are "low-level" mechanisms such as contextual allomorphy, (possibly opaque) morphophonological processes, and zero morphemes. This has the advantage that the gradual degrees of portmanteauicity in synchrony and diachrony can be modeled, synchronically as the accumulation of a variety of formal factors instead of stipulating a categorical distinction, and, diachronically, as a clear-defined switch by which (combinations of) non-portmanteaus are reanalyzed as portmanteaus.

The accumulation model incorporates the observation made by (Heath, 1991:81-82) that single portmanteauicity factors such as non-transparent allomorphic conditioning tend to occur in the same environments as true portmanteaus. It also provides a possible explanation for the markedness-degree of portmanteaus. If single accumulating factors such as zero-marking and allomorphy are independently marked (and relatively rare) their combined effect should turn out to be even more marked. Finally, in this model, true portmanteaus in the sense of our descriptive classification will require maximally non-transparent analyses (e.g. as the combination of two special allomorphs which occur only in the specific "portmanteau context").

The validity of these arguments of course depend on the empirical results of the project. If these imply that a substantial subpart of portmanteaus is opaque, this provides straightforward justification for a "Stumpian" special-mechanism approach to portmanteaus. But even though this is an empirical possibility which cannot be discarded a priori, the reductionist hypothesis is methodologically the best stance to examine its validity. A further important challenge for the reductionist approach is to account for properties which seem to set portmanteaus systematically aside from simple markers. Thus Trommer (2003b) claims that affix ordering properties of portmanteaus differ in a principled way from corresponding simple markers.

Of special impact for the formal representation of person portmanteaus are the class of markers which are identified in Trommer (2006) as "ambiguous exponents", i.e. agreement affixes which in some contexts express features of two arguments, and in other paradigm cells features of a single argument. Thus the Belhare affix **ka-** marks 1st person inclusive (hence [+1+2]) objects in some forms, and a combination of a 1st person (exclusive) subject and a 2nd person objects in other forms. What unifies these uses is that in both cases **ka-** expresses the person features +1 and +2, but in one case in a simple mapping and in the second case in a portmanteau manner. While Trommer (2006) argues that ambiguous exponents are simple markers exhibiting a special type of contextual allomorphy, the theoretical impact of this phenomenon can only be evaluated on the basis of more relevant data and a systematic evaluation of possible alternative analyses.

Question 4: What is the extent and distribution of portmanteau microvariation within groups of closely related languages? It is an often noted problem of morphological analysis that specific analyses are underdetermined by the data since, in contrast to syntax, the amount of data is relatively limited (for a given language normally finite). Further, in contrast to phonology it is much less obvious what are the relevant external conditions of language use

that restrict possible interpretations of the morphological data. This problem is also virulent for the typological and formal research on portmanteaus since determining the portmanteau status of a given marker requires crucially to determine the structure of the morphological paradigms in which it functions. To broaden our data base we will study microvariation data in two senses.

First, we will study the distribution of portmanteaus in small-scale language families which differ minimally in their morphosyntax and their use of portmanteaus. Models for this type of study (although only partially or not concerned with portmanteaus) are Noyer's (1992:chapter 1) case study of the prefix conjugation in Afro-Asiatic, the dissertation of Lakämper (2000) on person and number inflection in Quechuan dialects, and Cysouw's investigation of "cognate paradigms" throughout the world's languages (Cysouw, 2003:chapters 8 and 9). As these studies show, the set of alternative analyses for a specific inflectional system is cut down dramatically as soon as closely related dialects or languages are taken into account. An application to portmanteaus is discussed in Trommer (2008) who argues, based on the detailed study of microvariation in Nilotic in Cysouw (1998), that the appearance of Turkana **k-** in transitive and intransitive plural contexts cannot be accidental homophony because it is also found in closely related varieties of Nilotic even though the morpho-phonological details differ substantially between the languages.

