REGINE ECKHARDT
KLAUS VON HEUSINGER
CHRISTOPH SCHWARZE
(Editors)

Trends in Linguistics

Words in Time

Diachronic Semantics
from Different Points of View
Introduction: Historical linguistics as a transdisciplinary field of research

Regine Eckardt, Klaus von Heusinger, and Christoph Schwarze

1. Words in time

Word meanings are constantly changing, and the driving forces of meaning change are varied and diverse. Few semantic changes are determined by purely language-internal factors; the majority can only be understood by taking various kinds of external influences into account. Usually, linguistic investigations are concerned with language-internal processes, while investigations of language-external historical developments tend to ignore linguistic considerations. It is evident, however, that diachronic semantics can only be fully understood if we base the study of semantic change on a well-developed explicit theory of meaning, i.e., synchronic semantics, one that includes a proper place for lexical semantics of both content words and function words. From this point we can start our quest for principles of semantic change, drawing on expert knowledge about the world, on the domains of reference of words, as well as on the social and cultural environment of speakers who used the words.

When modern linguistic research in Germany slowly evolved as a new independent field of scholarly investigation in the early nineteenth century, Wilhelm von Humboldt still lived up to the ideal of the universally interested and informed investigator of language. Having received his academic education at a time when philosophy, history, philology, and even natural sciences were still hosted in the overarching Faculty of Arts, he not only envisaged Wissenschaft as one undivided field of intellectual effort but specifically viewed the investigation of language itself as a contribution to the larger project of general anthropology; the investigation of human nature in its cultural, psychological, sociological, historical and philosophical aspects.
Since Humboldt’s time, considerable scientific progress has led to specialization in the humanities, and former subfields have emancipated themselves into self-reliant disciplines with their own methods, perspectives, aims and theories. The investigation of meaning and meaning change was continued under the label semantics as an advanced specialized subfield of linguistic research. While the impressive results in linguistics over the last decades fully justify these specializations, diachronic semantics remains a challenging topic of investigation. Interdisciplinary efforts are necessitated by the very nature of the object of study and, at the start of the 21st century, the time seems ripe to reunite the various branches of the humanities, each advanced and matured since the days of Humboldt, in the common exploration of the nature of language variation and change. A first step must be to strengthen the awareness of scholars in different fields of the fact that many truths can be told about the common object of investigation – natural language and languages in their cultural, social, historical and psychological settings. Facts and laws of language that are focused on by one discipline can be irregular neglectables of the next discipline, and single scholars, absorbed by their own field of expertise, are in danger to base their work on a one-sided idealization of a multi-faceted reality.

The present volume offers a collection of studies in meaning change from linguistic as well as nonlinguistic perspectives conducted by scholars in linguistics, philology, philosophy, sociology, anthropology and history. The collection comprises, and hence allows to compare, in-depth investigations of language change from different perspectives and scientific paradigms. Each of the contributions demonstrates the methods and techniques of its own core field whilst showing the fruitful mixing of perspectives of more than one discipline. In bridging the gap between disciplines, we want to strengthen sensitivity for the many dimensions of language as a social, cultural, cognitive, formal and historical object.

2. Theories of language change and meaning change

2.1. The modern classics

Wilhelm von Humboldt is commonly named as the prototype of a universal scholar. He approached the investigation of language in the course of his wider project of seeking to elucidate the general nature of human character. He was therefore interested in languages as reflecting the culture of their
Historical linguistics as a transdisciplinary field

speech communities, and his perspective on the diachronic investigation of individual languages can be likened to the efforts of the biographer: to understand which events and circumstances had influenced and shaped the language as it presented itself at some given point in time. Under this perspective, facts about political, social and cultural history as well as philosophical and psychological insights (at the time hardly even separated) were much more than mere anecdotal annotations to grammar.

The Schlegel-Humboldt debate about the development of Romance languages from their Latin origin can serve to illustrate this fact (see specifically Trabant 1990: 128f. and Plank 1991, 2002 for a broader overview). The point of conflict was, in simple terms, how two synthetic languages – Latin and Germanic – could give rise to descendant Romance languages that were more analytic in character. Humboldt focused in particular on the Romance future tense and article system to witness the shift from more synthetic to more analytic. We must remember that at the time, synthetic languages – most saliently Ancient Greek – were viewed as the climax of language development. The emergence of analytic elements was hence viewed as language decay. Humboldt and Schlegel were concerned with the creative potential of language. To put it in terms of the biological metaphor of language creation: Can two parental languages as such give rise to an inferior offspring, or does ‘decay’ arise through nonlinguistic external influence?

Schlegel held the view that the development in question occurred as a mixing of two languages when Latin-speaking communities imitated the language of the politically dominant class of Germanic intruders. Humboldt (1827–29 [1907]: 292) opposed this and argued that the development must have resulted from social and political conditions internal to the Latin-speaking society:

Ihre sie charakterisierende Eigenthümlichkeit gieng nicht aus der Mischung Germanischer und Römischer Rede und Sprache hervor, sondern aus der durch die siegreiche Einwanderung fremder Stämme bewirkten Zerstörung des politischen Bestandes, der darauf folgenden Zerrüttung des ganzen Culturzustandes, und der diese Katastrophen begleitenden Verderbniss der Sprache (Humboldt).

1 “Their (= Romance languages) characteristic features did not emerge from the mixing of Germanic and Roman speech and language, but from the victorious immigration of foreign peoples and the destruction of the political system caused
Humboldt (1827–1829 [1907]: 292) hence seems to maintain the positivist view that languages, left on their own, can only improve: decay is caused by external catastrophes and resembles a disease from which languages can recover: “(so) glänzend sie sich auch wieder aus diesem neu entwickelt haben.”

Humboldt and Schlegel’s basic question was about language: is language change an inherently directed process? Any eventual answer would essentially have to rest on considerations about the political, cultural, social and linguistic constellations in Northern Italy and Southern France during the crucial period.

Steinthal and Lazarus are commonly quoted as Humboldt’s truest followers and heirs, but Humboldt’s universalist perspective on the subject of investigation was shared by linguists in the nineteenth century in general. Hermann Paul in Prinzipien der Sprachgeschichte states that “Die Kulturwissenschaft ist immer Gesellschaftswissenschaft. Erst Gesellschaft ermöglicht die Kultur, erst Gesellschaft macht den Menschen zu einem geschichtlichen Wesen” (Paul 1880: 7). In view of the fact that Paul saw linguistics as an inherently diachronic science where a purely synchronic perspective meant severe reduction of the topic (Paul 1880: 20ff.), his statement is an indirect command to embed linguistic research in a wider historical-cultural perspective. His own work reflects this conviction in many aspects, and his observations about language and languages are constantly related to reflections on the nature of man and society (e.g. Paul 1880: ch. 14 for an illustration).

At the end of the nineteenth century, Georg von der Gabelentz still presents himself as standing firmly in Humboldt’s tradition. In his 1891 Die Sprachwissenschaft, he provides a colorful description of the multidisciplinary early days of (German) linguistics: “Sie glich einer Colonie, deren erste Bebauer aus verschiedenen Gebieten zugewandert waren, und noch heute gereicht ihr solcher Zuzug oft zum Gewinne”. He continues by list-

thrust, as well as subsequent disruption of the cultural conditions and concomitant decay of language.”

2 “(although) they brilliantly redeveloped” (after this decay).

3 “Cultural science is always social science. Only society makes culture possible, only (his) society will turn man into a historical being”. Note that this passage – part of the introduction to the field of linguistics – does not even single out linguistics as an isolated subbranch of the humanities (then called ‘cultural science’).

4 “It (= linguistics) was like a colony, the founders of which all came from different directions, and until today, such migrants frequently are a gain for the field.”
Historical linguistics as a transdisciplinary field

ing the scholarly origins of the founders, among which we find classical philology, oriental studies, theology, law and medicine. The next paragraphs are devoted to describing the necessary skills and prerequisites of a good linguist. Von der Gablerentz (1891: §§2, 4) recommends not only knowledge in physiology and physics (for matters of phonetics and phonology) but more importantly psychology (because “Die Sprache ist unmittelbarer Ausfluss der Seele”, language is the most immediate reflex of the soul) and finally philosophy, logic and analytic skills in order to be able to turn the multitude of observed facts into a coherent and systematic grammatical theory. Von der Gablerentz (1891: 53) summarizes that “In unserer Wissenschaft ... gilt dies, dass man sich nicht ungestraft vereinseitigt, und dass kein Ab- und Umweg ungelohnt bleibt. Geschichtlichen, länder- und völkerkundlichen, philosophischen, ästhetischen, auch wohl naturwissenschaftlichen Interessen gebe man getrost ihr Recht: ...”.

Linguistics in the nineteenth century was hence tightly interwoven with the humanities. The study of a language was always about the language as a historical entity, as an object that could only be understood against the background of its development over time. What is true for the discipline as a whole holds true for its subparts as well: the investigation of semantics took its origin (and even the term semantics itself was coined) in the investigation of meaning change.

2.2. The emancipation of linguistics

In the twentieth century, both diversification and specialization took place. The cultural sciences split into highly specialized fields, independent of each other and of linguistics, which matured from a mere preparatory subject into an academic discipline in its own right. For the first time, linguistics provided new paradigms of thought that productively influenced other sciences. The so-called linguistic turn substantially shaped the field of philosophy, and structuralist methods inspired leading schools of research in several branches of the humanities.

5 “It is true for our science ... that one may not restrict one’s perspective without serious retributions, and that no digression remains unrewarded. Give in freely to your historical, geographical, ethnological, philosophical, aesthetic interests and those in the natural sciences: ...”
Within the newly emancipated field, historical research was thrust into the background in favor of synchronic linguistics as an empirical science which, in turn, split up into investigations into the phonological, morphological, syntactic, semantic and pragmatic aspects of natural languages. These specializations proved extremely productive, and the field has gained many valuable insights into the linguistic and cognitive abilities of humans by focusing on specialized facets of language.

Like the field as a whole, the discipline of semantics underwent a bifurcation. On the one hand, semantics flourished as a synchronic discipline, perceiving language as a fixed system in which most factors that might lead to variation are abstracted away in favor of a more or less rigid formal treatment of meaning. Compositional semantics in the tradition of the works of Gottlob Frege (1970) and Richard Montague (1974) demonstrated over the last century that semantic research in the sobering spirit of formal logic can lead to deep and intricate findings about the nature of human language and cognition. The approach has since been refined and extended considerably, giving rise to discourse semantics (e.g. Kamp and Reyle 1993), generative lexical semantics (Pustejovsky 1995) or formal pragmatics (Kadmon 2001; Mey 2001). In recent years, the competing paradigm of cognitive semantics has sought to reconcile insights into semantic investigation with psychological facts about categorization and information processing (see Dirven and Verspoor 1998 for an overview).

On the other hand, we can draw on a modern tradition of attempts to explain meaning change, usually cast in less rigid terms. The literature on meaning change offers impressive classificatory work, notably Bréal (1900), Meillet (1925), Stern (1931), Benveniste (1960), Ullmann (1967) or recently Blank (1997). These large-scale treatments of meaning change are complemented by specialized investigations into single patterns of change, notably metaphor, metonymy, lexical fields, grammaticalization and historical pragmatics. Some of these will be reviewed in more detail below.

Most recent investigations into language change adopt a purely language-internal perspective that has already proved advantageous in synchronic research. The universalist tradition of the discipline’s founders has been dismissed in favor of a concentration on the notions and techniques that have been brought to professional heights over the last 100 years. It is only a few isolated attempts that still reveal the value of conducting research at the borderline between diachronic linguistics and neighboring fields. Keller’s (1990) metaphor of the invisible hand brings insights from the study of social interactions to bear on the investigation of language
change. Other investigations of meaning change proceed in terms of general psychology like, for instance, the prototype-based theory of meaning change in Geeraerts (1997) or the account of metonymic change in terms of figure and ground of Koch (2001). In spite of their success, these works remain the exception in the modern literature on diachronic linguistics.

2.3. Recent linguistic investigations into meaning change

In *Recent developments in historical semantics*, Ferenc Kiefer (2001: 13–14) offers the following overview of current research in the diachronic investigation into meaning:

In sum, then, in historical semantics three main lines of research can more or less clearly be distinguished: (i) research based on semantic fields (the structuralist tradition); (ii) the application of prototype theory to historical semantics as well as cognitive linguistic accounts of metonymy and metaphor; and, finally (iii) the use of semantics as well as pragmatic principles in order to account for grammaticalization phenomena.

Apart from the recent productive co-operation of linguistics and psychology in the exploration of prototype phenomena, this quotation makes no mention of links to other cultural sciences. A closer investigation of current literature in diachronic semantics will confirm this impression.

With his famous theory of lexical fields, Jost Trier initiated a new direction in historical semantics (Trier 1931, 1973). In this line of research, the principles and methods of structuralism were systematically applied to diachronic investigation. Language is perceived as a closed coherent whole where changes never occur in isolation but, rather, affect the entire system. Trier’s ideas were taken up in much subsequent work and recast within a sound theoretical basis by Coseriu (1964, 1970). The findings promise interesting repercussions for our ideas of human categorization but, nevertheless, the focus of this paradigm is predominantly language-internal.

Explicit links between psychology and historical linguistics are drawn in the work of Dirk Geeraerts (1997). He proposes that semantic changes should be approached in terms of prototype theory, thus avoiding the obvious problems of categorical lexical semantics based on binary feature systems. Future exchange between (historical) linguistics and categorization
theory in psychology promises to offer valuable insights into the nature of human thinking and reasoning (see also Blank and Koch 1999).

The investigation of metaphor is another branch of linguistic research in which psychological and semantic interests meet. Metaphor as a mode of creative language use has challenged scholars at all times (Black 1962, Lakoff and Johnson 1980, Indurkhyia 1992). Moreover, in a recent line of research, it has been proposed that “emerging” metaphors can underly the meaning of new analytic grammatical constructions (Heine, Claudi and Hünnemeyer, 1991, Sweetser 1990).

Grammaticalization phenomena in general are a flourishing field of diachronic research. Lehmann (1982) is commonly perceived as an early paper defining the field, its basic terms and research agenda. Important later overviews include Traugott and Heine (1991), Hopper and Traugott (1993), Bybee et al. (1994), Pagliuca (1994), Giacalone Ramat and Hopper (1998), Campbell and Janda (2001), Traugott and Dasher (2002), and Wischer and Diewald (2002). Investigations into the emergence of grammar tie in well with the deepened understanding of the syntax-semantics interface that was achieved in the synchronic and modular approach to natural languages. Moreover, advanced computerlinguistic projects allowed separate lines of theorizing to become reunited in overarching frameworks. Application of pragmatic reasoning helped to successfully elucidate the semantic side of grammaticalization (Traugott and König 1991, Traugott and Dasher 2002). While investigations into metaphor and metonymy essentially relate change to conceptualization, grammaticalization phenomena draw attention to speakers’ communicative and interactional competence and present change as a result of social interaction.6 We will come back to this dichotomy below.

The advances in the field of diachronic semantics in the last decades cannot conceal the fact that restriction to just only language-internal considerations alone severely limits the insights that we can hope to gain. We will discuss three cases in which linguistic theorizing essentially rests on results provided by neighboring sciences.

The first case is the Humboldt-Schlegel debate reported above. At the time, scholars (tacitly or explicitly drawing an analogy to biological evolu-

---

6 It should be added that none of the named authors would claim that cognitive or communicative abilities alone can account for language change. Yet, in comparing different branches of investigation, it seems legitimate to highlight the distinctive assumptions.
tion) were trying to evaluate the hypothesis that language development was directed and that it led from minor to more and more improved and refined types of grammar, culminating in *synthetic* languages (Hopper and Traugott 1993: 19f.). The development from Late Latin to the Romance languages was a striking counterexample to this claim. How should one evaluate this well-documented development in the opposite direction? Humboldt’s answer, as witnessed by the quotation above, was that certain language-external factors could lead to language *decay* (i.e. developments in a more analytical direction), whereas language development in an undisturbed cultural setting was language improvement. But whatever conclusion one would have come to, it was clear that in-depth knowledge of language-external facts concerning the time period at issue was mandatory in addressing this question.

It would certainly be premature to claim that such issues are no longer of relevance for modern linguistics. The current agreement that grammars of languages all over the globe are of equal complexity and that no type presents an evolutionary endpoint could only be reached by carefully evaluating informed hypotheses about the influence of one or the other external historical event on language change.

The next case shows how such “big” questions turn into very specific issues with very concise answers. In recent years of grammaticalization research, a common feeling has grown to the effect that many of the competing hypotheses in the field can only be evaluated on the basis of more detailed historical knowledge about the wheres and whens of the instances of change in question. The nature of the *onset contexts* of reanalysis and change is one of the topics currently under debate (Kuteva 2001, Heine 2002, Diewald 2002, Traugott and Dasher 2002 among others).

Let us consider a concrete example. The German intensifier *selbst* (= English *-self*) developed a new use in the sense of a focus particle *selbst* (= English *even*) during the eighteenth century. Scholars agree that the point of emergence must be located somehow between almost synonymous sentence pairs like the following:

(1) *Venus selbst war nicht schöner als Anna*.
   ‘Venus herself was not more beautiful than Anna.’

(2) *Selbst Venus war nicht schöner als Anna*.
   ‘Even Venus was not more beautiful than Anna.’
In spite of the plausibility of this claim, it was still an open question under which circumstances a native speaker of German (between 1700 and 1800) would be whimsical enough to start using *selbst* in the sense in (2) rather than confining herself to the conservative use in (1). The issue of when a potential ambiguity metamorphoses into actual change has often been speculated about, often on the basis of disquietingly little empirical evidence. One could envisage a standard explanation on the basis of *error*, *uninformed language use* or *misunderstanding*.

Yet, a philological survey into the uses of *selbst* strongly suggests that we should trust more in both the competence of speakers and the flexible adaptive potential of natural languages rather than blaming it all on error (note how the notion of *decay* reappears at the horizon). A survey of contemporary documents suggests that ambiguous uses arose earliest in poetry and dramatic verse. Examples can be found where *selbst* in the old sense can only be justified by costly pragmatic accommodation while *selbst* in the modern sense – which speakers might have envisaged before as a latent possibility in their language – fitted perfectly (Eckardt 2001, 2003: ch. 6). Such examples offer plausible onset uses for modern *selbst* in semantic-pragmatic terms. However, the analysis looks wildly implausible in sociological terms. Clearly today, in 2003, language innovation in poetry and rhymed drama would hardly ever have a chance to spread. Can we hence trust the philological findings?

At this point, research in cultural history informs us that in 1750, such a spread was not as unlikely as today. At that time, we witness an increased public interest in literary writing in German, which was seen as part of defining German culture and science, as opposed to the common European humanistic tradition – expressing itself in Latin. A large part of the educated classes was engaged in these attempts in quite practical ways. Literary circles flourished, honourable citizens spent their leisure time with literary discussion and poetic competition (Dietze 1963, von Borries and von Borries 1991). A new use of a word like *selbst*, initiated by leading figures in the normative efforts to define good use of German, therefore had a very good chance of entering into educated and common language. Coming back to the initial linguistic hypothesis, we find that it was in fact strengthened by reference to language-external facts.