Second, complex morphological systems often display a notably different distributions of markers in different contexts. Thus the Turkana person portmanteau marker **k-** extends to 2nd person object forms and to intransitive contexts in most moods and tenses, but is restricted to transitive forms with a 1st person object in the so-called "subsecutive" mode (Dimmendaal, 1983:175). Likewise, some agreement markers in Algonquian (see also Wunderlich, 2003 on Cree) behave as portmanteaus (as they are restricted to very specific combinations of subject and object persons) in the verbal paradigm used in main clauses, but function as straightforward object agreement markers in the paradigm used in subordinate clauses. Obviously a full understanding of portmanteaus like this depends on taking into account all relevant subparadigms which imposes substantial constraints on possible analyses (Trommer, 2007).

Question 5: How do portmanteaus develop historically (and how do they disappear)?

A further empirical area which promises a deeper understanding of portmanteaus is their historical development and the historical development of paradigms containing portmanteaus. We also expect that the diachronic genesis of portmanteaus will contribute to a more adequate synchronic analysis. Any proposal for the reconstruction and diachronic development of paradigms is strongly connected to their synchronic analysis. The combination of diachronic arguments with the structural scrutiny of microvariation offers much stronger restrictions on possible analyses, because the synchronic analysis will have to function at every stage in the diachronic development proposed. In this sense, there is a bidirectional control between the (synchronic) microvariation and the (diachronic) reconstructed variants.

Concretely, there are two different diachronic aspects that are of importance for our project: *First*, portmanteaus (apparently) develop from erstwhile independent "simple" markers, either by some sort of fusion between two originally separate elements, or by extension of the possible references of an originally monovalent element into a multivalent element. Portmanteaus reanalyzed from the combination of different simple markers are likely to retain some of the properties of the earlier stage, which is important to accurately describe the synchronic situation. Moreover if a marker assumes more portmanteau-like properties at a given historical stage we expect that this change is conditioned by the structure of pronominal marking in the language (e.g. by adjacency of the original "simple" markers) and other grammatical factors

(e.g. the restriction of pronominal marking to a single affix-"slot", as argued by Wunderlich, 2006:6 for Kiowa).

Second, the paradigmatic structure of closely related languages also changes (Cysouw, 2003:chapters 8 and 9). Likewise, the range of possible referential values of a portmanteau is changeable, indicating another pathway for portmanteaus to become synchronically opaque. For example, in Guaraní, the combination first plural subject/second singular object is marked by a prefix **ro-**, which is very probably historically derived from the simplex **ro-** used for intransitive first person plural. Less obviously, this same prefix is also used for the combination first singular subject/second singular object. It seems unlikely that this combination is directly developed out of the intransitive first person plural. More likely, the development had two steps: 1PL → 1PL/2SG → 1SG/2SG. The second step is an example of paradigmatic extension of a portmanteau.

Finally, it is important to realize for any diachronic reconstruction that the existence of portmanteaus in one branch of a genealogical unit, and the absence in another branch, has of course two different possible historical scenarios as an explanation. It might be that the branch with portmanteaus has developed them, or the other branch has lost them. There is often a tendency to analyze portmanteaus as complexity developed out of earlier "simple" markers (as found in closely related languages), though it is just as likely that a system without portmanteaus has originated by losing the portmanteaus. There is probably no simple and foolproof way to decide between these two possibilities, but in most cases there will be clues available in the kind of microvariation pointing to one or the other scenarios.

3.1.2 Methodological Goals

Besides the concrete goals described above, there are various subsidiary goals, maybe better seen as meta-goals. These more general methodological objectives underly all of the concrete work to be carried out in the course of this project.

Combining diachrony and synchrony for morphological analysis   The main methodological goal of the project is to develop a synergy effect for morphological theory by applying historical reconstruction of microvariation and formal analysis of languages in their attested as well as in their reconstructed shape in parallel in a way that possible reconstruction restricts the possible analysis inside a chosen formalism and the choice of the formalism itself, and vice versa independently motivated aspects of the formalism restrict possible reconstructions and possible analyses of microvariation. In the literature there are very few approaches of this type. Lumsden (1987) does an interesting abstract morphological analysis of different historical stages of Old and Middle English, but obviously he treats only a well-documented transition in a single language. Noyer (1992) compares different stages of Afro-Asiatic, but does not really reconstruct: all data are based on documented sources and no intermediary stages of historical development are accounted for.

Teasing apart truly morphological aspects of paradigmatic change from external factors
Although phonological processes play a central role in diachronic developments, those processes alone would result in a random distribution of portmanteaus throughout the paradigm, which seems to be counterfactual. To improve on this basis, our analysis will include diachronic processes of morphological nature that describe preferred changes of the paradigmatic structure of person markers.

De-discretize language description and comparison The received approach to portmanteaus, like most categories in linguistics, consists of a strictly discrete distinction between portmanteaus and non-portmanteaus. Because of the assumption that such clear-cut distinctions exist in language, much of the intermediate variation is considered a nuisance, or a problem to the theory. By starting off from the assumption that the category of portmanteauicity is not a yes/no question, we circumvent this problem. However, the resulting challenge is to build a coherent morphological theory that can deal with such gradient notions

3.2 Arbeitsprogramm

3.2.1 Typological Survey

The cross-linguistic data we use will consist of a qualitative sample of about 30-50 languages which are claimed to have person portmanteaus. Thus the major design criterion for the sample is to include a maximal variety of person portmanteau types differing with respect to their analyzability (cf. section 3.1.1), their position in the overall cross-referencing paradigms of the respective languages, and the features they encode. Creation and evaluation of the sample will depart from the data in Cysouw (2003) and will be happen in close cooperation with Anna Siewierska.

3.2.2 Microvariation

We will investigate microvariation in Nilotic, Uralic and Iroquoian.

Nilotic The Nilotic languages are a sub-branch of Nilo-Saharan, one of the four large stocks in Africa, spoken in parts of Sudan, Kenya, Ethiopia, Uganda, Zaire and Tanzania. According to a widely accepted classification, Nilotic comprises three branches (Grimes and Grimes, 1996):

(1) **Branches of Nilotic**:

Eastern:		Bari: Bari
		Lotuxo-Maa: Lotuho, Maasai
		Teso-Turkana: Teso, Turkana, Karamojong, Toposa
Southern:		Kalenjin: Päkot, Nandi, Markweta
		Tatoga: Datooga, Omotik
Western:	Lwoo:	Northern: Anywa, Päri, Shilluk
		Southern: Lango, Acholi
	Dinka-Nuer:	Dinka, Nuer

Person portmanteaus are found in Southern and Eastern Nilotic, most pervasively in Eastern Nilotic where (except for Bari and Lotuho) both subject and object agreement are realized by prefixes. Characteristically etymologically related person portmanteaus resemble inverse markers in their distribution, are also used in specific intransitive plural forms, and show subtle distributional differences between different languages and in different paradigms (Cysouw, 1998). All these properties are highly relevant for central aspects of the project. Southern Nilotic shows less obvious, but nevertheless intriguing cases of portmanteauicity: In Päkot, where both subject and object are crossreferenced by suffixes, subject and object markers show

11

evidence for incipient fusion interacting in a complex way with the tonal morphophonology of the language. In Nandi, which as most of the Southern Nilotic languages has subject prefixes but object suffixes, specific subject-object combinations are indicated additionally by characteristic tones on the object affixes (cf. e.g. the 2sg object affix **-in** which in the perfective bears a high tone if the subject is 1st person, but a falling tone if it is 3rd person, Creider and Creider, 1989:98). This case is especially interesting since the person portmanteau is linearly adjacent to the object marker, but not to the subject marker, and we suspect that similar phenomena exist in other Southern Nilotic languages.