What may look like just another isolated word history in fact reflects a broader discussion in the field of grammaticalization research. Competing hypotheses about the driving forces in grammaticalization maintain that (a) metaphorical processes (see Heine, Claudi, and Hünnefelder 1991, and sub-
sequent work by Heine and collaborators) versus (b) communicative-pragmatic constellations are the most prominent cause of reanalysis and language change (the latter was first proposed in Traugott and König 1991 and refined in later work by Traugott and colleagues). The opposing positions take up the older dichotomy of language internal and external factors in a new, more refined sense. A metaphor-based account suggests that our common cognitive capacities are the driving force in language change. The account in terms of historical pragmatics assumes that language change is driven by our communicative and pragmatic competence, more generally the capacity for social interaction through language. We find ourselves in a situation much similar to the Humboldt-Schlegel dispute. Once more, an informed answer to this question will have to rest on solid ground work in other disciplines such as sociology, cultural sciences, psychology, history and literary studies; each applied in careful case studies offering a sound empirical basis for future theoretical work.

We take the current developments as an indication that diachronic linguistics in general, and diachronic semantics in particular, are at the threshold of a new cycle of research. Building on the solid foundation of almost 100 years of synchronic semantic research, the Big Questions in the field seem to re-emerge, calling for the universalist perspective that was shunned for a century. At this point, it might be beneficial to recall the holistic concept of language as a formal, cultural, historical and social object that informed the scientific work of the pioneers of the field in the nineteenth century, and to see how this view contributes to our understanding of meaning as a panchronic notion. Yet, the world has changed and our knowledge has increased. Today, universalist competence of a new quality can be achieved through the co-operation of researchers in different disciplines. We will, however, have to sharpen our awareness of possible points of exchange, illuminating shifts of perspective, and focus, methods, possibilities and limits of each discipline as it addresses language as an object of research.

3. Diachronic semantics from different points of view

Matters change as time goes by. First and foremost, the world and its objects undergo changes, old things vanish and new things emerge. Likewise, our beliefs are constantly adjusted, renewed, extended and corrected as we learn more about the world and its inhabitants. And finally, there is language change. We have the clear intuition that we can and should distin-
guish these modes of change and that each is exemplified by its own prototypical cases.

Stern (1931: 194) nicely discusses changes in the domain of referents as “not very interesting” for the purpose of developing a theory of meaning change. Focusing on the word *ship*, he points out that the word previously denoted wooden vessels, propelled by windpower or oars. In the wake of technical progress, the term *ship* was extended to steam boats, motor-driven ships and ships with nuclear propulsion. Yet, we have a clear intuition that these changes essentially rest on technical progress, paired with a completely conservative use of the word *ship*.

It is perhaps more difficult to draw the line between knowledge increase and meaning change. Still, we find examples of an increase of knowledge, perhaps even with repercussions on language use, which occur under constant meanings. For instance, at some time the comparison of the inner organs of whales with those of other fish and mammals suggested that whales should be classed with the latter rather than the former. This however was perceived as a correction of the previous classification of whales and not as a change in the meaning of ‘fish’. Earlier generations would call a whale a ‘fish’ because they lacked knowledge about its inner structure, not because they lacked proper knowledge about the meaning of ‘fish’.

*Meaning changes* proper are changes in the conventions that determine the referents of a word. When Latin *trahere* (‘to pull’) changed to French *traîre* (‘to milk’), speakers adopted the additional convention to restrict the word to actions of pulling at a specific place (an udder) and with a specific purpose. Evidently, there is no concomitant change in knowledge (milking cows and goats was a technique known before) nor in the world (neither goats nor cows nor farmers underwent substantial changes).

These observations pose several exciting challenges for the semanticist. On the empirical side, studies in word history will frequently address cases that cannot be cleanly assigned to one of the three modes of change. Changes in the world, in knowledge and in linguistic conventions are inextricably inter-woven and we cannot hope to gain a full picture as long as we restrict our attention to the clean cases. The papers in the first part *Changing beliefs, diversifying worlds, and flexible meanings*, written by Andreas Blank, David Kronenfeld and Gabriella Rundblad, David J. Wasserstein and Judith Meinschaefer, offer an inspiring vista into the intricate exchanges between political developments, economical factors, social constellations in and between language communities, conceptualization and language use.
On the theoretical side, they pertain to the task of developing an adequate notion of meaning. It is widely agreed that a word’s meaning arises in the tension between its referents and the pieces of knowledge that relate the word to other words of a language. Proper names offer one particularly clear case where the reference of a name determines its meaning. The complementary case can be exemplified by words that receive their meaning by explicit definition, like in *a sow is a female adult pig*. The vast majority of word meanings, however, emerges in a dynamic equilibrium between knowledge and reference that is hard to explicate. As a result, we face a potential paradox. Not all changes in referents induce meaning change. Not all changes in knowledge induce meaning change. And yet, knowledge and reference conspire to determine a word’s meaning. The contributions in the second section *The meaning of meaning change*, written by Hans Rott, Ulrike Haas-Spohn, and Regine Eckardt offer three different answers to this paradox.

The third part, *The force of grammar*, comprises studies in meaning change that highlight the beneficial combination of advanced linguistic tools and theories with the methods of other disciplines such as thorough philological analysis, research into social interaction and discourse analysis. While the great value of philological, social, and anthropological research in the investigation of meaning change is already established in several contributions in part one and two, the papers by Eva-Maria Gerö and Arnim von Stechow, Miriam Butt and Wilhelm Geuder, Walter Breu, and Susanne Günthner address changes that, in addition, require a high degree of linguistic sophistication in order to reach an adequate analysis. Studies of that kind impressively prove that the temporary retreat to synchronic linguistic theorizing did not necessarily lead the field away from diachronic research. On the contrary, we witness a renewed co-operation of linguistics and the humanities, investigating classical issues at a higher level.

3.1. Changing beliefs, diversifying worlds, and flexible meanings

We will now proceed to summarize the articles in the order in which they appear in the volume. The first two contributions, by Andreas Blank and David Kronenfeld and Gabriella Rundblad, both perceive meaning change...
as change in the conceptualization of our world. The authors share the view that object domains are conceptualized into semantic or conceptual fields which are structured by certain parameters. Languages map sets of words (lexical fields) onto these semantic/conceptual fields, and the relation between lexical structure and conceptual structure is shaped by cognitive processes such as metonymy and metaphor. Blank as well as Kronenfeld and Rundblad argue that meaning reflects our conceptualization of the world, and that change in the conceptualization results in meaning change. Hence, both contributions base their inquiry into diachronic semantics on a modern version of the classical theory of lexical fields (Trier 1931), extended and refined by notions from cognitive linguistics and anthropology.

In his contribution Words and concepts in time: Towards a diachronic cognitive onomasiology, ANDREAS BLANK demonstrated that onomasiology, i.e. the way we name the objects we refer to, renders a more truthful representation of the way we conceive and conceptualize objects than does the traditional perspective of semasiology (the investigation into the meaning of expressions). Blank’s investigation into conceptualization is based on detailed contrastive comparisons of how different languages name one and the same object. One of his examples is words for ‘small piece of wood for lighting candles, cigarettes etc.’. The English word match was formed as a metaphor from the word match ‘(lamp) wick’, while the French word allumette for the same object is derived from allum- ‘to light’ and a suffix with the meaning ‘instrument for...’, and the German Streichholz is a compound that combines Holz ‘wood’ with streichen ‘to rub’ – thus the composition refers to the act of lighting. Blank argues that onomasiological principles not only structure the synchronic lexical fields, but also restrict the process of meaning change. He concludes that the principles of lexical change can only be understood if we start from the cognitive level of concepts and proceed in the onomasiological direction towards the realm of words. This leads Blank to a diachronic cognitive onomasiology as the background theory for diachronic semantics.

Blank’s article adopts a typological perspective, demonstrating both that the same real object can be conceptualized in different ways in different cultures as well as investigating the implications of this for meaning change. In contrast, the contribution of DAVID KRONENFELD and GABRIELLA RUNDBLAD illustrates the successful interaction between structural linguistics and anthropology in the investigation of one language community. In The semantic structure of lexical fields: Variation and change, the authors start from Trier’s (1931) notion of lexical field and ap-
ply it to names of watercourses in English. They argue that the particular name for a certain type of watercourse can only be understood in opposition or relation to other names in the lexical field, and they proceed to a detailed and informed account of the diachronic development of the entire lexical field. While they closely follow Trier’s program in his seminal book, they present an application that lends an additional dimension to the method. Their research program addresses a conceptual domain that allows for inquiries into the language external, historical, social and economical factors that drive the changes of the lexical field.

The authors perceive language as grounded in a speech community where the use of words and terms is driven by the speakers’ interests and needs – a perspective that is taken up later in Eckardt’s contribution and set to work in her theory of reference. They use methods from cognitive anthropology and cultural linguistics (a position that is characterized more closely in their article) for describing the status of a particular word in a lexical field. Moreover, the authors combine methods of linguistic field research with philological analysis in order to characterize the lexical field of watercourses from late Old and Middle English up to the present: the vocabularies of contemporary speakers are assessed by using questionnaires and interviews. The semantic fields of speakers of earlier times are reconstructed indirectly by inferences based on contemporary usage, complemented by relevant contextual and functional information from source texts and dictionary information.

One major finding of Kronenfeld and Rundblad’s analysis is that the lexical field of watercourses at earlier times was organized on the basis of the features QUANTITY, SPEED and QUALITY, while the contemporary system is mainly organized by the feature QUANTITY. This falls in line with the observation that the contemporary lexical field comprises fewer names than at earlier stages. In addition, the authors show that differentiation mainly concerns middle and small watercourses. As an example, they focus on the words burn and brook, words that are used more or less synonymously for ‘river’ in modern English. The word burn (Old English burna ‘a bubbling or running watercourse with clear water’) was more frequently used in Old English than brook (Old English brôc by metonymic transfer from ‘marsh’ meaning ‘marshy watercourse’). At a later stage, brook became more frequent than burn. The historical explanation is that early colonization preferred places with clear and running water (= burn), while later colonization had to take less favorable places with rather opaque and muddy water (= brook). In terms of historical semantics, the most interest-
Regine Eckardt, Klaus von Heusinger, and Christoph Schwarze

The crucial point is that these changes in everyday life not only effect changes in the frequency of use of one or another word (which would be a trivial result) but, moreover, determine subsequent semantic developments, for instance the subsequent use of one rather than another word as a general cover term.

David Wasserstein’s contribution *khalîfa – A word study* treats the semantic evolution of the Arabic word *khalîfa* ‘caliph’. The word, originally labelling the position of the unique, divinely justified political leader of the Islamic world, is evidently closely interwoven with the society’s political reality, history, and the interests of leading figures in society. Similar key words in politics (and religion) frequently lead a fascinating semantic double life: while the content of the word is unofficially defined by interested and influential parties, official decisions can subsequently be justified with reference to the word’s semantics, pointing out that the course of events proceeds ‘true to the word’. Wasserstein’s article offers a detailed and informed case study of this kind. We witness a case in which the clean distinction between meaning change and reference change, proposed at the beginning of this section, breaks down completely. The very emergence and the subsequent evolution of the notion and referents of *khalîfa* originates from human political activities, and linguists ordinarily refrain from investigating similar cases because they would lead them beyond the safe ground of their own methods and expertise. It is not surprising that the study of *khalîfa* is written by a historian rather than by a linguist, offering well-prepared material for semantic investigation.

Wasserstein’s study raises several intriguing questions for the linguist. Are the changes in the use of the word under study changes of its meaning? Or is the variation in the use of *khalîfa* just a reflection of the polysemy inherent to the root *kh.l.f*, from which the word is derived? Or should the story of *khalîfa* be understood as a variation in elliptic usage? As can be gathered from Wasserstein’s analysis, *khalîfa* ‘deputy’, ‘successor’ is a relational noun and the variation, at least in part, might correlate with elided material. Different answers could be envisaged, to a certain degree depending on the theoretical position adopted by the specific scholar. The contributions in the second part of the book will demonstrate the discrepancies that can arise even in a formal semantic treatment of the meanings and meaning changes of ontologically extremely simple words like ‘gold’, ‘water’ or ‘jade’ (natural kind terms). The meaning and development of a word like *khalîfa* evidently exceeds such cases in complexity by several dimen-
Historical linguistics as a transdisciplinary field

It remains a challenging goal for semantic theory to account for the kind of variation documented in Wasserstein’s contribution.

The case study of Kronenfeld and Rundblad on names for watercourses highlighted the interdependencies between economic interests and language use. Political and historical expertise were indispensable for an adequate account of \textit{khalîfa} by Wasserstein. \textsc{Judith Meinschaefer}’s study *Words in discourse – On the diachronic lexical semantics of ‘discours’* exemplifies yet another method of accessing earlier language stages. On the basis of a scrupulous analysis of historical texts, Meinschaefer offers a detailed account of the stages of use of French \textit{discours}. The article takes its starting point from a careful philological evaluation of the use of \textit{discours} in the works of Michel de Montaigne and Jean-Jacques Rousseau, counterchecked by lexicographic information in historical dictionaries. Yet, the strong linguistic background of the author leads her to new ground.

First, Meinschaefer’s article goes beyond the unquestioned application of philological text analysis in that she offers justification for this method on the basis of a linguistic analysis of the contexts of use of the term under scrutiny. When we read and try to understand texts of the past, we are in fact confronted with a fundamental problem. Initially, all that contemporary readers have at their disposal is their semantic competence of the present day language. As their experience in reading historic texts increases, the readers may come to feel that they are acquiring a specific competence for older stages of the language. But how can we be sure that this is not just an illusion? Historians and literary scholars usually trust their intuition or appeal to the so-called hermeneutic method, which basically is a circular verification of interpretative consistency. Both positions are in contradiction with the analytical tradition to which modern linguistics is committed. Meinschaefer shows how a painstaking distributional analysis can help to make our understanding of word meanings in historical texts more transparent. The reader may decide whether this approach actually frees us from the hermeneutic method or merely makes that method seem more acceptable.

A second point of interest is tightly connected with the first. The domain of reference of the word under study, French \textit{discours}, is entirely located in the realm of culture. This puts the topic of Meinschaefer’s contribution close to Wasserstein’s, and it is revealing to compare the different approaches taken by the historian, the linguist and the scholar in literature (in part demonstrated in Meinschaefer’s article). Of course, Kronenfeld and Rundblad’s anthropological perspective (adequate for what one might call
‘geographic kinds’), and the philosopher’s interest in natural kind terms, exemplified by the articles of Rott and Haas-Spohn, mark other prominent positions in the landscape of historical word study. Ordinarily, linguists and philosophers prefer to study problems of lexical meaning with respect to words that denote artifacts or natural kinds. Meinschaefer bridges the apparent abyss by integrating the object of her study into an ontology familiar to semanticists: *discours* refers to a kind, more specifically, to a “cultural kind”.

The third point of interest rests upon the second. Since *discours* refers to a cultural kind rather than to some irreducible aspect of intellectual life, Hilary Putnam’s (1975) concept of division of labor can successfully be applied to the diachronic study of this word. In analogy to the definition of the meaning of *water* or *gold*, we must distinguish between the experts’ meaning of *discours* and common usage. Meinschaefer investigates the writings of Michel de Montaigne and Jean-Jacques Rousseau who can justifiably be viewed as experts in the domain of cultural kinds, and even more as agents in the development of expert knowledge: both authors exert influence on the meaning of the word by restricting the previously existing amount of polysemy. Common usage then follows the experts’ restrictions to some extent, but the speakers do not fully acquire the experts’ knowledge. This may lead to further semantic innovations such as, in the case of *discours*, the emergence of a negative connotation in Modern French.

3.2. The meaning of meaning change

One of the major challenges for the historical semanticist consists in teasing apart changes in the world, changes in knowledge, and changes in meaning. This point was illustrated by several of the case studies in part one of the book, such as, for instance, Judith Meinschaefer’s and David J. Wasserstein’s papers which discuss closely interwoven changes in meaning and matter. The papers in the second part of the book illustrate that a clean, meaningful and empirically adequate distinction between meaning change, belief revision and reference change is anything but trivial, even if we concentrate our attention to a class of expressions that are semantically fairly simple.

The papers by Rott, Haas-Spohn and Eckardt address the meaning of natural kind terms and illustrate that different guiding questions will lead to different preferences with respect to the most appropriate semantic account.
The working linguist will be most interested in an account of word meaning and change of meaning that reflects the empirical linguistic facts correctly and in a simple manner. In the philosophy of science, however, the question of language change is inextricably connected with the question of scientific progress: there is a shared belief that scientific progress typically means that we learn more about the same thing, referred to by a name with constant meaning. Only in exceptional cases would we admit that science underwent a paradigm change where not only knowledge, but the entire way of talking about a given subject has changed (see Kuhn 1962).

Yet, according to one commonly held view, the meaning of a word arises holistically from the entirety of speakers’ world knowledge concerning that word (the position is discussed in detail in the contribution of Hans Rott). Obviously, this view comprises an unhealthy over-eagerness to diagnose a Kuhnian paradigm change whenever we acquire a new bit of knowledge about some term: in terms of this theory, scientific progress can only mean that we know different things about different notions rather than (as one would intuitively like to say) learning more about the old notions. The three contributions by Rott, Haas-Spohn and Eckardt demonstrate the tensions that arise between the linguistic and the philosophical perspective in addressing these issues.

HANS ROTT’s contribution *Theoretical concepts in flux – Conceptual knowledge and theory change* investigated the relation between world knowledge and meaning. Rott recapitulates the classical insight that there seem to be two kinds of facts that can be reported with respect to a given word. A sentence like *bachelors are adult unmarried men* seems to report a fact about the meaning of *bachelor* and is tautological for anyone who can claim to master the word. In contrast, a sentence like *bachelors develop strange eating habits* reports an empirical fact about bachelors. Rott relates this distinction to the philosophical dichotomy between *analytic* and *synthetic* judgements. In a historical survey, he perspicuously traces the notions of *analyticity* and *syntheticity* in the works of Immanuel Kant, Gottlob Frege and William O. Quine. At least the former two philosophers developed these notions primarily on the basis of the example of mathematical truths (*analytic*) in contrast to scientific findings (*synthetic*), and it is a challenging task to apply their notions to the case of truths about word meanings, in contrast to truths that are communicated with words (i.e. word meanings). Rott’s summary not only offers a lucid overview for readers without a strong background in analytical philosophy, but it also brings to light the constant truths underlying the never-ending debate on the relation
between analytic sentences and lexical knowledge, thus paving the way for his own proposal.

Rott approaches the classical distinction between analytic and synthetic statements on the basis of the techniques and methods of formal logic. More specifically, his account is formulated in terms of belief revision and theory change. Rott takes advantage of the fact that investigations into logic, model theory and nonstandard logics over the last decades have resulted in very clear, explicit and concise ways to distinguish languages, interpretations, theories and theory change. He proposes a systematic way to relate changes in language, changes in meaning, and changes in theory in a coherent and adequate manner. Following the classical method of indirect axiomatic characterization of the core properties of a given phenomenon, Rott explicates links between the property of being an analytical statement, being stable under minor belief revisions and the process of major change of an epistemic state. The resulting account is an indispensable gauge for any empirical diachronic investigation in areas where the history of language and the history of ideas are closely interwoven and hard to separate. Rott modestly restricts his attention to the meanings of words that are part of scientific discourse, but in the context of the overall volume, his theory gains impact in the stimulating confrontation with more demanding case studies such as those presented in part one.

With her contribution Meaning change as character change, ULRIKE HAAS-SPohn takes a different line in the meaning representation of natural kind terms. She takes up the proposal of Putnam (1975) according to which the meanings of natural kind terms such as, for instance, gold or water are constituted by direct reference: any piece of gold will represent the entire extension of the kind, and naming a piece of gold gold will yield a name that refers to all and exactly this extension. Haas-Spohn reconstructs Putnam’s account in terms of Kaplan’s (1977) two-dimensional semantics, originally designed to capture the meaning of context-dependent words like I, here, and now. Haas-Spohn carefully recapitulates the core ideas and techniques of both works, thereby making her paper readily accessible to a wide readership with only elementary knowledge of logical semantics.

The most important advantage, Haas-Spohn proposes, lies in the fact that her account allows us to disconnect the link between identical reference and necessary identity of natural kind terms. She can explain why the insight that “water is H₂O” constituted a true piece of scientific progress at the time of discovery even though the natural kind terms water and H₂O are coextensional by logical necessity (following the original theory in Put-
Her proposal in fact conceives of language as an inherently historical entity and it amounts to making the respective natural language itself a context-dependent object. The meanings of words can change between earlier and later contexts of use. Scientific progress, according to her proposal, sometimes actually does amount to an adjustment of meaning, yet in a carefully circumscribed manner. Haas-Spohn’s contribution hence offers a detailed and well-justified answer to the question “how do meanings arise and change?”. She manages to capture central observations in language philosophy and the philosophy of science and offers a convincing synthesis of apparently conflicting standpoints.

The contribution Meaning change in conceptual Montague semantics by REGINE ECKARDT likewise relates to the seminal paper by Putnam (1975). However, she criticizes Putnam’s referential theory of meaning in at least two respects. First, Eckardt insists that the account’s narrow focus on natural kind terms, perhaps tolerable from the philosophical point of view, is untenably limited for linguistic purposes. Second, she argues that Putnam’s theory forces one to adopt a counter-intuitive notion of meaning change. Looking at real word histories, the author demonstrates the discrepancies between meaning changes in an intuitive sense, and meaning changes in the terms of Putnam.

Eckardt agrees with Putnam that frequently, a word’s meaning seems to be determined by ostensive reference to selected or prototypical exemplars of a kind. Yet, she maintains the view that ‘natural kinds’ should not be given a special status in the communicative practices of speakers. She proposes that natural kind terms in fact represent just another way of classifying the world into discrete categories and should not be treated in a separate theory of lexical semantics.

The author develops a theory of word meaning in which a word’s meaning is established on the basis of reference to a typical exemplar plus a mode of categorization. The mode of categorization is determined relative to the interests, knowledge and needs of the interacting speakers. The interaction between referent and speakers’ interests in establishing word meaning faithfully mirrors the findings of Kronenfeld and Rundblad who illustrate this process in their empirical study of watercourse names. These examples, as well as others offered in the article, suggest that the resulting account can treat a major part of the lexicon of a language rather than being restricted to a limited class of words in the expert language of natural sciences.
The resulting theory retains Putnam’s view that expert languages have a special status in the negotiation of meanings, but these experts are now coherently integrated into a uniform theory of word meaning. Putnam’s stereotypes – his account for the ordinary man’s kind of lexical knowledge – can likewise be located as an integral part of such an overarching theory of lexical meaning. This seems to be a more satisfactory perspective than Putnam’s suggestion to treat stereotype knowledge as a secondary, inferior mode of meaning representation for the purposes of everyday language. Finally, Eckardt evaluates the account against the history of the word jade (and its Chinese predecessor), a real instance of language use and knowledge acquisition. The resulting picture is compared to Putnam’s treatment of the case, pointing out that the philosopher’s theory will necessitate a counter-intuitive extension of the notion of meaning change.

The two preceding papers are written against a common scholarly and terminological background and yet show two divergent strategies for reconciling semantic theory with the facts of historical developments. Reading either contribution against the background of the other, we see how the dialogue between historical semantics and analytical philosophy can highlight different aspects of the phenomenon under scrutiny, depending on the guiding research question. The philosopher will justifiably call to mind that there are certain indisputable fixed points in human thinking and reasoning that no semantic theory should weaken or blur. The linguist, in contrast, will want to exploit the systematic philosophical groundwork within a broader range of phenomena, and will use core philosophical insights to develop theories that render a less idealized picture of communicative reality.

3.3. The force of grammar

The articles in the first two parts of the volume focus mainly on language-external forces in meaning change. The third part comprises contributions that demonstrate how advanced grammatical and semantic theorizing conspires with expertise in discourse analysis, classical philology, and sociological field research in the investigation of grammaticalization processes, language contact phenomena and reanalysis. These contributions introduce an additional level of complexity in that the authors have to keep track not only of words and their meaning but also of so-called functional elements, meaningful parts of the sentence that, moreover, play the syntactic-
semantic role of joints between other parts of the sentence. Changes in meaning occur in coherence with changes in syntax, and the question of which one be the driving force is still unsettled. Case studies like the following hence rest on fully developed theories of morphology, syntax and semantics but also require philological expertise, sociolinguistic methods and techniques of field research. Notably, studies about on-going language change like those of Breu and Günthner can, moreover, take advantage of the fact that the researchers are intimately acquainted with the social, historical, cultural and political situation of the contemporary speech communities under investigation. Consequently, we anticipate that future cooperation between linguistics and the neighboring disciplines in the field of grammaticalization research will lead to results of a new quality.

The contribution *Tense in time: The Greek perfect* by Eva Gerö and Arnim von Stechow is a study on a classical problem of historical linguistics. Addressing the evolution of the Greek perfect, the authors combine the methods of formal syntax and semantics with the philological knowledge of old texts. The formal representation of the meaning contribution of tense and aspect markers constitutes one of the core topics in formal semantic research, and older theories that rest on metaphoric circumscriptions of tense and aspect meanings have been replaced by systems that explicate tacit reference parameters and the complex cross-references between the temporal location of states and events that are lexicalized in tense and aspect systems in natural languages. The present article represents one of the first attempts to bring these results to bear on the investigation of language change, and the authors demonstrate an impressive cooperation between expertise in classical philology and semantic proficiency in the diagnosis and representation of aspect meanings.

As a result, their account offers more than just another way to represent the evolution of the Greek perfect. The authors trace the gradual generalizations and subtle shifts that lead from one language stage to the next and are able to explicate how minimal semantic changes, interacting with contextually driven adaptations of meanings and pragmatic inferencing, can result in considerable semantic shifts at the surface level. Gerö and Stechow’s analysis also relates aspect marking in Old Greek to tense and aspect marking in other languages; the authors explicitly refer to English, German and Swedish (the languages for which the authors’ model of tense semantics was originally developed). The resulting picture allows us to combine a language internal view of the tense/aspect system of a given language with
a more universal explication of a constant small set of parameters, notions and relations that underlie the tense systems of many languages.

We should also point out that Gerö and Stechow’s contribution can be read at two levels. Readers with a background in semantics will appreciate this paper as an example of how the infelicitous separation of theory-oriented linguistics and classical philology may be overcome. Those readers who do not have this background, on the other hand, are not expected to take lectures in formal semantics in order to understand the paper. Informal summaries between formal parts allow an intuitive understanding of the authors’ notion of ‘Extended Now’ and how it can help to resolve a much-debated puzzle in the history of Greek, and the more technical parts may be appreciated as the formal verification of the prose rendering of the account.

MIRIAM BUTT and WILHELM GEUDER’s contribution Light verbs in Urdu and grammaticalization is, just like Gerö and von Stechow’s study on the Greek perfect, an investigation which combines results of current linguistic theory with philological knowledge about a group of languages which is well documented over an extremely long period. The authors address the development of light verbs in Urdu and related contemporary and earlier Indo-Aryan languages. After a very careful introduction to the prosodic, syntactic and semantic characteristics that distinguish light verbs in Hindi and Urdu from auxiliaries and full lexical verbs, Butt and Geuder proceed to the question of whether Indo-Aryan light verbs are, as has been claimed, the result of an on-going process of grammaticalization. They observe that these verbs do not fit very well into the pattern known as the grammaticalization cline. In answer to this puzzle, they argue that the origin and the surprising stability of light verbs and their transparent relationship with phonologically and morphologically identical full verbs can be better explained by assuming lexical and syntactic variation based upon polysemy.

Whatever the status of light verbs with respect to grammaticalization eventually may be, it is clear that light verbs enter into a combination with the main verb that is syntactically, as well as semantically, much more intricate than what we know from common auxiliaries and tense constructions. A large part of the article focuses on the subtle interactions at the syntax-semantics interface that allow the light verb to contribute to both argument structure and content of the overall construction in a way that derives semi-transparently from the original content verb. According to Butt and Geuder, the meanings of Indo-Aryan light verbs emerge by dropping one or more semantic features of the full verb in such a way that the
verb can no longer describe an independent event. Light verbs contribute to the description of the event referent of the main verb in a way that can be compared to adverbal modifiers in languages like English or German. Yet, the interaction between main and light verbs is evidently of higher complexity in that the light verb and the lexical verb build up a shared argument structure. Butt and Geuder can show that the special nature of light verbs is also reflected by their behavior over time. In contrast to other pathways of verb grammaticalization, light verbs seem to mark a stable point in grammar; they remain unaffected by further change over millenia.

WALTER BREU’s contribution *Bilingualism and linguistic interference in the Slavic-Romance contact area of Molise* presents a classical study of language change by language contact. He investigates the language contact situation in Molise (Southern Italy) between the Slavic dialect of a community of immigrants who entered Italy around 1500, and Italian. The study hence takes up the traditional theme of language change in a bilingual community that was a main topic even in Humboldt’s time. However, while Humboldt and his contemporaries were still most concerned with the moral qualities of speakers and their languages, Breu’s investigation focuses on the co-existence of two different conceptualizations of the world in the mind of the bilingual speaker. On the basis of data that were collected over several decades in the contact area, Breu has the possibility of tracing and documenting the pathways and limits of analogical levelling at all linguistic levels with a degree of detail that is beyond reach in the investigation of language change in classical languages.

In the Molise language community, Italian is the adstratum (or superstratum) language, while Molisian Slavic is the substratum language. Breu, illustrating his claim with well-chosen examples, proposes that bilingual speakers combine the two grammatical systems in the most economic way rather than using two different grammatical systems. He shows that if the two languages differ with respect to polysemy, speakers tend to level this distinction between the two systems in favor of a constellation in which both languages observe an isomorphic mapping between concepts and words. For instance, the Italian word *prima* ‘first’ and ‘earlier’ corresponds to two words in the Slavic dialect, namely *prvo* ‘first’ and *prije* ‘earlier’. While conservative speakers still carefully obey this distinction, the majority of current (bilingual) speakers use *prvo* for both meanings ‘first’ and ‘earlier’, corresponding to their use of Italian *prima*. Breu calls this change *polysemisation*. Extending the perspective to linguistic objects in general, Breu also considers interferences at the morphosyntactic level such as de-
clension or the tense-aspect system and convincingly argues that the cause of a lexical or grammatical change in one language is often a parallel construction in the contact language. In reviewing numerous cases of adjustment between the two systems, Breu notes that there are certain intriguing restrictions: if the adstrat language Italian has more words for a semantic field than Molisian Slavic, i.e. if Italian is more differentiated, he rarely observes a levelling process with respect to a higher differentiation of Molisian Slavic and innovation of new forms.

Breu finally traces the innovation of functional elements such as articles, and his findings bring to light yet another factor motivating words to undergo grammaticalization. Breu notes that the development of an indefinite article in Molisian Slavic is also covered by polysemy. The indefinite article na ‘one’ is developing from the homonymous numeral parallel to the Italian pattern, where the indefinite article uno, una has the same form as the numeral for ‘one’ uno, una. Molisian Slavic indeed exploits a potential polysemy of the word denoting ‘one’ rather than building up a full article system in analogy to the Italian superstratum. This, Breu points out, is proved by the fact that Molisian Slavic has not developed a definite article: in Italian, the definite article and the demonstrative are quite different forms and offer no basis for transfer. This shows that the co-existence of similar linguistic forms with different functions has not only a special function language internally, as demonstrated by Günthner for German conjunctions and by Butt and Geuder for Urdu light verbs, but it is also a driving factor in interference between two contact languages.

Finally, SUSANNE GÜNTHNER’s contribution is a detailed account of an on-going process of meaning change. In her article Lexical-grammatical variation and development: The use of conjunctions as discourse markers in everyday spoken German, Günthner argues that the conjunctions weil ‘because’ and obwohl ‘although’ have grammaticalized into discourse markers. She discusses conversational data from colloquial German collected in the last 30 years and proves that the variations in the use of weil and obwohl are not – as prescriptive school grammars would have it – due to mistaken or erroneous language use. To the contrary, Günthner shows that two homophonic variants of weil, a discourse marker variant and a conjunction variant, can be clearly distinguished by their different meanings, scope properties, prosodic behavior, syntactic behavior as well as pragmatic implications. A similar distinction is demonstrated for obwohl.

The most prominent difference is that the conjunction weil introduces a subordinated verb final sentence (... weil er Grippe hat ‘... because he has
flu’) while the discourse marker weil introduces an independent verb-second sentence (... weil (-) sie läuft total deprimiert durch die Gegend ‘... because (-) she is walking around looking totally depressed’). The conjunction weil is prosodically integrated into the main clause, while the discourse marker weil constitutes its own prosodic phrase. In terms of semantics, coordinating weil offers a cause for the truth of the fact denoted by the main clause whereas the discourse marker offers the reason why the speaker makes the assertion in the other clause. More subtle semantic distinctions can be drawn regarding the scope interaction of weil and negation. Finally, the author shows that there is a small closed lexical class of conjunctions with similar behavior (weil, obwohl, während, wobei). The descriptive part of Günthner’s article demonstrates impressively that a detailed linguistic analysis of different uses of a word can elucidate grammatical systematicity and structure in cases where schoolgrammars retreat to a simple black-and-white picture.

While Günthner’s work includes a valuable linguistic description of a piece of present day German grammar, the true merits of the article lie in its developmental perspective. The analysis is based on a corpus of spoken dialogues of colloquial German spanning 30 years which allows minimal intermediate steps in the development to be traced. We hence have the fascinating possibility of witnessing an on-going language development. It has frequently been noted (see Janda 2001: 316f. and references therein) that the investigation of contemporary language variation and change has several considerable advantages over the investigation of past changes. Not only can the researcher gather data with a degree of detail that can practically never be achieved in written corpora, but we also have the possibility of eliciting negative judgements (ungrammaticality, unintelligibility, inappropriateness). Most importantly, the investigating linguist is acquainted with the non-linguistic context of variation to a degree that would require the work of several specialists if it were to be reconstructed for earlier language stages.

We, as well as the quoted scholars, do not suggest restricting attention to current variation and giving up interest in earlier changes altogether.

---

8 Note that even corpora of present day language are in most cases restricted to specific kinds of texts, mostly newspapers. Only the internet offers resources that come close to spoken colloquial language, but it seems unclear whether advances in technology will be conservative and retain written web texts long enough to turn them into a useable resource for diachronic research.
Only under a long-term perspective can we single out and understand the equilibria of grammatical systems (like the tense/aspect system that emerged in the process described by Gerö and Stechow). Only the long-term perspective reveals the points of inertia (like the light verbs considered in Butt and Geuder). Only under a long-term perspective can general trends in conceptual variation become visible (like those documented by Breu). Only the long-term perspective allows us to single out stable systems of categorization in the tension between ‘natural’ properties and ‘relevant aspects’ of the objects in our environment (as in the case of semantic field of watercourse names in English, presented by Kronenfeld and Rundblad). Similarly, it is only Günthner’s experience in the long-term perspective that allows her to name the lasting trends emerging from her data. Studies in on-going language variation and change, however, will set new standards for the level of finegrainedness of further investigations into past language variation and language history. This highlights the pressing need for the diachronic linguist to cooperate with experts in the humanities, history, psychology, cultural sciences and anthropology in order to achieve equally detailed descriptions and evaluations of past language changes.

The present volume aims at provoking this kind of co-operation by presenting, and hence drawing attention to, investigations into meaning change conducted under different perspectives. We present in-depth studies illustrating the scholarly perspective of different disciplines that all pertain to meaning change. All are written by professional researchers in their field, exemplifying their methods, assumptions and results. Several contributions are the result of interdisciplinary co-operation of two authors, or of a single author with strong support from external consultants. More interestingly though, we see various clusters of articles which, in synopsis, highlight particular questions about meaning change which, in turn, reveal their full depth only under a multidimensional perspective. The collection will hence increase sensitivity of points of mutual exchange and beneficial dovetailing of different fields. Historical semantics is a research field that offers, and also requires, different points of view. Only if it is conducted in this spirit can it reveal to the scholar the best that it has to offer.

References

Benveniste, Émile
Black, Michael
Historical linguistics as a transdisciplinary field


Blank, Andreas

Blank, Andreas and Peter Koch (eds.)

van Borries, Ernst and Erika van Borries

Bréal, Michel

Bybee, Joan, Revere Perkins, and William Pagliuca

Campbell, Lyle and Josef Janda (eds.)

Coseriu, Eugenio
1964 Pour une sémantique diachronique structurale. In Travaux de linguistique et de littérature 2 (1), 139–186.

Coseriu, Eugenio, Gunter Narr, and Erich Brauch

Devitt, Michael and Kim Sterelny

Dietze, Walter

Diewald, Gabriele

Dirven, René and Marjolijn Verspoor
Eckardt, Regine

Frege, Gottlob

von der Gabelentz, Georg

Geeraerts, Dirk

Giacalone Ramat, Anna and Paul Hopper

Heine, Bernd

Heine, Bernd, Ulrike Claudi, and Friederike Hünnefelder

Hopper, Paul J. and Elizabeth Closs Traugott

von Humboldt, Wilhelm

Indurkhya, Bipin

Jackendoff, Ray

Janda, Richard D.
2001 Beyond “pathways” and “unidirectionality”. On the discontinuity of language transmission and the counterability of grammaticalization.
Historical linguistics as a transdisciplinary field


Kadmon, Nirit

Kamp, Hans and Uwe Reyle

Kaplan, David

Keller, Rudi

Kiefer, Ferenc

Koch, Peter

Kuteva, Tania

Kuhn, Thomas S.

Lahiri, Aditi, Alexander Patschovsky, and Christoph Schwarze

Lakoff, George and Mark Johnson

Lehmann, Christian

Mey, Jacob L.
Meillet, Antoine

Montague, Richard

Pagliuca, William
1994 *Perspectives on grammaticalization*. Amsterdam/Philadelphia: Benjamins.

Paul, Hermann
1880 *Principien der Sprachgeschichte*. Halle: Niemeyer.