Uralic   The Uralic family comprises the Ugric languages, Saamic-Fennic, Samoyedic and a number of other languages. All Uralic languages show subject agreement, but only the Ugric languages, Samoyedic and Mordvin exhibit object agreement (Abondolo, 1998):

(2)   **Branches of Uralic**

Samoyedic	Northern	Nganasan, Enets, Nenets
	Southern	Selkup, Kamas, Mator
Finno-Ugric	Ugric	Khanty, Hungarian, Mansi
	Finno-Permic	Permic: Komi, Udmurt
		Finno-Saamic: Saami, Finnic
		Mordvin
		Mari

For the project, Uralic is especially relevant since it seems to show portmanteau agreement systems in statu nascendi. Only for Hungarian, the existence of portmanteaus is widely accepted (but see Trommer, 2003d). In contrast, Samoyedic and the other languages with object agreement illustrate transparency of subject and object marking to various degrees providing possible evidence for pre-portmanteau stages of historical development.

Iroquoian   The Iroquoian languages, a family of closely related languages in Northern America, exhibit one of the world's most extensive paradigms of person prefixes, including numerous transitive portmanteaus. It is immediately obvious only from looking at these prefixes that they are related, though the details of the development of the person markers have not been worked out. The historical tree of Iroquoian (following Chafe and Foster, 1981) is shown below (ignoring the less-well documented varieties Nottoway, Susquehannock and Laurentian). To complicate matters, both Cayuga and Onondaga are claimed to have had various re-contacts after the original split, leaving tracing in their grammatical and phonological system. A first attempt to tackle the problem to analyze the person marking can be found in Cysouw (1997).

(3)   **Branches of Iroquioan**

Southern	Cherokee	
Northern	Tuscarora, Cayuga	
	Lake Iroquoian	Huron
		Inner Iroquoian: Seneca, Onondaga, Oneida, Mohawk

3.2.3   Data Collection

For all languages in the cross-linguistic sample, and the relevant languages from Uralic, Nilotic, and Iroquoian, the following data will be retrieved from published sources:

- full paradigms of person portmanteaus and related subject and object cross-referencing paradigms, including all morpho-phonological variation and alternations

- (morpho-)phonological processes relevant for the proper segmentation of the paradigms

- basic grammatical characteristics which are (potentially) relevant for the analysis of the cross-referencing system (e.g. case marking, agreement with DPs which are not arguments of the agreeing verb, word order, morpheme order, extent of pro-drop)

The first point addresses the central data relevant for the project. Collecting phonological information (point 2) is important to evaluate how far an apparent portmanteau can be disassembled into different markers (cf. the discussion in 3.1.1). The third point is crucial to identify structural factors which might be important for the understanding of possible portmanteau types. For Uralic, Nilotic, and Iroquoian we will in addition collect detailed information on the known historical development of these language families, and diachronic (morpho-)phonological changes which are relevant for disentangling phonological and morphological aspects of portmanteau and paradigm genesis.

To document the retrieved data, the project will use an ontology for paradigms which is under development in the DFG-project "Mikro- und Makrovariation: Hierarchieeffekte in Kiranti und Algisch" (TR-521-/3-1) to create an electronic database which contains the basic data (in IPA notation), the detailed phonological and syntactic information necessary for a proper analysis and possible morphological segmentations and annotations. Extraction from the sources and integration of the data in the database includes the following steps:

- Extraction of basic morphological, syntactic and phonological information

- Standardization of transcriptions

- Morphemic segmentation and identification of person portmanteaus

These steps which will require a substantial part of the practical workload in the documentation part of the project will be carried out by student assistants. The database will allow easy comparison of different analyses for the same language, but also of related morphological data in different languages, and will be published online after the completion of the project. Documenting data and segmentation together follows and extends recent practice in typological databases which also include raw data (e.g. the Surrey Agreement database, http://www.smg.surrey.ac.uk/ and the Autotype database; Bickel and Nichols, 2002). Including different possible segmentations addresses the central problem that the proper segmentation of person portmanteaus cannot be fixed a priori or in a completely theory-neutral way.