Plank, Frans

Pustejovsky, James

Putnam, Hilary

Stern, Gustaf
1931 *Meaning and Change of Meaning: With Special Reference to the English Language*. (Acta Universitatis Gothoburgensis 38,1.) Goeteborg: Elander[Dr.].

Sweetser, Eve E.

Trabant, Jürgen

Traugott, Elizabeth Closs and Richard B. Dasher

Traugott, Elizabeth Closs and Bernd Heine (eds.)
Historical linguistics as a transdisciplinary field

Traugott, Elizabeth Closs and Ekkehart König

Trier, Jost


Ullmann, Stephen

Wischer, Ilse and Gabriele Diewald
Words and concepts in time:
Towards diachronic cognitive onomasiology

*Andreas Blank †*

1. Introduction

Recent issues in diachronic lexical semantics have shed new light on an old-fashioned linguistic discipline, i.e. onomasiology.\(^1\) This is less surprising when one is aware of the fact that, in their groundbreaking essay on metaphor, Lakoff and Johnson (1980) rediscovered the existence of metaphorical schemas, the so-called “conceptual metaphors”, such as LINGUISTIC CONCEPTS ARE CONTAINERS, ARGUMENT IS WAR etc., which assemble words and idioms having a common conceptual source and a common conceptual goal. As we will see below, this approach is already halfway towards cognitive onomasiology. Furthermore, cognitive linguistics is grounded on assumptions about language quite similar to those of onomasiology (see section 3). One of the aims of the present paper is to discuss the potential of modern onomasiology in light of cognitive linguistics. The second aim is to develop a framework of diachronic cognitive onomasiology. It will be shown that both, onomasiology and cognitive linguistics, reveal their full explanatory power in a diachronic perspective.\(^2\)

---


\(^2\) Diachronic cognitive onomasiology has been conceived in collaboration with Peter Koch and has been theoretically developed and practically applied to Romance languages with our co-workers in the Tübingen Decolar-Project (Paul Gévaudan, Barbara Ventarola, and Antonia Neu). The aim of Decolar (*Dictionnaire étymologique et cognitif des langues romanes*) is a) to list the words for concepts designating body parts as well as human perception and qualities attributed to man in 14 Romance languages and idioms, and b) to describe as accurately as possible their diachronic genesis and to classify them according to the respective source concepts. In a parallel project “Lexikalischer Wandel – Polygenese – kognitive Konstanten” the same method is applied to a larger sample of languages of the
2. From semasiology to onomasiology

2.1. The semasiological and the onomasiological approach to the lexicon

Onomasiological lexicology goes back to the early 20th century’s *Wörter und Sachen*-movement in linguistics, whose intention was to discover the different expressions existing in one or more languages for a given concept and to explain their etymology and the motivations for their creation (cf. Blank, in print a; for a detailed bibliography cf. Quadri 1952). It served also as a methodological background to the great enterprises in linguistic dialectology during the 20th century. This extremely fruitful line of study lost its vitality under the influence of modern (structuralist) semantics, whose view was decidedly semasiological.

The difference between the two approaches can be characterized by the following schema:

(1) Semasiology and onomasiology (cf. Quadri 1952, 168)

```
semasiological perspective

to seize
   TAKE HOLD OF
   SUDDENLY
   TAKE INTO

to comprehend
   PERCEIVE THE
   MEANING OF

onomasiological perspective

to understand
```

From a semasiological point of view (< Gr *séma* ‘sign’), we investigate the different senses, i.e. the polysemy of, e.g. E *to seize*, while onomasiology (< Gr *ónoma* ‘name’) asks for the denominations of, e.g. the concept *UNDERSTAND*.\(^3\) In diachrony, two levels of abstraction should be discerned:

---

\(^3\) Concepts can be linguistically realized by a paraphrase, as done here, or by a simple word. In the latter case, this word has a double function: it first denotes a
individual studies of words or of concepts and more theoretical approaches to types of semasiological or onomasiological processes. Individual semasiological diachronic approaches describe the history of a particular word in time, e.g., how *to seize* has acquired the metonymic sense of ‘to take into custody’ and the metaphorical sense of ‘to understand’, but also its derivatives (*seizure* and *seizing*) and idioms (*to seize an opportunity*). On the level of theoretical lexicology, semasiology is the perspective behind typologies of the mechanisms of lexical innovation, such as metaphor, metonymy, types of word formation, idioms, etc. Onomasiological studies try to discover the different lexical “pathways” through which a particular concept has been designated by going back to the respective source concepts. They focus mainly on continuous changes in the way we express concepts and thus help to discover recurrent schemas for designating a concept or a group of concepts as, e.g., the metaphorical expression of MENTAL PERCEPTION through words for PHYSICAL MANIPULATION (cf. Sweetser 1990: 28–44). The onomasiological perspective is also chosen for typologies of the motives of lexical change (cf. Zgusta 1990; Blank 1999a).

While from a semasiological point of view one can (and maybe should) decide whether one concentrates on cases of semantic change or on word formation, idioms or loanwords, etc., onomasiological studies of particular concepts must always address the whole scale of types of lexical change. Onomasiological change thus includes all types of lexical change (cf. Koch, in print).

2.2 Universality or culture-specificity of concepts?

A fundamental problem of onomasiology arises from the fact that we must first of all define a conceptual system before starting an onomasiological analysis. Where do we obtain these concepts? Are conceptual systems universal or language-specific? Do they have a binary branching structure like

---

mental idea (referred to in SMALL CAPS), but on the other hand, it is of course identical to an existing lexeme in a given language (in *italics*), which indeed is one of the denominations of this concept. The denominations of a given concept are, in fact, more or less synonyms.

4 Cf. section 4, and in more detail Blank (1997) for semantic change, Blank (1998d) for word formation, Blank (1996) for idioms; Blank (1999b) and (2001a) for a comprehensive overview.
the famous “arbor porphyriana” (cf. Raible 1997: 31f.) or must they admit other kinds of structuration? The first question leads straight to the more general question, whether our language determines our concepts and our vision of the world or whether the world determines our language. I cannot discuss this fundamental and indeed more philosophical than linguistic issue here. It seems however clear that any kind of Platonian realism in linguistics is doomed to failure and that speech communities create their own conceptual systems, or in other words a “world” of their own, which is then subsequently verbalized. This is to say that concepts are neither universal nor are they really language-specific: they rather are culture-specific and thus extralinguistic phenomena. In Upper Engadinian Rheto-Romance, e.g., we find the following words for SNOWING (cf. Liver 1989: 792; HWR: s.vv.).

(2) a. neiver ‘to snow’
   b. bischar ‘to snow with small, icy flakes (esp. with strong cold)’
   c. brisclar ‘to snow softly’
   d. cuflar ‘to snow heavily with wind’
   e. cuflergnar ‘to snow softly with wind’

The fact that Engadinian has at least five words where Standard English and Standard German have only one or two, obviously derives from the alpine climate which makes subtle linguistic differentiations necessary. We have access to this “foreign world” by paraphrasing the concepts in question (as we did by defining the five words) and thus we are able to conceive the denoted facts, but, as the example shows, the existence of a simple lexeme or a lexicalized word-formation is good evidence for the existence or the prominence of a concept in a given speech community.

The example allows some conclusions regarding the methodology and the aims of onomasiological research:

1. Literally every referent and every concept can be verbalized by any language. It is, however, more interesting to study which concepts are usually and constantly expressed in a given language (Heger 1964: 514). Only from this perspective can we gain insight into the way a speech community conceptualizes the world.
Words and concepts in time

2. The conceptual system we choose for onomasiological studies should largely correspond to the semantic structure of the envisaged language (Heger 1964: 515). Thus, for analyzing Rheto-Romance we need a more subtle system of meteorological concepts than for analyzing English. Comparative onomasiological studies must define and rearrange the conceptual system according to the language-specific differentiations.

As a consequence, we should distinguish in semantics between an extralinguistic, although not universal conceptual system (what Humboldt has called the “Weltbild der Sprache”; cf. Hallig and Wartburg 1963: 52) and a language-specific semantic structure on the level of the signifieds. This distinction may seem to be too subtle and hypertrophic (cf. the critique in Taylor 1999: 23ff.), but it nevertheless appears necessary when one looks at the examples cited above: the opposition of neiver, bischar, brisclar, cuflar, cuflergnar necessitates a set of intralinguistic semantic features in Rheto-Romance in order to distinguish the words semantically on the level of the envisaged language. These features remain purely extralinguistic for E to snow, where only one intralinguistic feature is needed to distinguish it semantically from to rain or to sleet. The concept to which to snow is linked nevertheless comprises all the concepts named by the five Rheto-Romance words, but this information remains irrelevant on the level of the English semantic system.5

2.3. The structure of conceptual systems

Diachronic semasiology investigates the lexical and semantic development of words. Words have an internal structure, i.e., the synchronic relations between their senses, and a number of external structures, i.e., word class, derivational class, and compounds, lexically related items etc. The simplest semasiological approach to words is the alphabetic order of a dictionary.

A major motivation for onomasiological approaches is the fact that alphabetic lists of words conceal the semantic and conceptual structures of languages (Hallig and Wartburg 1963: 53). So long as single concepts or a smaller group of concepts are investigated it suffices to make sure that the

---

concepts have been well established in the sense of what has been said in section 2.2. Problems arise when it comes to larger onomasiological studies, as they require a full-fledged conceptual system. As mentioned above, simple binary taxonomies like the “arbor porphyriana” would not suffice to describe the complexity of the human vision of the world. As a consequence, the different conceptual systems that have been suggested (e.g. Buck 1949; Dornseiff 1954; Hallig and Wartburg 1963; Schröpfer 1976ff.; Vernay 1991ff.) combine mainly two types of relations:

a) **taxonomy**: relations between concepts which are more or less similar exhibiting a number of common features (e.g. TIGER, LION, LEOPARD, PUMA) so that they can be subordinated to a more general concept which logically includes them (FELINE or CAT).

b) **engnymy** (cf. Koch 2001): a system of concepts that exhibit a sub-type of contiguity, such as part/whole, cause/consequence, producer/product, activity/place etc. Engynomic structures specify mainly cognitive models of knowledge such as frames, scenes or scripts (cf. Koch 1999b; Blank 2001b).

Taxonomic and engynomic structures are synchronic counterparts to some of the ten associative relations that, in a semasiological perspective, are used to verbalize concepts (cf. section 4.1.). Examples for how they combine in onomasiological structures can be drawn from any onomasiological dictionary. In section 2 (Mankind, sex, age, family relationship) of Bucks *Dictionary of Selected Synonyms in the Principal Indo-European Languages* we find a large number of taxonomic structures, such as MAN₁ (Human Being) – MAN₂ (vs. Woman) – WOMAN or PARENTS – FATHER – MOTHER, but also combinations with engynomic structures, such as HUSBAND – WIFE – MARRY – MARRIAGE or engynomic structures alone, such as ORPHAN – WIDOW. The more complex a conceptual structure is, the greater is the role of engynomic structures as, e.g., the frame HEAD, whose different subframes (FACE, ZONE OF THE EYE, ZONE OF THE NOSE, ZONE OF THE MOUTH) and their concepts (EYE, EYEBALL, IRIS, PUPIL, EYELID, EYELASH, EYEBROW) are completely related by contiguity.

---

6 For the importance of these two relations cf. also the network-models of semantic knowledge from Collins and Loftus (1975) to Pustejovsky (1995).
3. Towards diachronic cognitive onomasiology

Traditional onomasiology as well as the 19th century semasiological semantics have been strongly influenced by psychology and thus have more or less explicitly discussed a number of ideas that constitute now the foundation of modern cognitive semantics (cf. Geeraerts 1988). This new line of study has emphasized that human cognition employs several basic mental operations, such as the grouping of contiguous elements into domains, the association of similar and opposite elements, the analysis of complex scenarios into clear-cut smaller scenes, the forming of figure-ground schemas, or the recognition of recurrent elements etc. As these primary mental operations can be considered “human” in a biological sense, it appears most plausible that our languages reveal traces of these principles in the way they verbalize concepts throughout their history.

Empirical studies done in cognitive linguistics give evidence for the universally fundamental character of the HUMAN BODY, of SPACE, of BASIC DIRECTIONAL and PERCEPTUAL CONCEPTS, such as UP/DOWN or LIGHT/DARK, and of some other basic concepts.7 However, many of these studies do not make clear whether they take an onomasiological or a semasiological point of view or even both (such as the studies on the word over and on the concept ANGER in Lakoff 1987), moreover they are often based on a rather small language sample and sometimes just on English. While the first is a methodological imprecision that may obscure the results of the study, the latter risks circular argumentation: basic cognitive concepts are postulated on the grounds of poor linguistic data and their cognitive relevance is then “proved” by finding them realized in the investigated language (Krefeld 1997: 5ff.): one cannot, e.g., postulate that body parts are basic cognitive concepts just because in English they serve as a source of so many metaphors (arm of the law, head of the department, foot of the mountain, heart of the city, etc.). Only if these metaphors prove to be recurrent in a greater sample of related and non-related languages, can one hypothesize a common cognitive or anthropological grounding.

7 A rather eclectic choice of exemplary studies: Berlin and Kay (1969); Rosch (1973), (1975); Rosch et al. (1976); Bierwisch and Lang (1987); Lakoff and Johnson (1980); Lakoff (1987); Johnson (1987); Langacker (1987/91); Dirven and Taylor (1988).
One way to avoid circularity thus consists in broadening the empirical basis, as has done Anna Wierzbicka in search of what she calls "semantic primitives", i.e., concepts that are realized in every human language, such as, e.g., ME, YOU, SOMEONE, WANT, A KIND OF, etc. (cf. Wierzbicka 1994, 1996). From an anthropological perspective, Wierzbicka asks whether a concept is universally verbalized, but she does not ask by what lexical means this concept is verbalized. In combining onomasiology with the main focus of cognitive linguistics, the latter question, however, acquires central interest: if a given concept, e.g., the LEADER OF A GROUP, is expressed the same way in different languages, e.g., by the word for HEAD, or if it is expressed the same way in distant phases of the history of one language, then we may conclude that this way of expressing the concept represents a cognitively salient and privileged conceptualization. In drawing such a conclusion, it is of crucial importance that there be evidence for polygenetic processes in order to exclude as much as possible adstrate influence or developments in an earlier stage of a common ancestor language.

Cognitive onomasiology, thus, requires both, an enlarged sample of languages in order to avoid circularity as well as a deepened insight into diachronic lexical processes in order to understand processes of conceptualization that, over time, have become opaque. Combining diachronic lexicology with onomasiology and applying it to more than just one or a few languages allows us to show, in an empirically justified way, which conceptualizations are proper to a single or very few speech communities versus those that can be found universally and thus may match a biological predisposition of perceiving the world. Cognitive onomasiology hence can procure us deeper insight into the way our mind works. It is important to say that "universally recurrent conceptualization" is not something that must be found in every language of the world nor even in most of them: first, there are always some speech communities that, for some reason or other, prefer a cognitively unprivileged way of conceptualizing a given aspect of the world, and second, several cognitively salient ways of conceptualization may exist in parallel and compete with each other.

The following section 4 gives a short overview of conceptualization principles and of some major lexical processes. Section 5 then offers three case studies which illustrate the range of diachronic cognitive onomasiology.

4. A comprehensive typology of lexical change
4.1. Ten types of associative relations

Revisiting older (and more recent) onomasiological studies, one is impressed by the mass of data and by the wide range of languages and dialects that have been investigated. Traditional onomasiological studies usually list metaphors and metonymies and distinguish between loanwords and semantic loans, but they are less specific about other types of semantic change and usually are unaware of the semantic aspects of word formation and idioms. Yet recent studies in lexicology have not only proved the existence of a broader range of types of semantic change, but have also emphasized the semantic aspect of word-formation, idioms, and other types of lexical change.\(^8\) From an onomasiological point of view, the common denominator of the major processes of lexical innovation is the fact that a speaker tries to verbalize a given concept by associating one or more other concepts that have already been verbalized in the speaker’s language.

A good example to illustrate how one concept can be variously conceptualized is the case of the small piece of wood used for lighting candles, cigarettes, etc.: the usual word in English is *match*, a metaphor from *match* ‘(lamp) wick’ (< OF *mesche*): the new object has been conceived as the functionally and formally similar object. In French we find *allumette*, which originally designated a splinter destined to transport fire. Morphologically, *allumette* is a suffixation of the verbal base *allum* ‘to light’; the suffix adds the sense of ‘instrument for ...’. The inherently associated concept, thus, is that of *to light*. The semantic change from ‘splinter’ to ‘match’ is a case of semantic restriction or specialization, as a match is a kind of splinter. German *Streichholz* is morphologically a compound that combines two contiguity associations: *Holz* ‘wood’ refers to the material matches are typically made of and *streichen* ‘to rub’ refers to the movement one performs when lighting a match. Spanish has two words for *match*: *fósforo* is a loan from Gr *phosphoros* ‘firebringing’ and thus characterizes metonymically the most salient aspect of the object; *cerilla*, literally “little piece of wax”, refers to matches that are made of wax instead of wood; the suffix makes it diminutive.

The example shows that, if we want to understand fully the relation between the target concept (MATCH), the source concept (WICK, TO LIGHT,

WOOD/TO RUB, FIREBRINGING, LITTLE PIECE OF WAX) and the words expressing the target concept (match, allumette, Streichholz, fósforo, cerilla), then we have to develop a semantic typology for all types of lexical innovation, such as semantic change, derivation, compounding, gender change, conversion, the different types of idioms, as well as word blending, agglutination, deglutination, onomatopoeia as well as loan words and their subsequent semantic changes. Recent studies (cf. footnote 8) have shown that, although these types of lexical innovation are formally completely different, they rely semantically on a small set of associative relations between source and target concepts. All associative relations can be reduced to the three Aristotelian principles of remembering, i.e. similarity, contrast, and contiguity.

Of these three principles, similarity shows by far the greatest diversity: The best known process based on similarity is that of metaphor which relates two concepts that exhibit a more or less peripheral perceptual or functional analogy or another common aspect, as e.g. E foot ‘terminal part of the leg, on which the body stands’ and ‘lowest part of a hill, ladder, etc.’. As the type of relation behind lexical metaphor can be found in word formation as well as in idioms, we call this associative relation “metaphorical similarity”.

There are cases of semantic change where conceptual similarity is greater, e.g., in Sp tigre ‘tiger’, which in Southern America means also ‘jaguar’, or E hound ‘dog trained to pursue game’, whose older meaning is ‘dog in general’. In both cases, we can identify similarity between the senses and the concepts behind them, and this similarity is much stronger than in the case of metaphor. Indeed, in addition to similarity, the concepts of DOG and DOG TRAINED TO PURSUE GAME as well as TIGER and JAGUAR also show a taxonomic relatedness, as defined in section 2.3. In the first case, the relation is hierarchical, DOG being the concept superordinated to DOG TRAINED TO PURSUE GAME. In the second case, the relation is non-hierarchical, TIGER and JAGUAR being both equally subordinated to the concepts CAT or BEAST OF PREY. The similarity between both animals is obvious, so that we may call the association between subordinated concepts on the same hierarchical level “co-taxonomic similarity” (cf. also Cruse 1986: 137).

In the case of E hound, the similarity remains more abstract, since, logically, DOG TRAINED TO PURSUE GAME is included in DOG. On the other hand, it is also obvious that the specialized sense is due to a kind of prototypical relation (cf. Blank 1997: 387), so that the similarity between the
prototype (dogs for hunting) and other dogs is the associative basis of this change. The relation between prototypical members, non-prototypical members, and the whole category is even more important in compounding (cf. Blank 1998d: 18–22). Synchronously, however, the hierarchical, taxonomic aspect is preponderant, so that we can call the relation underlying the change of E hound ‘dog in general’ > ‘dog trained to pursue game’ “taxonomic subordination”. The opposite relation, where the process goes from the subordinated to the hierarchically higher concept, is called “taxonomic superordination”.

Similarity is a gradual phenomenon ranging from peripheral similarities to very strong conformity. The highest degree of similarity is “conceptual identity”, i.e. an association of the same concept or the highlighting of prototypical features of a concept, as observed in cases of tautology or in certain types of word formation (cf. Blank 1998d: 9 and 17f.).

Finally, similarity plays a role in an association that is not situated on the conceptual, or better: not only on the conceptual level, but works rather on the level of signifiers. This happens in cases of popular etymology, but also – voluntarily or involuntarily – in word blending, as in E motor + hotel → motel or in L altus + Germ *hauha- → OF haut ‘high’. We call this relation “formal similarity”.

“Contrast” as an associative principle is by far less important than its counterpart similarity. It figures in some cases of word formation and rarely in semantic change and idioms. As contrast among signifiers is the fundamental principle of every semiotic system, only conceptual contrast is relevant to lexicology and to cognitive onomasiology. Although lexical innovation by contrast occurs rarely, we can distinguish two subtypes: “co-taxonomic contrast” (or antonymic contrast), which designates an association of something directly opposed (as in E bad ‘not good’ > E (Slang) ‘excellent’), and “antiphrastic contrast” for cases of association of more indirectly opposed concepts (as in F pensionnaire ‘guest in a boarding house’ > F (Argot) ‘convict’).

Completely different from similarity and contrast are associations by “contiguity”, which is the principle behind the engynomic structures as defined in section 2.3. While contrast and similarity are relations that demand a certain amount of reflection on the concepts involved and are sometimes purely “academic”, as the two concepts may have nothing to do with each other, contiguity relations result from reality insofar as we have learned that

---

9 To the interdependence of contrast and similarity cf. Blank (1997: 142f.).
there is a spatial, temporal, or logical connection between the concepts or that we can presuppose such a connection. We call this type of relation “conceptual contiguity”.

Again, as with similarity, there is a non-conceptual type of contiguity which is relevant for one type of semantic change (lexical absorption or ellipsis) and probably for word formation and idioms, as well. This is due to the fact that simple words are combined to larger lexical units and that they may influence each other mutually. One example: automobiles were first called motor cars to distinguish them from other cars. With time however, motor cars became more common than other cars, so that car was used instead. Semantically, this is an absorption of the sense of motor car into the simple lexeme car, which formally is a part of the complex unit motor car (cf. for details Blank 1997: 288–292). The relation between the parts of complex lexical units is called syntagmatic contiguity.

4.2. A cross-classification of processes and relations

At first glance, the ten types of associative relations may appear rather abstract. And indeed, on the level of lexical description, they are surely more difficult to deal with than the terms “metaphor” or “metonymy”, which describe purely semantic processes in which the word itself remains unaltered. The advantage of the ten relations appears only when it comes to describing the different types of word formation and other types of lexical innovation, where a morphological process (suffixation, compounding, conversion, or zero-derivation) and a semantic process coincide. These processes do not have traditional names, unlike the classical figures of speech. The distinction however is of absolute necessity for a cognitive onomasiology as described above – and, in a larger sense, for any work in synchronic and diachronic lexicology – as it allows a clear specification of the manner of verbalization (compound, idiom, conversion, semantic change, etc.) and of the type of relation between source and target concepts (metaphoric similarity, conceptual contiguity, etc.). As a result, we obtain a cross-classification of the manner of verbalization and the type of relation.

Some examples:

(3) E tea ‘a beverage’ > ‘afternoon meal’

\[
\begin{align*}
\text{manner of verbalization:} & \text{ semantic change} \\
\text{type of relation:} & \text{ conceptual contiguity (= metonymy)}
\end{align*}
\]
(4) a. It *ragazzo* ‘boy’ + *-ino* → *ragazzino* ‘little boy’
suffixation / taxonomic subordination (= diminuition)

b. It *ragazzo + -one* → *ragazzone* ‘big boy’
suffixation / taxonomic subordination (= augmentation)

(5) F *bien* ‘well’ → *(le) bien* ‘the good’, ‘property’
conversion / conceptual identity

(6) G *Wüste* ‘desert’ + *Schiff* ‘ship’ → *Wüstenschiff* ‘camel’
compounding / conceptual contiguity + metaphorical similarity

(7) F *mener qn. en bateau* ‘to fool someone’
syntactic idiom / metaphorical similarity

(8) E *motor + hotel* → *motel*
word blending / conceptual contiguity + taxonomic subordination + formal similarity

Semantic innovation, word formation, idioms, blends, etc., are one way to cope with the necessity of verbalizing a concept. Another way is to borrow the word for the concept from a language that has already verbalized it. At all times and in most languages of the world borrowing has been a very common strategy: according to Walter and Walter (1991: 9ff.), ca. 8000 of the 60,000 words contained in the *Petit Larousse* have foreign origins – not including the borrowings from Classical and Medieval Latin. Loanwords are often subject to further lexical change during or after the borrowing process (cf. Blank 1995: 46–53). A sort of attenuated form of borrowing is the so-called “loanshift” (Haugen 1950), a substitution of a foreign word or idiom by indigenous lexical material. In this case, indeed, all processes described above are possible, with the only difference that the foreign language serves as a model more or less faithfully applied:

(9) E *window* ‘rectangular field in a computer screen containing a specific application’ → G *Fenster*
semantic change / metaphorical similarity (= semantic loan)
(10)  E home page ‘first page of a website, often containing its directory’
      → F page d’accueil
      lexicalized syntagm / metaphorical similarity + conceptual contiguity
      (= loan transfer)

At this point, we can return to the onomasiological perspective, as we are now able to describe precisely the lexical pathways by which the words for MATCH in English, French, German, and Spanish came to mean what they mean today. For this purpose, we use the following schema which goes from a given TARGET CONCEPT, to its SOURCE CONCEPT, In some cases, one likes to go beyond this source concept, which then is a TARGET CONCEPT, for a process that leads to another SOURCE CONCEPT, A case can be made that at the center of each phrase a lexical process links source and target form, and an associative relation links source and target concept:
(11) MATCH in English, French, German, and Spanish

<table>
<thead>
<tr>
<th>✦ TARGET CONCEPT₁</th>
<th>M A T C H</th>
<th>‘short, slender piece of wood or other material tipped with a chemical substance which produces fire when rubbed on a rough or chemically prepared surface’</th>
</tr>
</thead>
<tbody>
<tr>
<td>☪ TARGET FORM₁</td>
<td>E match</td>
<td>F allumette</td>
</tr>
<tr>
<td>❓ PROCESS / RELATION₁</td>
<td>semantic change / metaphorical similarity</td>
<td>semantic change / taxonomic subordination</td>
</tr>
<tr>
<td>➤ SOURCE FORM₁</td>
<td>E match</td>
<td>F allumette</td>
</tr>
<tr>
<td>➦ SOURCE CONCEPT₁ = ✦ TARGET CONCEPT₁,₁</td>
<td>WICK</td>
<td>SPLINTER DESIGNED TO TRANSPORT FIRE</td>
</tr>
<tr>
<td>❩ SOURCE FORM₁,₁</td>
<td>OF mesche</td>
<td>OF allumer + -ette</td>
</tr>
<tr>
<td>✦ TARGET FORM₁,₁</td>
<td>ME mac-che</td>
<td>F allumette</td>
</tr>
<tr>
<td>➤ PROCESS / RELATION₁,₁</td>
<td>loan / conceptual identity</td>
<td>suffixation / conceptual contiguity</td>
</tr>
<tr>
<td>➦ SOURCE FORM₁,₁</td>
<td>OF allumer + -ette</td>
<td></td>
</tr>
<tr>
<td>➦ SOURCE CONCEPT₁,₁</td>
<td>WICK</td>
<td>TO LIGHT</td>
</tr>
</tbody>
</table>

5. Diachronic cognitive onomasiology: three case studies

The following three case studies are intended to demonstrate our conception of diachronic cognitive onomasiology and especially its advances in the precision of lexical analysis. We will see how universally salient conceptualizations can be sorted out from a large corpus of languages, how different conceptualizations consequently lead to a trend in verbalizing
concepts completely different from similar fields and, finally, how the conceptualization itself can vary in the history of one language.

5.1. Universal conceptualizations: the PUPIL revisited

In 1949, the Italian linguist Carlo Tagliavini published an article on “Some names for the PUPIL, an onomasiological study with special regard to the Hamito-Semitic and the African languages” [my translation]. Analyzing the names for this concept in far more than 100 languages and idioms, Tagliavini discovered nine main strategies of verbalizing the PUPIL (the part of the eye), some of them comprising subgroups which contain variations of the main conceptualization. The major source concepts, according to Tagliavini, are BALL/EGG/APPLE, BLACK, CENTER, STAR/LIGHT, NUT/PIP/PEARL, MIRROR, SEE/LOOK, LITTLE MAN/GIRL/BOY/PUPPET, and, more rarely, syllable reduplication (cf. Tagliavini 1949/82; for a short and slightly reorganized overview cf. Blank and Koch 1999b).

Tagliavini’s onomasiological study is ingenious and uniquely outstanding, in that he clearly points out the major source concepts. We learn that there neither exists one single way of conceptualizing PUPIL, nor is there an infinity of source concepts, but a rather strictly limited number of types. From the broad language sample we are allowed to conclude that these are universally recurrent strategies for naming the PUPIL, or, in other words, individual innovations that have been successful insofar as other speakers adopted them because they considered them to be convincing. Furthermore, a kind of empirical hierarchy can be stated: LITTLE MAN/GIRL/BOY/PUPPET is by far the most common strategy in Tagliavini’s corpus (44 idioms), followed by NUT/PIP/PEARL/LITTLE STONE (36), BLACK (24), BALL/EGG/APPLE (21), SEE (16), STAR/LIGHT (14), CENTER (7), and MIRROR (4).10 At the end of his study, Tagliavini states, referring to Bertoldi, that here “distant nations have occasionally met on the identical roads of human imagination” (Tagliavini 1949/82: 568 [my translation]).

10 Two restrictions must be made: 1. the hierarchy maybe is due to Tagliavini’s choice of languages which is not fully representative for the languages of the world; 2. the evidence for some of Tagliavini’s examples seems to rely on singular attestations in dictionaries and sometimes does not hold in light of modern lexicography as, e.g., Arab adaq whose source concept rather is TO LOOK AND GLANCE (cf. DMWA, s.v. adaqa) than SWELLING as postulated by Tagliavini.
From a cognitive point of view, one can predict that, if speakers adopt a new way of expressing the concept PUPIL, with high likelihood they will use one of the strategies discovered by Tagliavini. In this sense, his study is exemplary for diachronic cognitive onomasiology. Nevertheless, we can enhance the value of his study by classifying all lexical and semantic processes that have appropriated words that originally meant, e.g., ‘ball’ or ‘black’, as denominations of the PUPIL, in order to clarify the lexical pathways speakers choose to designate a new concept, as well as the intermediate stages. Applied to Tagliavini’s study this view will add, among others, the following refinements:

1. **Distinction between complex and simple denominations:** There is a difference between more and less explicit denominations as, e.g., Hung szemfeketéje ‘the black (part) of the eye’, Maori karu pango ‘the eye-black’ vs. Coptic kake ‘the dark one’. Although the source concept remains the same, we have, on one side, semantically explicit compounds combining two conceptual contiguities (PUPIL – BLACK and PUPIL – EYE) and, on the other side, a conversion with conceptual contiguity where no explicit reference on the eye is found. The same holds true for Sard candela di ogu ‘candle of the eye’ and Engadinian stallina ‘little star’: while the first case is a lexicalized syntagm that, semantically, combines metaphoric similarity with conceptual contiguity, the second is a diminutive in metaphorical use.

2. **Distinction of different diachronic stages:** This problem is already overt in Engadinian stallina: in a diachronic perspective, we must distinguish a) the formation of the diminutive stallina ‘little star’ (from staila + -ina, lexicalized with this sense, cf. DTL, s.v. staila), which is, semantically, a taxonomic subordination (a LITTLE STAR being a kind of STAR), from b) the semantic change from ‘little star’ to ‘pupil’. The correct source concept of PUPIL in Engadinian thus is not STAR, but LITTLE STAR, which itself refers to STAR.

This distinction is of even greater importance in the following case: Tagliavini’s material shows clearly that, in most of the languages which use the BALL/EGG/APPLE-conceptualization, BALL, EGG, or APPLE are not the source concepts for PUPIL, but rather EYEBALL, which, to some extent, is more convincing from a perceptual point of view (cf. F prunelle ‘little
Andreas Blank

plum’ > ‘eyeball’). The correct source concept for PUPIL in these cases is EYEBALL. This is not just a matter of punctiliousness, but modifies our list of universal source concepts for PUPIL. Nevertheless we can state that Tagliavini found out, without being fully aware of it, that BALL/EGG/APPLE > EYEBALL > PUPIL constitutes a typical conceptual drift. I illustrate this drift with three examples taken from Tagliavini’s corpus:\footnote{The intermediate sense ‘eyeball’ of Sard láddara is reconstructed by us, but is supported by the parallel Pt bugalho ‘gallnut’ > ‘eyeball’ (cf. DES, s.v. gáddara).}

(12)

<table>
<thead>
<tr>
<th>Δ TARGET CONCEPT</th>
<th>PUPIL</th>
<th>‘the opening in the iris of the eye’</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ TARGET WORD₁</td>
<td>Ir uball na suile</td>
<td>Syryenian sin-kol’k</td>
</tr>
<tr>
<td>PROCESS/RELATION</td>
<td>semantic change / conceptual contiguity</td>
<td>semantic change / conceptual contiguity</td>
</tr>
<tr>
<td>SOURCE WORD₁</td>
<td>Ir uball na suile</td>
<td>Syryenian sin-kol’k</td>
</tr>
<tr>
<td>CONCEPT-SOURCE = Δ TARGET CONCEPT₁,₁</td>
<td>EYEBALL</td>
<td>EYEBALL</td>
</tr>
<tr>
<td>□ TARGET WORD₁,₁</td>
<td>Ir UBALL NA SUILE</td>
<td>Syryenian SIN-KOL’K</td>
</tr>
<tr>
<td>PROCESS/RELATION₁,₁</td>
<td>SYNTAGM / METAPHORIC SIMILARITY + CONCEPTUAL CONTIGUITY</td>
<td>compound / conceptual contiguity + metaphoric similarity</td>
</tr>
<tr>
<td>SOURCE WORD₁,₁</td>
<td>Ir uball + suile</td>
<td>Syryenian sin + kol’k</td>
</tr>
<tr>
<td>SOURCE CONCEPT₁,₁</td>
<td>APPLE + EYE</td>
<td>EYE + EGG</td>
</tr>
</tbody>
</table>

Let us finally have a look at the major pattern for denominating PUPIL: at first glance, the source concept LITTLE MAN/GIRL/BOY/PUPPET seems
somewhat strange, but it is explainable as the small reflection of oneself in
the other’s eye. We have, thus, an inseparable combination of similarity
(ONESELF – THE LITTLE PICTURE) and conceptual contiguity (THE LITTLE
PICTURE – THE ORGAN ITSELF). Here again, cases of word formation as,
e.g., Iranian mardom-e tsesm ‘little man of the eye’, Arab sbi-ul’aiin ‘little
boy of the eye’, Kimbundu camóna já méssu, Sp niña de ojo ‘little girl of
the eye’, must be distinguished from semantic changes as, e.g., L pupilla
‘orphan, little girl’, Albanian minzë ‘girl’, or Gr kôrê ‘girl’.12

Concluding this paragraph we can state that a detailed diachronic lexico-
logical analysis of the denominations found in a language sample not
only focuses on the process itself in order to give lexicologically satisfying
results, but sharpens our view for intermediate concepts and thus improves
the precision of the postulated source concepts.

5.2. Cognitive saliency and diachronic drift: TREE and FRUIT

In section 5.1. we have seen that different strategies universally compete
for the verbalization of a concept. Some strategies may, however, rank
higher according to the number of languages in a corpus which use them.
We can then say that in the group(s) of languages investigated a certain
trend towards certain strategies can be detected. In the case of PUPIL, how-
ever, we cannot really explain this trend, unless by saying that the little im-
age of oneself in the other’s eye is more salient than other conceptualiza-
tions.

In the example we will discuss in this section, perceptual and conceptual
saliency is determined by a cultural background which serves as an expla-
nation for the trend we observe. In his study on “tree and fruit” Koch
(1999a) analyzes the denominations of PEAR and PEAR TREE as well as
BEECH and BEECHNUT in 27 languages of the world. As a general tendency
he finds that, as in English, PEAR TREE is verbalized on the basis of PEAR
(13) and BEECHNUT on the basis of BEECH (14):

(13) a. Pt pera → pereira (suffixation / conceptual contiguity)
    b. Turk armut → armut aπacī (compound / taxonomic subordi-
    nation + conceptual contiguity)

12 The latter may of course be absorptions of former complex forms which are not
documented.
c. Pers golabi → derast-e golabi (lexicalized syntagm / conceptual contiguity + taxonomic subordination)

(14) a. Sp haya → hayuco (suffixation / conceptual contiguity)
   b. Jap buna → bunanomi (compound / taxonomic subordination + conceptual contiguity)
   c. Arab ʿan → qamar azʿan (lexicalized syntagm / conceptual contiguity + taxonomic subordination)

The reason for this complementary situation is obvious: in the case of PEAR/PEAR TREE, the fruit is more important and therefore cognitively more salient, i.e., the tree serves as a background; in the case of BEECH/BEECHNUT the situation is inverted: the wood is the figure, the fruit is the ground. The conceptualization of trees and their fruits thus depends largely on our experience and on their corresponding relevance to us. From this perspective, it is not surprising that in the majority of the languages analyzed by Koch the more salient concept is verbalized as a simple lexeme, while the less salient is verbalized as a complex lexeme on the basis of the simple one (19 in the case of BEECH/BEECHNUT, 20 in the case of PEAR/PEAR TREE). Koch states that “new designations of trees and fruits are not created in a totally arbitrary way” (1999a: 343).