3.2.4   Diachronic Analysis

The diachronic analysis will encompass the complete person paradigms (including the portmanteaus). The diachronic development of each individual person marker will be tracked through the microvariation of the closely related languages. By combining the (assumed) historical subgrouping and the known sound changes, we will try to fit as much of the attested

variation of the person makers onto the historical tree. We expect there still to be a lot of incompatible similarities (i.e. person markers that are more similar between two languages as expected from the tree), which might either be the result of parallel innovation, or of re-contacts. One of the central supplementary arguments that we will use to resolve historical developments is the paradigmatic structure of the person markers: in each stage of the language the structure of the paradigm will have to be modeled by the formal analysis. And each change in the paradigmatic structure should not just be as minimal as possible (diachronic parsimony), but also induce minimal changes in the formal model (formal parsimony).

3.2.5 Formal Analysis

Formal analyses will be carried out for all relevant languages in Nilotic, Uralic, and a substantial subset of the crosslinguistic sample. The formalism which is used is Minimalist Distributed Morphology (MDM, Trommer, 1999, 2003c, 2003d), a version of Distributed Morphology which explicitly rejects special types of representations for portmanteaus, and is therefore well-suited to implement the reductionist approach to person portmanteaus developed in section 3.1.1. Since MDM allows only a single morphological operation (vocabulary insertion), it substantially reduces the set of possible analyses for specific data, which is of importance to keep an evaluation of different segmentation possibilities for person portmanteaus (resulting in different formal analyses themselves) manageable. The formalism for vocabulary insertion developed in Trommer (2008) which minimally differs from the MDM formalism by allowing specific portmanteau items is a convenient reference point for comparing the developed analyses to alternative non-reductionist approaches to portmanteaus.

References

Abondolo, D. (1998). Introduction. In Abondolo, D., editor, *The Uralic Languages*, pages 358–386. London and New York: Routledge.

Aissen, J. (1999). Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory*, 17(4):673–711.

Baerman, M. (2005). Verbal person/number syncretism. In Haspelmath, M., Dryer, M., Gil, D., and Comrie, B., editors, *World atlas of language structures*, pages 122–125. Oxford: Oxford University Press.

Bickel, B. and Nichols, J. (2002). Autotypologizing databases and their use in fieldwork. In *Proceedings of the LREC Workshop on Resources and Tools in Field Linguistics*, pages 25–26.

Bittner, A. (1988). Reguläre Irregularitäten. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 41:416–425.

Bittner, A. (1996). *Starke 'schwache' Verben—schwache 'starke' Verben*, volume 51 of *Studien zur deutschen Grammatik*. Tübingen: Stauffenburg.

Bresnan, J., Dingare, S., and Manning, C. D. (2001). Soft constraints mirror hard constraints: Voice and person in english and Lummi. In Butt, M. and King, T. H., editors, *Proceedings of the LFG 01 Conference, The University of Hong Kong, Hong Kong*, pages 13–32. CSLI Publications.

Brittain, J. (2001). *The Morphosyntax of the Algonquian Conjunct Verb*. New York: Garland Press.

Bruening, B. (2001). *Syntax at the Edge: Cross-Clausal Phenomena and the Syntax of Passamaquoddy*. PhD thesis, MIT.

Chafe, W. L. and Foster, M. K. (1981). Prehistoric divergences and recontacts between cayuga, seneca, and the other northern iroquoian languages. *International Journal of American Linguistics*, pages 121–142.

Creider, C. A. and Creider, J. T. (1989). *A Grammar of Nandi*. Hamburg: Helmut Buske Verlag.

Cysouw, M. (1997). A history of Iroquoian gender marking. Ms. available online at http://www.eva.mpg.de/ cysouw/pdf/cysouwIROQGEN.pdf.

Cysouw, M. (1998). Pronoun descriptions: Nilotic. Ms. available online at http://email.eva.mpg.de/ cysouw/pdf/cysouwNILOTIC.pdf.

Cysouw, M. (2003). *The Paradigmatic Structure of Person Marking*. Oxford University Press, Oxford.

Cysouw, M. (2005a). Inclusive/exclusive forms of 'we'. In Haspelmath, M., Dryer, M. S., Gil, D., and Comrie, B., editors, *World Atlas of Language Structures*, pages 162–169. Oxford University Press, Oxford.

Cysouw, M. (2005b). Morphology in the wrong place: A survey of preposed enclitics. In Dressler, W. U., Kastovsky, D., Pfeiffer, O. E., and Rainer, F., editors, *Morphology and its Demarcations*, pages 17–37. Benjamins, Amsterdam.

Cysouw, M. (2005c). Syncretisms involving clusivity. In Filimonova, E., editor, *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*, pages 73–111. Benjamins, Amsterdam.

Cysouw, M. (2005d). A typology of honorific uses of clusivity. In Filimonova, E., editor, *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*, pages 213–230. Benjamins, Amsterdam.

Cysouw, M. (2006). The asymmetry of affixation. In Gärtner, H. M., Beck, S., Eckard, R., Musan, R., and Stiebels, B., editors, *Puzzels for Krifka*. ZAS, Berlin.

Cysouw, M. (2008). Linear order as a predictor of word order regularities. *Advances in Complex Systems*.

Daniel, M. (2005). Plurality in independent personal pronouns. In Haspelmath, M., Dryer, M. S., Gil, D., and Comrie, B., editors, *World Atlas of Language Structures*, pages 146–149. Oxford University Press, Oxford.

Dimmendaal, G. J. (1983). *The Turkana Language*. Dordrecht: Foris.

Dressler, W. (1985). Suppletion in word-fomation. In Fisiak, J., editor, *Historical Morphology*, pages 97–112. The Hague: Mouton.

Fertig, D. (1998). Suppletion, natural morphology, diagrammaticity. *Linguistics*, 36:1065 – 1091.

Gordon, L. (1986). *Maricopa Morphology and Syntax*. University of California Press, Berkeley.

Grimes, B. F. and Grimes, J. E. (1996). Ethnologue language family index to the thirteenth edition of the ethnologue.

Halle, M. and Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In Hale, K. and Keyser, S. J., editors, *The View from Building 20*, pages 111–176. Cambridge MA: MIT Press.

Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19(1):1–33.

Heath, J. (1991). Pragmatic disguise in pronominal-affix paradigms. In Plank, F., editor, *Economy of Inflection*, pages 75–89. Berlin: Mouton de Gruyter.

Heath, J. (1997). Lost wax: Aprupt replacement of key morphemes in australian agreement complexes. *Diachronica*, 14(2):197–232.

Heath, J. (1998). Pragmatic skewing in 1-2 pronominal combination in native american languages. *IJAL*, 64(2):83–104.

Helmbrecht, J. (2004). *Personal Pronouns: Form, Function, and Grammaticalization*. PhD thesis, Habilitationsschrift, Erfurt.

Hockett, C. F. (1954). Problems of morphemic analysis. *Language*, 23(4):321–343.

Lakämper, R. (2000). *Plural- und Objektmarkierung in Quechua*. PhD thesis, University of Düsseldorf.

Lakämper, R. and Wunderlich, D. (1998). Person marking in Quechua — a constraint-based minimalist analysis. *Lingua*, 105:113–148.

Lumsden, J. (1987). *Syntactic Features: Parametric Variation in the History of English*. PhD thesis, MIT. available under: http://dspace.mit.edu/handle/1721.1/14702.

Macaulay, M. (2007). On prominence hierarchies: Evidence from Algonquian. Ms. University of Wisconsin-Madison.

Mel'čuk, I. (1994). Suppletion: Toward a logical analysis of the concept. *Studies in Language*, 18:339–410.