Nevertheless, some exceptions of this general rule show that, once again, we have no cognitive “one-way road”, but just a typical drift of verbalization. Lexical solutions which synchronically do not or do not clearly mark this drift are gender change (15), metonymy (16) and the creation of morphologically independent lexemes (17). Especially intriguing is Polish (18) for PEAR/PEAR TREE, with a development that even contradicts the general drift of verbalization:
Words and concepts in time

(15) a. L *pirus* ‘pear tree’ ≈ *pirum* ‘pear’
    b. OGr *ápios* ‘pear tree’ ≈ *ápion* ‘pear’

(16) Russ *grusza* ‘pear’ > ‘pear tree’ (semantic change / conceptual contiguity)\(^{13}\)

(17) Danish *bøg(etræ)* ‘beech’ – *olden* ‘beechnut’

(18) Polish *gruszka* ‘pear tree’ → *gruszka* ‘pear’ (suffixation / conceptual contiguity)\(^{14}\)

The existence of examples that do not fit into the general framework does not weaken the hypothesis of a cognitive grounding of the denominations for TREE and FRUIT, but emphasizes that, as a general rule, naming of concepts is not governed by strict constraints, but rather by a set of higher and lower probabilities from which we can “predict the range of [...] designational options” (Koch 1999a: 343f.). If a language exhibits one of the more untypical strategies, change – if it occurs – will tend towards the more typical solution. Evidence for this historical dynamic in Koch’s sample is given by the majority of Romance languages as well as by Modern Greek: these gave up the Latin and Old Greek gender alternation for PEAR/PEAR TREE (cf. 15) and adopted a more explicit marking of the relation which, of course, follows the typical drift of verbalization:

(19) a. F *poire*, Sp, Pt, Cat *péra*, Engadinian *paira* ‘pear’
    → F *poirier*, Sp *peral*, Pt *pereira*, Cat *perer(a)*, Engadinian *pairer* ‘pear tree’
    b. NGr *apidí* ‘pear’ → *apidá* ‘pear tree’

5.3. Change of conceptual boundaries: from less to more salient categorizations of the ARM

The development described in Romance underlines the existence of a typical cognitive constellation which favours a new conceptualization and verbalization of the PEAR TREE. The concepts PEAR TREE and PEAR themselves

\(^{13}\) As this is, according to Koch (1999a), the diachronic direction of the process, Russian follows the general drift of verbalization for PEAR/PEAR TREE.

\(^{14}\) Polish *gruszka* seems to be a case of diminuitive formation, but the suffix can also possibly have a relational function, as is the case with It *-ina* in *faggina* ‘beechnut’ (← *faggio*).
however remain unaltered. How a conceptual system itself changes under the influence of a cognitively more salient categorization of a conceptual frame is described by Krefeld (1999), once again with examples for change from Latin to Romance. Krefeld first states that there existed no original word in Latin to designate properly the concept ARM as we would conceive it. The three Latin words, armus ‘upper part of the upper arm and shoulder’, (h)umerus ‘upper arm and shoulder’ and lacertus ‘muscular upper arm’, indicate a categorization of the human body that overrides the distinction between the torso and the body extremities. The speakers of Latin thus conceptualized and verbalized the human body in a rather idiosyncratic manner, as shown in the following drawing:

(20) The conceptualization of the human body in Latin (cf. Krefeld 1999: 266)

During the classical period this schema, whose motivations are difficult to understand, was already accompanied by a more modern model that conceptualizes the ARM more as a whole and as an extremity ending with the shoulder joint, not including the shoulder itself. Accordingly, the Greek
loanword brachium ‘forearm’ and ‘arm’ was introduced, while the meaning of (h)umerus was reduced to ‘shoulder’. The Romance languages abandoned completely the older Latin categorization and continued exclusively what Krefeld calls the “torso-extremity-model” (1999: 259f.). Thus, Romance makes a clear conceptual distinction between the ARM as an extremity (e.g. F bras, It braccio, Rum bra=t, Engadine bratsch) and the SHOULDER as a part of the torso (e.g. F épaule, Sp hombro, Pt ombro, espádua, Sard pāla, kočcu). From a cognitive point of view, this appears to be a more salient conceptual segmentation, as the perceptively clear-shapped distinction between the torso and the extremities is respected. Again, we observe the adoption of a more typical strategy, but this time a change in the conceptual structure itself was involved.

6. Conclusion

Diachronic cognitive onomasiology investigates the main strategies that exist in a language sample for conceptualizing and verbalizing a given concept, and it tries to explain them against a cognitive background in terms of salient perceptions, prominence, convincing similarities, etc. It looks for source concepts that seem to be universally recurrent, lays bare the associative relations between source and target concepts, and describes the lexical processes used by the speakers. It thus requires a double framework of associative relations which can combine with virtually any process of lexical innovation.

This theoretical foundation also allows the description and explanation of changes towards a cognitively more prominent strategy as well as of reorganizations of conceptual structures. Furthermore, to a certain extent we can predict which strategy will most probably be taken by speakers when they produce lexical innovations. In this sense, a history of concepts that integrates semasiological change on the level of the individual designation contributes to a better understanding of how reality (or what people hold it to be) is perceived and interpreted and thus may serve as a modest contribution of linguistics to a better understanding of the human mind.

References

Berlin, Brent and Paul Kay
1969  *Basic Color Terms: Their Universality and Evolution*. Berkeley: California University Press.

Bierwisch, Manfred and Ewald Lang

Blank, Andreas


Blank, Andreas and Peter Koch


Blank, Andreas and Peter Koch (eds.)


Blank, Andreas, Peter Koch, and Paul Gévaudan

Buck, Carl D.

Collins, Allen M. and Elizabeth F. Loftus

Cruse, D. Alan

DES = Max Leopold Wagner

Dirven, René and John R. Taylor

DMWA = Hans Wehr
Andreas Blank


Johnson, Mark

Koch, Peter


Krefeld, Thomas

Lakoff, George

Lakoff, George and Mark Johnson
Langacker, Ronald W.

Liver, Ricarda

Lüdi, Georges

Pustejovsky, James

Quadri, Bruno

Raible, Wolfgang

Rosch, Eleanor

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem

Schröpfer, Johannes (ed.)

Sweetser, Eve E.

Tagliavini, Carlo
Taylor, John R.


Vernay, Henri


Walter, Henriette and Gérard Walter


Wierzbicka, Anna


Zgusta, Ladislav

The semantic structure of lexical fields: Variation and change*

David Kronenfeld and Gabriella Rundblad

1. Introduction

Using as our base the lexical field ‘a natural watercourse’, we will attempt a description of the interrelation between the synchronic state of language and the diachronic changes ever evident in it. We will show how the lexical field was structured in the earlier periods of English, and how and (possibly) why it has come to take on its current shape. To achieve our goal, an anthropological approach has been used, whereby the cultural influence on language is used to help delineate the semantic and lexical changes and the reasons behind them.

2. Semantic and lexical fields

Lyons has defined “the sense of an expression” as “the set, or network, of sense-relations that hold between … a lexical expression and one or more other lexical expressions in the same language” (1995: 80). These sets, or as we shall refer to them here – fields, are based on relations of contrast and inclusion.

The theory of semantic fields (or conceptual fields) and lexical fields builds primarily on the work of Jost Trier who was the first to seriously attempt “to introduce Saussure’s principles into semantics” (Ullmann 1962: 7–8). It is uncertain whether Trier distinguishes between lexical fields and semantic fields (see discussion in Lyons 1977: 250–261); Lyons, on the other hand, clearly does, and we shall therefore use his definition here. A

* We would like to thank Eugene Anderson and Robert Moore for their help in distributing and collecting various research forms, and all the students for participating in our research (see Appendices A and B).
semantic field will vary from one language to another and from one period to another, depending on the way the speakers conceptualize the world around them. In order to be able to communicate about concepts, we impose a set of lexemes over the semantic field – a lexical field – but it is possible that one lexical field may not cover all parts of a semantic field. But, even more commonly, more than one lexical field will be used for any one semantic field, resulting in overlaps between fields (both lexical and semantic).

Following Saussure, Trier (1934) argued that individual words acquire their meaning through their relationship to other words within the same semantic field – that is contrast and inclusion – and any extension of the sense of one word would according to Trier automatically narrow the sense of the neighboring words.

Trier’s theory assumes that lexical fields are easily definable closed sets, but as pointed out by Lehrer (1974: 17) and as is evident from the current study, this is not always the case – fields tend to have fuzzy borders. Therefore delimitation of fields and their members is often arbitrary. Some critics have also claimed that field theory is only applicable to abstract fields (Lyons 1977: 257–259), whereas Lyons has claimed the opposite – the theory is more applicable to “concrete conceptual fields, where the lexemes have identifiable denotata, than it is to abstract fields, where they do not” (1977: 259).

We can study semantic fields by mapping the structure of related lexical fields and if we compare a lexical field at a given point in time with the same field at another point in time we can establish any sense changes that have taken place over time. The current paper will try to describe and explain the changes within the semantic field for “a natural watercourse” using a study of its lexical field. First, however, we have to provide the theoretical background to our study.

3. Synchrony and diachrony

As Saussure long ago noted, language, particularly including meaning, is systematic, and the system is a cognitive one based on patterns of opposi-

---

1 Trier’s ideas were soon taken up to be further developed – for further discussion see Lehrer 1974, Lyons 1977: 250–261, Ullmann 1957: 155, 161. See also Berlin and Kay 1969.
The semantic structure of lexical fields

Saussure, as a historical linguist, saw that language change was not a summation of changes in isolated elements (words or sounds) but was instead a change in patterns of related elements — where, as Trier noted, each element is displaced against other elements in the pattern, and thus where a change in one element necessarily occasioned adaptive changes in related elements (1916: 79–100). It was this insight which led to Saussure’s delineation of synchronic linguistics and to his understanding that diachronic change had to be understood as a sequence of synchronic states — and thus to his insight that effective understanding of diachronic change required prior understanding of the relevant synchronic systems.

Linguists after Saussure have mostly applied his insights to the structure of signifiers and the interrelations among them. Signifieds and their structure have been largely ignored within linguistics. Similarly ignored (at least up until Greenberg’s work) have been systematic relations between synchronic states and diachronic change — the issue that we will be addressing in this paper. Linguistic anthropologists, for whom field languages were as much a medium of investigation as a target or topic, were among the few groups to look at systematic relations among signifieds; a major tradition for such work was that of ethnoscience, which evolved into today’s cognitive anthropology (see D’Andrade 1995, Colby Fernandez and Kronenfeld 1981, and Colby 1996 for overviews; see Tyler’s 1969 collection for classic articles; and see Kronenfeld 1996 for later semantic work). In this work, the meanings of terms (morphemes, lexemes, “segregates”)

2 Though, we note our disagreement with Trier regarding the issues of closed sets and automatic narrowing.

3 In his discussion, Kronenfeld includes critical examinations of the componential analysis work (aimed at determining sense relations from the distribution of attributes across referents) of Lounsbury (1964a and others), the earlier prototype work of Lounsbury (1964b, 1965, 1969), Berlin and Kay (1969), and Berlin (e.g., 1972), related work of Rosch and others in psychology (e.g., Rosch et al. 1976), and the fuzzy set approach of Kempton (1981). He relates his theory to earlier work, and delineates where it offers significant advances.

4 As Frake (1962: 31–32) pointed out, for semantic purposes the relevant units are not morphemes; hot and dog, while “good” morphemes, do not explain the meaningful unit hot dog (there is no dog involved, hot or otherwise). Lexemes often work for semantic purposes; hot dog is a single lexeme and is semantically meaningful. But there still remain linguistic entities such as ham ‘n cheese sandwich (in American English) which, while multi-lexemic (it is indeed a sandwich made with
anchored in the structure of the semantic contrasts within their domain – sometimes hierarchical or taxonomic structures (linking contrasts via relations of inclusion), sometimes componential ones (defining terms via the intersection of defining attributes). This work very directly represented an implementation of Saussure’s notions of pattern in the area of word meanings where he clearly had intended it to apply (1916: e.g., 122–137). This work has still mostly been synchronic – though, with some salient exceptions, such as Friedrich (1966, 1970). To Saussure’s notions of pattern, Kronenfeld (1996) has added a delineation of the role played by presumptive (or presupposed) pragmatic relations among prototypic referents within any given domain.

In any study – whether synchronic or diachronic – it is important to understand language’s socially constructed nature. That is, language is grounded in speech communities; children learn language from their exposure to speech in these speech communities, linguistic resources get created or adapted to express the communicative needs of members of these com-

---

5 In this work lexical fields have not been systematically and/or explicitly distinguished from semantic fields. However, actual treatments have generally been clear (if only implicitly so) about the difference, and about, in that context, their focus on semantic relationships. In a similar vein, the distinction between semantic relations among categories and pragmatic relations (whether actual or presumed) among the referents of those categories has largely been ignored – with the price, often, of unnecessary theoretical muddling.

6 Friedrich’s work is noteworthy in the present context for its application of the systematic semantic approach of early ethnosience – along with equally systematic bio-geographical data – to the kinds of analytic problems addressed by the early Wörter und Sachen movement. Those problems, though, differ somewhat from our present concern – by being focused on recovering data on essentially unknown cultures from reconstructed linguistic data. We are focusing more on using the interplay between more or less known cultures and concrete historical linguistic data to examine the cultural and semantic processes that drive lexical change – even through we do, then, try (in the manner of Friedrich) to feed that information back into a richer understanding of the conceptual lives of members of our targeted culture(s).
munities. From this perspective, and joined with insights gained from Greenberg’s work on universals and marking (1966, 1968), has come our recognition of the importance of frequency phenomena, both in connection with marking relations, and, more generally, with the maintenance of terms in a language.

4. Objectives

Adopting this approach, we are taking the lexical field – here, for the domain of natural watercourses in English – as our primary entity of change, and looking at change relative to that lexical field. Such change sometimes can be a matter of specific terms being replaced by new terms within the still existing structure (or paradigm), or of terms shifting their position within the paradigm; other times change can be a matter of changes in the structure itself. We will offer examples of each of these kinds of change.

Our approach to the study of a semantic domain, including its paradigmatic structure, its relationship to the pragmatic and communicative experience of the language’s speakers, to speakers’ cultural givens, to frequency of usage, and to prior structures and givens (i.e., history), derive in large part, though not exclusively, from work in ethnoscience and cognitive anthropology. The particular details of our approach to these issues are based largely on the semantic theory developed by Kronenfeld (1996), stressing

---

7 A given language, such as English, is grounded in a large number of mutually overlapping speech communities. These speech communities variously represent region, class, cohort, occupation, recreational groups, religious groups, civic groups, and so forth. The overlap is represented by the typical belonging of individuals to more than one speech community. The extensiveness of the overlap, joined with shared history (as represented by common experiences of members of the two communities within the larger society, including their shared experience of the common language), keeps the various communities linked together in a pretty much shared language; but the differences among the communities often give rise to varying degrees of dialectical variation. Most individuals wind up doing some code-switching among their various communities – to a greater or lesser degree depending on the social spread involved.


9 Kronenfeld’s treatment of the paradigmatic (in Saussure’s sense, see Kronenfeld and Decker 1979) structure of terminological categories developed out of anthropological work on componential analysis (e.g., Lounsbury 1964a, Goodenough
the influence and “power” of defaults as laid out in Rundblad’s work on markedness (1998). The theory goes considerably beyond earlier componential, extensionist, and prototype work in its attention, inter alia, to communicative function, to the paradigmatic nature of semantic extension, to the varied bases (denotative, connotative, figurative) of semantic extension, and to the flexible but systematic ways in which terms in usage are related to potential referents in the pragmatic world.

Our approach is in some ways similar to the “cultural linguistics” approach (within the context of the cognitive linguistics paradigm developed particularly by Langacker [e.g., 1987, 1991] and Lakoff [e.g., 1987, 1990]), but with important differences. Some of the issues raised regarding earlier ethnoscience work apply here too – including a different view of the nature and role of prototypes – and, obviously, we do not use the Langacker and Lakoff analytic machinery. Some relevant further differences include the following: Kronenfeld’s theory builds on a distinction of semantic relations among terms (cf. Saussurean “signs”) of contrast and inclusion from the pragmatic schemas (of action, connection, function, etc.) that the terms reference. His theory sees semantic extension as a function of paradigmatic contrasts within a domain, and as representing a comparison of the prototypic schema relations implied by the terms with the relevant schema relations entered into by the potential target referents. Following Kronenfeld’s theory (1996), our analysis will build on the distinction of semantic relations among terms (cf. Saussurean “signs”) of contrast and inclusion from the pragmatic schemas (of action, connection, function, etc.) that the terms reference, postulating that semantic extension is a function of paradigmatic contrasts within a domain, representing a comparison of the prototypic schema relations implied by the terms with the relevant schema relations entered into by the potential target referents. We will also stress the sharp distinction of collective representations (including, inter alia, grammar, cultural models [sometimes spoken of as cultural schemas], and shared se-

1956, Wallace and Atkins 1962, and Romney and D’Andrade 1964). The treatment of semantic extension developed out of Lounsbury’s (1964b) work on Crow- and Omaha-type kin terminologies and Berlin and Kay’s work on color, and contrasts with the fuzzy set approach of Kempton (1981) as well as with the kind of prototype approach seen in Rosch’s work (e.g., Rosch et al. 1976). Much of our present perspective is applied to the wide-spread and well-studied, if somewhat special, domain of kinship terminologies in Kronenfeld 2001 (where other relevant kinship work is cited).
The semantic structure of lexical fields

mantic knowledge) from individual schemas and representations (whether of individual action, collective knowledge, or whatever) – and their very different formal properties, functional constraints, etc.

It takes usage to maintain terms in a language, and it takes cultural importance of the referents of the term and the reflection of that importance in conversations to generate adequate levels of usage. Through this route, the vocabulary of a speech community comes to reflect pragmatic conditions – in our present case, the historic uses of watercourses (implying local elaboration of specific types) and contemporary urban disuse (implying a loss of specific types in favor of more generic categories). Local variation in saliently important uses joined with the frequency-of-use threshold (needed to maintain a term in the language of some speech community) can result under some conditions (depending on the social structure of relevant discourse) in a shared structure within the same given language having regional variation in the actual terms which fill given slots in the structure.

In this paper we will first lay out our understanding of the range of contemporary English-language usage regarding natural watercourses, including the range, distribution and use of terms and the semantic structure in which these terms participate. We will then use this contemporary basis as a reference point for interpreting historical data regarding term meanings and frequencies, and thus for giving a richer picture of historical vocabulary change than has previously been offered in this area. We will show, in particular – besides changes in specific terms – a) a change both in the generic or default watercourse term and in the implicit (unmarked) referent presumed for the unmarked term, b) changes in the relative importance of the different attributes which distinguish kinds of watercourses from one another, c) a change in the size of a typical speaker’s watercourse vocabulary, and d) a consistent pattern of variability in the terms which fill one or more slots in a consistently shared (at any given synchronic moment) pattern. We will illustrate how changes in population density and economic activity affect the uses (and relative salience) of different kinds of watercourses, and thus the frequency with which different terms get used in different regions and at different times.

Our research is based on several different types of material. For both present and past varieties of English, we have used various American English and English English corpora (which will be introduced and discussed later). In addition to these corpora (one of which, *The Survey of English* 10 See for example Nerlich and Clarke (1988); Clarke and Nerlich (1991).
Dialects [SED], forms the basis of our understanding and interpretation of present-day English English, data on contemporary American English were elicited for the sole purpose of our investigation (Appendices A and B). These data, while still very spotty, intuitive, and anecdotal, do reflect a systematic approach to data elicitation that, eventually, will yield a much firmer base. We will show the research instruments designed for this elicitation in Appendix A and will illustrate for a couple of specific communities the kind of data they yield in Appendix B. The approach embodied in these research instruments has guided our understanding and interpretation of less formally collected data.