Müller, G. and Trommer, J., editors (2006). *Subanalysis of Argument Encoding in Distributed Morphology*, volume 84 of *Linguistische Arbeits Berichte*. Institut für Linguistik: Universität Leipzig.

Noyer, R. R. (1992). *Features, Positions and Affixes in Autonomous Morphological Structure*. PhD thesis, MIT. available under: http://www.ai.mit.edu/projects/dm/theses/.

Osthoff, H. (1899). *Vom Suppletivwesen der indogermanischen Sprachen*. Heidelberg, Kommissionsverlag von Alfred Wolff.

Ronnerberger-Sibold, E. (1980). *Sprachverwendung, Sprachsystem, Ökonomie und Wandel*, volume 87 of *Linguistische Arbeiten*. Tübingen: Niemeyer.

Siewierska, A. (2004). *Person*. Cambridge University Press, Cambridge.

Silverstein, M. (1976). Hierarchy of features and ergativity. In Dixon, R., editor, *Grammatical Categories in Australian Languages*, pages 112–171. Canberra: Australian Institute of Aboriginal Studies.

Stump, G. T. (2001). *Inflectional Morphology*. Cambridge: Cambridge University Press.

Trommer, J. (1999). Morphology consuming syntax' resources: Generation and parsing in a minimalist version of Distributed Morphology. In *Proceedings of the ESSLI Workshop on Resource Logics and Minimalist Grammars*.

Trommer, J. (2003a). Direction marking as agreement. In Junghanns, U. and Szucsich, L., editors, *Syntactic Structures and Morphological Information*, pages 317–341. Berlin and New York: Mouton de Gruyter.

Trommer, J. (2003b). *Distributed Optimality*. PhD thesis, University of Potsdam.

Trommer, J. (2003c). Feature (non-)insertion in a minimalist approach to spellout. In *Proceedings of 39*, pages 469–480.

Trommer, J. (2003d). Hungarian has no portmanteau agreement. In Siptár, P. and Pinón, C., editors, *Approaches to Hungarian*, volume 9, pages 283–302. Akadémiai Kiadó.

Trommer, J. (2006). Plural insertion is constructed plural. In Müller, G. and Trommer, J., editors, *Subanalysis of Argument Encoding in Distributed Morphology*, volume 84 of *Linguistische Arbeits Berichte*, pages 197–228. Institut für Linguistik: Universität Leipzig.

Trommer, J. (2007). On portmanteau agreement. Talk given at the Leipzig-Harvard Workshop on Morphology and Argument Encoding, Cambridge MA, September 2007; Slides available under http://www.uni-leipzig.de/~jtrommer/Harvard/trommer.pdf.

Trommer, J. (2008). ø-agreement in Turkana. In Richards, M. and Malchukov, A., editors, *Scales*, volume 86 of *Linguistische Arbeits Berichte*, pages 0–0. Institut für Linguistik: Universität Leipzig.

Veselinova, L. (2003). *Suppletion in verb paradigms: Bits and pieces of a puzzle*. PhD thesis, Department of Linguistics, Stockholm University.

Veselinova, L. (2006). *Suppletion in verb paradigms: bits and pieces of the puzzle*. John Benjamins, Amsterdam; Philadelphia.

Wunderlich, D. (1996). A minimalist model of inflectional morphology. In Wilder, C., Bierwisch, M., and Gärtner, H.-M., editors, *The role of economy principles in linguistic theory*, pages 267–298. Berlin: Akademie-Verlag.

Wunderlich, D. (2003). A minimalist view on inflectional paradigms: the expression of person and number in subjects and objects. MS., University of Düsseldorf.

Wunderlich, D. (2006). Towards a structural typology of verb classes. In Wunderlich, D., editor, *Advances in the theory of the lexicon*, pages 57–166. Berlin: Mouton de Gruyter.

Zúñiga, F. (2002). *Inverse systems in indigenous languages of the Americas*. PhD thesis, University of Zürich.