5. The structure of the lexical field

Today’s ‘natural stream’ lexical field contains only a few words. Among the more widespread and frequently used words we find *river, stream,* and *brook.* In addition to these, we find *beck* in English English (especially Northern English English), while *creek* is common in American English. A few more words can be found, if we include highly regional and local words\(^{11}\) – words that are mainly used for (very) small streams.

If we, however, look at the equivalent lexical fields\(^{12}\) for the Old English and Middle English periods, the picture is conspicuously and significantly different. For the Old English period alone, some 100 words of dif-

---

\(^{11}\) Due to the dialectal nature of several of the words we shall deal with here, some readers might not recognize all the words and words might not even be listed in dictionaries of the standard varieties.

\(^{12}\) It should be stressed that we are not claiming that individual speakers had access to or used all of the terms available to them within a lexical field; but we are able to show the developments and decreases in size for the lexical fields through the history of English, from which we can postulate that the vocabulary of individual speakers also must have decreased. There are necessarily differences to be found with regard to both size of vocabulary and the actual composition of the vocabulary that depend on whether we are dealing with individual, regional, and/or social variation. Given the conspicuous lack of words meaning ‘watercourse’ in texts from the earlier periods, we obviously cannot gather information on individual (or, to a large extent also social) variation; but our use of river-names (as described below) does provide a reasonably good picture of regional variation.
different origins, that is, native\textsuperscript{13} Celtic, Latin, and later Scandinavian words, can be found (Rundblad 1998). During the Middle English and Modern English periods, some new words, such as run\textsuperscript{14}, were created from already established elements or borrowed, as in the case of rill\textsuperscript{15}, from other languages, such as French, German, and Dutch, and were added to the field. Even so, the number of words that either have ceased to be used in the sense ‘a natural stream’ or have ceased to be used entirely far exceed the number of new words.

In order to understand the operation of semantic fields (or domains) we have to consider not just the relevant vocabulary of the language as a whole but also, as far as possible, the active vocabularies of individual speakers. Active vocabulary here refers primarily to words in people’s speaking vocabularies, but also to words they have an active and rich understanding of even if they do not use them; excluded are words that they have heard of and vaguely understand, but have no particular sense of – commonly referred to as the speakers’ passive vocabulary. This active vocabulary is important because it provides the basic material from which speakers infer their sense of relevant oppositions, inclusion relations, defining attributes, and pragmatic connections. This speaker’s “sense of relevant…” provides the basis on which each speaker constructs her or his internal representation of language, including the semantic structure; and language itself is comprised of the systematically shared internal representations of structure and content. Since people do not learn the structure directly or independently of their active vocabulary, the semantic structure of the language cannot be larger or more complex than what is implied by general levels of individual usage – even if there be some considerable variation in usage from individual to individual.

For contemporary speakers we can directly assess their vocabularies using questionnaires and interviews (see Appendix A, B and Table 3 in A-

\textsuperscript{13} The number of native (Anglo-Saxon or Germanic) words alone in use during the Old English period was 67, not counting variant forms or derivatives (Rundblad 1998).

\textsuperscript{14} Relevant meanings stated in the OED are: ‘A small stream, brook, rivulet, or watercourse; a channel or overflow’ and ‘A flow or current of water; a strong rush or sweep of the tide’.

\textsuperscript{15} ‘A small stream; a brook, runnel, rivulet’; The word is particularly used of small streams of water that have temporarily formed in soil or sand after rain or a tidal ebb (OED).
Appendix C) or various language corpora (see Footnote 32 and Table 1 in Appendix C). For speakers of past forms of the language we rarely have any direct evidence, but instead have to make inferences based in part on reasoning from contemporary usage and in part on relevant contextual and functional information that we can extract from various sources. We have here chosen to base our analysis on data compiled by Rundblad (1998, Section 10.1) in her work on the use of watercourse words in river-names during the Old English and Middle English periods (see Tables 2, 4 and 5 in Appendix C).16

The active usage of most contemporary individual English speakers seems only to involve two or three of the natural watercourse terms: always river, often stream, and commonly one out of the remaining terms. The particular other term included seems to vary primarily with geographical region – at least, both within England and within the United States. The passive vocabularies of individuals typically are somewhat broader, but still in many cases remain substantially smaller than the vocabulary of the language as a whole. We presume that a similar situation obtained in the past, but the much larger total vocabulary we find for earlier periods of English – especially in a much smaller and more geographically restricted population of English speakers in these past periods – suggests somewhat larger individual active vocabularies than seem presently the case.17

During the earlier periods, villages and villagers depended on streams and rivers for agriculture, fishing, milling and transportation, often having chosen to place their homes near a local stream (cf. Edlund 1987). Today we

---

16 Rundblad compiled a database of river-names using the English Place-Name Society’s (EPNS) volumes covering 14 English shires (Bedfordshire, Buckinghamshire, Cambridgeshire, Devonshire, Essex, Hertfordshire, Huntingdonshire, Middlesex, Northamptonshire, Surrey, Sussex, Warwickshire, Worcestershire, and East and North Riding of Yorkshire). She used the very first occurrence of a word in a river-name as a distinct indication that the word was in use at that point in time; for example, the first occurrence of Old English burna in River Bourne (Ca) dates back to 1577, suggesting that in 1577 this noun was used and known (EPNS 19: 2).

17 We say “seems” here only because our actual systematically elicited full first-hand data is still quite geographically restricted, while our broader firsthand data is sketchy and indirect. We expect improved data in the future to enable us to remove the hedge. We are postulating that similar data for other communities will show, in a similar fashion, alternative patterns in these other communities. We do want to note also that we are making use of published corpora data for modern English, and that these data are quite consistent with our firsthand data.
simply turn on our taps, buy our fish and flour at the supermarket, and take the car whenever needed.

Because of the reliance on streams for food and other essentials, people seem likely to have regarded streams as essential, life-giving and possibly “living” parts of their environment. The number of different uses that streams had and the differing attributes that were critical to each of these uses make it likely that people in these past periods were making more distinctions among kinds of streams than is currently the case for most of modern English. Therefore, the names given watercourses and words used to talk about them appear to offer a clear picture of how salient and important the relation between people and geographical elements could be and probably was. The names, words, and especially the wide range of lexical items available show how streams were viewed and what elements in, around, or about them were essential or significant enough to trigger the creation and to influence the usage of the terms.\textsuperscript{18} The coding of this specific pragmatic or functional information in specific stream terms gives us a way, in any given locale, of assessing relevant relations of contrast and inclusion and inferring relevant distinctive features.

The elements found to influence the perception we have and have had of water channels and the water flowing in them include quality, quantity, speed, sound, movement, and type (cf. Edlund 1987). The use of the most common words during the two earlier periods seem to have been governed mainly by three parameters: quantity, that is the width and/or depth\textsuperscript{19} of the

\textsuperscript{18} Unlike present-day usage, where most streams and rivers have only one name – a name familiar to people living in the areas through which it flows – an official name since people need to agree on names of places and rivers to ensure successful communication, streams and rivers had, in the earlier periods, several different names each as they stretched from one area to another and the names given would vary from area to area, from village to village (Rundblad 1998: 99–100). So whereas rivers and streams today might have a name, specifying (or suggesting) a certain size or appearance, that stands even for parts of the river/stream that do not have the size or appearance specified, the naming practices of earlier time enabled a more “correct” naming in that different parts could receive different names suitable to those particular parts.

\textsuperscript{19} It has been suggested that streams are three-dimensional (width, length, and depth) phenomena, but the importance of depth is debatable. The majority of the words have but one quantitative specification – the width of the stream. Nevertheless, some of the words in the present study clearly entail specifications of depth, as some terms can only be used for deep streams (these terms generally stem from
stream; quality, for example, clear or muddy water; and type (type of riverbed, location etc). Whereas several of the terms were not specified in regard to quality and type, all of them seem to be equipped with some specification of size. In Modern English, we find that the same three elements still influence the usage; however, the one parameter that plays the most prominent part is quantity. Hence the tendency for a size-focused structure of the lexical fields of the previous periods has further developed.

It is worth noting that most of the earlier cultural uses – especially the ones that have subsequently receded in importance – seem to have pertained, primarily, to streams in the middling\textsuperscript{20} to small range. This suggests that for these earlier periods the default or presumptive size, absent further information, was probably middling – versus today’s large.

Today, at least, we can distinguish four sizes: large, middling, small, and very small. Typically, speakers today make a primary distinction between large and smaller. Often the smaller category is subdivided into middling versus small categories. That is, while many speakers simply have terms for large ones and small ones, other speakers make a large, middling, small distinction, but in a way that suggests that the primary division is large versus other, and that the “other” then is subdivided. Very small watercourses are more problematic; in many places they are ephemeral and

\textsuperscript{20} Middling is used to refer to something in the middle size range, with emphasis on the relationship to what is called large or small rather than on any absolute size.
unnamed. The large ones are always rivers. The small (versus large – including middling) ones vary regionally, and can be streams, brooks, creeks, etc. If a middling category is distinguished, it can be as a *stream* versus one of the other small watercourse terms, or it can be some other regional distinction pitting one of our basic small watercourse words either upward against something else, such as *branch*\(^{21}\) (which is seen as mid-sized in some parts of the US) or downward, as the middling term, against something smaller, as *creek*\(^{22}\) or *crick* (a variant form of *creek*, which occurs elsewhere in the US) for American English, and terms like *brook*\(^{23}\) and *beck*\(^{24}\) versus the smaller *sike* (or *sitch/siche*)\(^{25}\) and *stell*\(^{26}\) in English English.

The word *stream* is problematic in a special, structurally interesting, way.\(^{27}\) It can refer to a small to mid-sized watercourse, as indicated above. But, also, for many speakers (and implicitly for some dictionaries – for example, *The American Heritage Dictionary* (AHD)\(^{28}\) and *The Oxford Eng-
lish Dictionary (OED) [cf. Footnote 28]) – stream does not clearly denote any specific size category in the sense that river, creek, brook, etc. do. Instead, in this usage (that is, for these speakers) it serves as a kind of generic ordinary language term for the abstract category that we are speaking of in this paper as ‘a natural watercourse’. Stream’s abstract quality can perhaps be seen also in the extreme rarity with which the word shows up in either watercourse names or place names. The linkage between the two usages can perhaps be seen in the connotative presupposition of middling size that stream (in its abstract generic sense) seems to carry in the absence of any modifier (such as ‘large’) to the contrary or any contradictory contextual information. Usages such as Field and Stream (the name of a magazine) seem consistent with this sense of stream as a generic, but a generic with a default presupposition of middling size. Consistently, stream’s derivative, streamlet, seems in some regions to be preferred for smaller (and especially very small) watercourses.

Stream stems from Old English stréam, and goes back to Proto-Germanic *straunaz from the Proto-Indo-European root *srou ‘to flow’ (OED). The word is highly productive, more productive than any of the other words, since it is often used metaphorically, for example a stream of people and a stream of information; in addition, it has given rise to new words, such as, streamlet, but more importantly to the verb stream.

Possibly, it was the creation of this verb that contributed to or caused the abstractness of the noun by emphasizing the flowing motion of water rather than the quality or quantity of the water, the type of watercourse, the sound, speed or movement (that is, specific movement, as mere flowing is not a specific type of motion) of the stream. It is hard to say where stream fits in in the past; its usage levels seem not to have been greatly different from what they are today, and the historical data on watercourse names suggests that its anomalous status (as often more of an abstract than a concrete term) has been a continuing thing (Rundblad 1998). Though it is the second most frequently used word for watercourses in both the US and in England (see Table 1), it is difficult to establish approximately when the word began to increase in usage as it has not been found in any settlement names (cf. Gel-

29 ‘A small stream; a brook, rill or rivulet’ (OED).
30 Note that among the other terms, the word closest in productivity is river, which is used in almost the exact same way as stream (cf. a river of thoughts etc). However, river lacks a verbal counterpart.
ling 1984) despite the fact that it definitely was in use in Old English (see Table 2).

Stream’s sometimes generic status, and its default presupposition of middling size, both seem somewhat anomalous in the context of current usage patterns. The generic sense of stream is further interesting because in today’s English usage the unmarked (and hence, by one measure, generic) dominant watercourse term clearly is river. We can clearly see in Table 1 that in both the American English corpora, Brown and Frown, and the English English corpora, SED and BNC, river is by far the most used term of all the watercourse terms possible. A similar picture clearly emerges from our data from our Riverside, California and Winter Park, Florida college students that is presented in Appendix B. River is a relatively new term in English, borrowed from French (Old French rivere). Along with river, other (related) terms used for small streams were borrowed: riveret (Old French riverete) and rivulet (possibly of Italian origin, rivoletto) (AHD; OED). The terms riverlet and riverling are, on the other hand, English creations (OED; Rundblad 1998: 88–89). As already mentioned, river is used for large natural watercourses; however, the generic nature of the term and possibly also its high frequency rate has made it possible for the term to be used regardless of size as the superordinate term for ‘a natural watercourse’, as illustrated by the comments made by one of the informants in the SED (Vol 4: 393).

In contemporary usage stream is clearly the next most frequent term after river, and the only term (besides river) whose use seems to run across all English-language speech communities. In this sense, in today’s English, stream can be seen as the generic smaller alternative to river. The data for two local versions of American English (our California and Florida stu-

31 Brown is the American English text corpus from 1961 compiled at the Brown University. Frown is the modern version of Brown (American English from 1992) produced by the University of Freiburg. BNC stands for the British National Corpus, which is the property of the BNC Consortium. SED refers to the collection of vocabulary items elicited in the Survey of English Dialects. It should be noted that the SED’s question to the informants was constructed in such a way that the word river would not be elicited resulting in the conspicuously low number of instances for river as well as the high figures for the other words.

32 ‘A small river or stream; a rivulet, rill, or brook’ (OED).

33 ‘A small stream or river; a streamlet’ (OED).

34 ‘A small river; a brook, stream, or rivulet’ (OED).

35 ‘A riverlet’ (OED).
students) in Appendix B, however, appear more problematic for *stream* than they were for *river*, showing, for those specific populations, *creek* and *stream* to be more or less equivalent (with *brook* more in the background). But, regarding the data in Appendix B and in Table 1 (Appendix C), we must remember that *creek* is very much a regional term in American English, and in English English it is not even considered part of the standard variety (Rundblad unpublished paper 1996); *stream*, then, is left even in the Appendix B data as well as in the American and English corpus data (Table 1) as the major English-wide term for smaller watercourses.

One effect of the modern simplification of English watercourse terms seems to involve a (presently occurring) shift for *stream* from its previous position as the abstract generic watercourse to a possibly emerging future position as the more concrete default smaller (than *river*) watercourse. In this process *stream* appears to be well on its way toward losing its abstract “watercourse in general” quality and becoming the basic small watercourse term (de-abstraction toward concreteness and size specification). In this emerging pattern, among the smaller watercourses, in regions (perhaps rural or suburban ones only) locally specific terms, such as *creek* in American English and *brook* in English English, may be kept as a specific and specified (that is, marked) alternative to the more generic default (or unmarked alternative) represented by *stream*.

In general, the tendency for speakers today seems to be to classify perceptively, focussing on perceptively conspicuous features such as size (and to a smaller extent on speed and type), rather than functionally – that is, classification and labelling in terms of use (of the watercourse) – as was

---

36 Specifically, in Appendix B, Table 3b shows that *stream*, *creek*, and *brook* are considered pretty much equivalent for the Florida students, although Table 1 indicates that *brook* is significantly less familiar to them than are *stream* and *creek*. In terms of what is a “kind of” what, we see (for these students) in Table 2a of Appendix B a hierarchy of inclusion, with *river* including *stream* which, in turn, roughly includes *creek*, which pretty much includes *brook*; fewer of the students said *creek*, *brook*, and *stream* were equivalent (and *creek* equivalent to *river*), but none had *river* as a kind of anything else and only very few reversed the *stream*, *creek*, *brook* ordering. Finally, the Form 2b responses for the California students, though with a very small sample size, show a clear picture for *river* (versus *stream*, *creek*, and *brook* – as familiar and large), but a much less consistent picture of the relationship between *stream* and *creek* (some said they are equivalent, some said that streams are is larger, and some said streams are faster), and no consensual picture of *brook* relative to anything except *river*.
apparently common practice in earlier times. However, in the emerging pattern just described, it is possible that the more specific forms represent a growing recognition of some more specific properties (such as a particular size, or speed, setting, or use); alternatively, the more specific forms could represent a foregrounding of some kind of historic or social attributes of the local region (via the use of terms tied either to the region’s past or to a place/time with which local people identify).

For urban areas the degeneration of the vocabulary for this domain seems to be progressing further than in rural or suburban areas. It seems, perhaps, to be heading toward a two-term form, involving only river and stream, where the two are distinguished only by size. This process in urban areas of losing the relatively specific terms from a domain, and reducing the domain to only a few relatively abstract terms is one that Berlin (1972: 83–84) has noted as common for folk botanical domains.37

Since river was not part of the English vocabulary until the Middle English period, there must have been other terms that filled its large watercourse slot before as the slot hardly can be a new creation (that is the importance of size seems to go back as far as we can tell). Hence, we need to find out which terms might have been used. There are three nouns that are most likely to have been used for large watercourses.38 First, there is the term flood. Based on the etymology of flood (the word stems from Old English flóð ‘a body of flowing water, a river’ and Proto-Germanic *flóduz created from the Proto-Indo-European root pló ‘to flow’), and the cognates of the word (for example, Swedish flod, German Fluss are still used for ‘a river’), we could postulate with some warranty that flood indeed was used for watercourses – particularly for large ones (see Rundblad 1998: 58; EDD II: 418; ASD: 291, 294; EPNS 25: 178; EPNS 36: 188; Löfvenberg 1942: 66–67; OED). However, we cannot say with certainty that the word was entirely restricted to large rivers, though this was according to the OED the preferred usage. Nevertheless, in Old English the term was also used in the sense of ‘the tide’, which was (or soon became) the dominant one and hence flóð was soon used only in this sense. Therefore, we must conclude that due to the change of meaning, flóð cannot have been the

---

37 In the case of Berlin’s study, this progression in the loss of terms/categories does not mirror the order in which the terms emerged; the relatively abstract terms develop relatively late.

38 None of these terms are part of the Modern American English lexical field for ‘a natural stream’ (cf. AHD).
‘river’ of either of the earlier periods as it could hardly have become established as the primary word for ‘river’ at the same time as it was going through an evident and conflicting sense change. Our conclusion is supported by the Figures in Table 2 (Appendix C).

The other two terms are similar in that they both are metonyms stemming from nouns that originally meant ‘water’: Old English ēa and wæter, Modern English ea (or rhee, yeo) and water (see Rundblad 1998: 57–58, 62; AEEW: 381; ASD: 223, 1160; EDD II: 223–224; EDD VI: 403; Löfvenberg 1942: 55–56; OED). Given their original meaning, the terms could naturally have been applied to streams of all sizes. But it seems reasonable to presume that it would take quite a large amount of water to trigger the use of these terms, which would mean that apart from being used for rivers, that is large streams, they were used for seas and lakes. In the case of Old English ēa, there is also its cognate Old Norse word á, which has been borrowed into English (see Rundblad 1998: 57–58; Ekwall 1928: 1, 138–139; EPNS 25: 1, 142–143; OED). This term, unlike ēa, was used for streams (that is mid-sized watercourses) rather than rivers, a use similar to that still present in the Scandinavian countries. Given the co-use of the two nouns in England, it is likely that the inevitable blend of the words caused Modern English English ea to be applicable to both large and mid-large streams.39

Table 2 states that, during the earlier periods, ēa and water indeed were among the most frequently used terms; ēa was often more frequent than stréam. Nevertheless, they seem to have enjoyed nothing like river’s dominant status. Instead, the equivalent to contemporary river in terms of usage for the earlier periods seems for a long time to have been brook, and before that bourn(e)/burn (using, here, their modern spellings), but both were generally mid to small size.40

Brook is still one of the most frequently used words in English English, but though it is still part of the active vocabulary in America (especially in the Northeast (AHD)), its presence is considerably less than in England (see Table 1). The word is Old English bróc meaning ‘muddy, deep-cut, opaque watercourse with clayey riverbed’ (Rundblad 1998: 54–56; see also 

---

39 Most likely the acceptance of ēa for middling size was facilitated by the dominance of river for large streams.
40 According to the OED, the ancestor of bourn(e)/burn (Old English burna (see below)) could in Old English be used to denote ‘a large stream, river’; however, the EPNS (25: 63–64) gives no indication of such a usage.
The semantic structure of lexical fields

Cole 1991: 37–47). The word originally meant ‘fenland, marsh’ though this meaning is entirely unaccounted for in English records and hence the sense development of the word may have taken place already before the Anglo-Saxon invasion (OED). In Modern English, the word simply stands for ‘a natural small stream’, and consequently the word must have gone through yet another sense shift, which most likely was completed in the late Middle English period, whereby the term became more general and more easily applicable. The new applicability of the word seems to have boosted its usage and is thus the reason why the term rapidly increased in usage (an increase that was initiated already during the later part of the Old English period). As have stream and river, brook has given rise to new creations, used of small streams: brocket and brooklet (OED).

Before we can fully account for the shifts and changes of brook and the impact these have had on our lexical field, we need to introduce bourn(e) (Southern English English and American English) or burn (Northern English English). The Modern English word (which is highly dialectal) stems from Old English burna meaning ‘a bubbling or running watercourse with clear water (and gravelly bed)’ (Cole 1991: 45), which towards the end of that period came to be used in the more general sense ‘a brook, stream, river’ (Rundblad 1998: 56–57).

Hence, the semantic shift (or rather shifts, since the very first meaning of the word seems to have been ‘spring, fountain’ (Löfvenberg 1942: 20–21; OED) of burna matches that of bróc. Thus, we are faced with two words that a) went through a metonymical transfer (from ‘marsh’ and ‘spring’ respectively), b) had a highly specific meaning during the Old English period (‘marshy watercourse’ and ‘clear watercourse’), and c) went through a second shift towards a general unspecified meaning, ‘a stream’.

---

41 On the sense developments of brook and the change in usage patterns, see Rundblad 1998: 54–56, 105–121.
42 Though in Kent and Sussex it is still used to denote ‘a water-meadow’ (EDD I: 414).
43 Both words has the meaning ‘a little brook, streamlet, rivulet’ (OED).
44 In Northern English English it means ‘a (small) stream, rivulet, brook’ and in Southern English English ‘an (intermittent) stream, a brook’ (EDD I: 360, 451; OED), whereas it is not really used in the Midlands. It should be noted that the term is not part of either standard American English (bourn(e) is very infrequent in American English and more part of the passive vocabulary of its speakers (cf. Table 1)) or Standard English English.
Despite the similarities between the words in the past, today one (*brook*) is still frequently used, both in English English and American English, whereas the other one (*bourn(e)/burn*) is seldom used in either (and especially not in the more standardized varieties of these two national dialects). We will return to this.

However, there is another difference that needs explaining; during the Old English period, *burna* was more frequently used than *bróc*, but in the Middle English period the positions were reversed, and *brook* had come to replace *bourn(e)/burn* (see Table 2). The reasons behind the changes in usage and the replacement are to be found in the colonization patterns for the Old English and Middle English periods (Rundblad 1998: 168–171). In her study of names, Cole (1991) found that the areas first colonized in England mainly contained watercourses where the water was clear and the bed was sandy or gravelly. In fact, the appearance of these watercourses, most likely contributed to the very reason why these areas were colonized first as the water in these streams must have been purer and fresher. Hence, the best word to use for such watercourses must have been *burna*. Thus, *burna* came to be the default word in Old English, the equivalent of today’s *river*. After the first colonization, the areas containing burns were already settled and the people had to start colonize less desirable areas. These areas also contained watercourses (or they would not have been colonized), but these often flowed through fenlands and the water, though drinkable, was less appealing in that it often was opaque and muddy. The new colonization pattern caused *brook* (Middle English *broke*) to be used more and *bourn(e)/burn* (Middle English *burne*) less as this term no longer was felt to be generally applicable. As a result of the shift in frequency of use, the replacement was completed, leaving *brook* the default. It should be stressed that the replacement of *bourn(e)/burn* by *brook* was completed before the second semantic shift of the words toward generalization.45

In regard to the feature specification for our terms, we can state that unlike today’s *brook* and *bourn(e)/burn*, which only are size-specific, *burna* and *bróc* had both qualitative and quantitative specifications, as well as specification for type and possibly speed (watercourses stemming from

---

45 Possibly due to the higher frequency rate of *brook*, this term went through the shift earlier or faster than *bourn(e)/burn* (Rundblad 1998, 54–57, 168–171; see also Ekwall 1928: 43; EPNS 25: 63). Or possibly, *burna* was replaced because it did not acquire a general meaning at the same speed as *bróc* did. That the two events are related is clear, but the exact order of them is hard to prove.
springs generally move faster than those originating from marshes). Of these features, given our discussion of the uses of the words, the most prominent and important one is quality. The colonization of England focused on the availability of fresh water as it was essential for survival.

Despite its acquired high frequency, brook failed, however, to remain the default word in English; in Modern English we find river and stream in that position. As already mentioned, stream is today the second most frequently used word for watercourses. In the earlier periods it was still among the more common ones, but its frequency rate never rose above 8.1% (see Table 2). There exist three reasons for suggesting that stream was not terribly frequently used in forming particular proper names. a) Burna and brook alternated as defaults in proper names since they were more applicable (given the particular importance at the time of water quality) than stream, which must have been too general in these local naming contexts—even if it must have been exactly this same general sense that promoted and increased its general usage. b) Most rivers had proper names before the Anglo-Saxons arrived and the mid-sized streams were named soon afterward; thereafter, even though most watercourses consistently received new names and additional nicknames these often remained unrecorded. c) River was adopted from French, and, in the majority of the cases, added to the existing (Celtic) names. Hence the use in proper names of stream (as well as the other terms) was limited. Nevertheless, the overall use of stream came to expand dramatically (facilitated by its productivity and applicability), whereas the use of brook decreased, thereby letting stream take over the place as the word for ‘a natural stream’ in American English and in Standard English English.

Like the majority of watercourse terms in Modern English, stream comes only equipped with specification as regards size, that is a quantitative specification. Unlike many of the other terms, stream already lacked during the earlier periods specifications of the other types that were available and used, such as quality and type, as in the case of burna and bróc.

Another watercourse term of importance is beck, which as shown in Table 1 (Appendix C) is the fourth most frequently used word in English English after river, stream, and brook (though the term is not part of American English) and the most frequently used term in northern England (SED).

---

46 As pointed out by Rundblad (1998: 99-100), many stream and rivers have denominations such as The River or The Brook, but it is doubtful whether these names should be regarded as names or even nicknames.
The word was borrowed from Old Norse, *bekkr*\(^{47}\), and as such it was at first mainly used in the Danelaw areas of Northern England. In river-names (and place-names) it occasionally came to replace Old English *bróc*, *burna*, and *bece* (with which it is cognate\(^{48}\)) (EPNS 25: 26). The word can first be found around 1300, but, like so many other Viking words, must most likely have been adopted much earlier. It is still mainly restricted to the northern parts of England where it is used for fairly large watercourses, but unlike *burn*\(^{49}\), it is applied especially to watercourses with a stony bed and to mountain streams (hence a *beck* is a fast-flowing watercourse) (see Table 3; cf. SED, EDD). This latter special sense of *beck* probably springs from the competition between *beck* and *burn*, a competition that required and triggered (or – grew in response to) a more subtle distinction in semantic content. Hence the features distinguished for *beck* are and were quantity, type, and possibly speed as mountain streams tend to be fast-flowing. Unlike *brook* and *bourn(e)/burn*, but like *stream*, the feature specification of *beck* has remained intact.

In Table 4 and 5 (Appendix C), we have tried to show the relation between the words and their development in regard to feature specification. We will here have a closer look at Modern Northern English English. Some terms have already been mentioned: *beck* is the default term, despite its specification for type (and speed) – a northern watercourse is (stereo-) typically a *beck*; *brook* is a non-specific term that seems to be used for streams that are not typical becks or burns, with which it shares its size-specification; *burn* is preferred for flatland streams; finally, *stream* is another general term, seemingly used for watercourses somewhat smaller than brooks\(^{50}\) (and hence becks and burns), unless used as the abstract default.

If we look at words for watercourses smaller than becks, brooks, and burns (and possibly streams), we first find *sike/sitch* which is commonly used for (very) small streams, especially of those flowing through flatland, and can be used for both natural and artificial watercourses. As illustrated

\(^{47}\)'A stream, brook, rivulet’ (EPNS 25: 26; OED).

\(^{48}\)However, the cognate *bece*, which had the specific meaning ‘a brook, beck or small rapid stream’, no doubt had some influence on the outcome and specification of sense for *beck* too (cf. EPNS 25: 23–24, 26; Rundblad 1998: 54).

\(^{49}\)As already stated, *burn* is used in the sense ‘a (small) stream, rivulet, brook, especially one running through flatland’ in Northern English English (EDD I: 451; OED).

\(^{50}\)See below.
by the SED informants, it is contrasted with *beck*, not with *burn* as could be expected, given the common denominator flatland. This could be interpreted as reinforcing the claim for *beck*’s default status, and/or as a result of the fact that sikes can be tributaries of becks. The word goes back to Old English *sic* (especially the forms *sitch/siche*) and Old Norse *sík* (especially the form *sike*) both used in the meanings ‘marsh’ and ‘watercourse’\(^{51}\), where the latter is an extension of the former, following which the semantic definition has remained unchanged.

Whereas sikes only occasionally are seen as tributary, gotes are commonly used for tributaries of becks, streams, and brooks (cf. SED). Another difference between the two lies in the fact that unlike *sike/sitch*, which is now equally used for natural and artificial streams, *gote*, ‘a small artificial watercourse leading to a mill or reservoir, a mill-race or water-channel; the outlet from a stream’ (EDD II: 691), has gone through a semantic shift and is now increasingly preferred for artificial watercourses.\(^{52}\) The change in feature specification for artificiality can be seen in Table 5 where T stands for ‘type’, where no brackets signal that the word is equipped with a semantically and pragmatically significant specification as regards the type of riverbed, whereas brackets signal that there is an indication that the type of riverbed might be (for example, has been and still might be) significant. We can also see how type specification works for *dike* and *drain* which are both used for artificial watercourses, though both can (very) occasionally be used for natural ones too.

Based on Table 3, we could say that in English English, depending on region, *beck*, *brook*, and *stream* are the reference points of all the other words; hence burns, creeks, dikes, drains, gills, gotes, sikes, etc. are kinds

\(^{51}\) Old English *sic* stems from Proto-Germanic *sík*- meaning ‘marsh; stream’ and was used only for very small watercourses (Rundblad 1998, Section 5.8). Jonsson claims that Old Norse *sík* was used of small watercourses in certain areas of Scandinavia, suggesting that the meanings ‘marsh’ and ‘watercourse’ might be derived from an original meaning ‘trickling water’ for Proto-Germanic *sík*- (Jonsson 1966: 272–274). Rundblad (1998, Section 5.8) points out though that there are no instances found in support of an original meaning ‘trickling water’. Nevertheless, the fact that, in England, both *sic* and *sík* were used of watercourses but not of marshes could support the claim that PGerman *sík*- might in fact have been polysemous and ‘might have influenced the second meaning as regards its applicability to both natural and artificial watercourses’ (Rundblad 1998, Section 5.8).

\(^{52}\) Modern English *gote* is Old English *gotu* ‘a watercourse, channel, stream’ (see Rundblad 1998, Section 6.2.2.3; ASD: 496; EPNS 25: 206, 212; OED).
of becks, brooks, or streams. These three words are also used for the larger types of watercourses, or rather the largest watercourses excluding watercourses of the river size (the lack of data for the very largest watercourses and the terminology for them is due to the nature of the data collection for and focus of the SED). Hence, burns, creeks, dikes, drains, gills, gotes, sikes, etc., are all smaller than becks, brooks, and streams.

Of the three words, stream seems strangely enough to be used for somewhat smaller watercourses than the other ones. However, this may be an illusion as we also need to reckon with the fact that brook is, in regional speech, more frequent (see Rundblad 1998) – and that, when people are asked for detailed information on watercourse terminology, reasons for usage based on social or communicative norms (which are generally hard to define) may be verbalized as semantic distinctions where no distinctions exist. However, it is also possible that watercourses that do not quite adhere to the general requirements of, or features for, the smaller watercourses simply get the label stream.

Also, it seems as if stream and brook are more often contrasted than stream and beck. This could suggest that the two terms, which are the most frequently used ones and the basic ones in the southern dialects (and in Standard English English) might be in competition and that the increasing use of stream (see Rundblad 1998) may trigger a change, a change that, as in the comment from Middlesex (see Table 3; SED, Vol. 3: 421), might alter the feature description of the terms resulting in stream becoming “larger” and brook “smaller”. (See also the discussion of the current replacement of brook by stream in Rundblad forthcoming). Beck, being “the” northern word for mid-sized watercourses, seems to suffer less from a possible past/future competition with stream.

Of the smaller watercourses, the main feature of the terms seems to be that they are small; the only contrasting between the words in the form of size occurs between dike and drill where the latter is the smaller. Otherwise the words contrast in form of the type of watercourse.

6. Conclusion

The various issues we have discussed point to a stronger version of the suggestion that we have already made above – that in the past size was not the predominate attribute of watercourses that it is today. Size was relevant, but the default size was middling, as opposed to today’s large. Middling streams were important in ways indicated but also were much more com-
mon in people’s experience than were large streams, and were much more variable (from one to the other in their uses) than were large ones. The previously much greater variety of watercourse terms, almost all in the middling to small size range, suggests strongly that other information than size was driving the vocabulary elaboration. Many of the terms used in earlier times (and the streams they were used for) were restricted to specific uses, such as milling and drainage;\(^{53}\) there are, for example, Old English *\textit{flé-ama}^{54}\) and (ge)\textit{lád}^{55}\) (English English \textit{fleam} and \textit{lade/load/lode} respectively), as well as English English \textit{ditch/dike} which are still used for drains (both of natural but primarily artificial origin) (Rundblad 1998: 59, 62, 97–98). Other terms suggest connections restricted to fishing, as in the case of Old English \textit{stell} which originally was used in the sense ‘fishing pool’, but which later was metonymically extended to ‘fish-catching place in a river’ and to the present use ‘a brook’ (Rundblad 1998: 94–95).

The active vocabularies of individuals seem very likely to have embodied some other salient contrasts in addition to today’s size one; the considerations we have raised, joined with Rundblad’s data, suggest moreover that some of the other contrasts might have been considerably more foregrounded than the size one. The varying attributes that Rundblad has found to be associated with these various watercourse terms suggest what is likely to have been the content of these contrasts.

These, in turn, offer us some potential insight into the local livelihoods of the people who lived in the various regions where one or the other of

\(^{53}\) Regarding terms used in the specific sense ‘drain’, please note that several of these terms do not fall within the category of natural watercourses but are artificial ones and hence will not be discussed further.

\(^{54}\) It has been suggested that Modern English \textit{fleam} stems from Old English *\textit{flé-ama} (from the Old English verb form *\textit{fléaman} ‘to flow’) and Middle English \textit{flem} (Rundblad 1998, Section 6.2.3.1; EPNS 25: 175–176; EDD II: 394). Originally the word was applied to natural watercourses and especially those that worked as mill-streams, but today it is used in the sense ‘an artificial channel, watercourse, mill-stream’, but is still applicable to natural watercourses in some dialects (EDD II: 394; OED).

\(^{55}\) \textit{Lád} ‘a channel or waterway leading to a mill’, \textit{load/lode} ‘an artificial channel for water; a watercourse, drain; canal; an aqueduct, channel; an open drain in fenny districts’, go back to Old English (ge-)\textit{lád}, is ‘path, road, passage over a river, watercourse’ (Rundblad 1998, Section 5.6; ASD: 406, 604–605; DEPN: 284; EDD III: 498, 639; Ekwall 1928: 234–235; Löfvenberg 1942: 248; Kristensson 1970: 104; OED).
these middling terms became dominant. This variability in the linguistic effects of local experience would seem, in turn, to explain the great variability we find in contemporary English fillers of the small to middling stream slot. As the local economic importance of smaller streams receded, the frequency with which their labels were used decreased, and often became insufficient to support more than one active small-stream term in any particular local usage (or speech community). That single term that survived would most likely be whichever of the earlier range of terms whose usage had been most frequent (and, hence, which term had been unmarked). The result of this process of vocabulary contraction would be to leave a small to middling stream slot with a lot of regional variation, in a situation where each local term in situ in England represents a little information regarding local economic history. Colonial usage, obviously, would reflect the predominant source areas for immigration, but also, in new coinages or elaborations (such as, perhaps, crick as a derivative of creek, or branch in some of its local senses), might reflect something of local conditions at the time of settlement as well.
Appendix A

Research questions

Each form opens with some general information about the study (but nothing that would be leading questions). Each form closes with some basic background information, including place of birth, where resided, native language, and languages spoken.

Form 1

1) Please list all the English words you know for kinds of watercourses (rivers, streams, etc.).
2) For each word, please give me a couple of sentences using the word.

Form 2a

The actual lists of words in the following questions are picked to cover terms known to be in local use in the area, plus maybe a few questionable ones. The lists, thus, vary from region to region.

1) For each of the following words for kinds of watercourses, please indicate whether 1) it is in your active speaking vocabulary, or 2) if not, whether you know what it means and understand it clearly when others use it, or 3) if not, whether you are at least vaguely familiar with it but unsure what it means, or 4) if not, whether you are totally unfamiliar with it. Put a check in the proper column.

<table>
<thead>
<tr>
<th>term</th>
<th>1) active</th>
<th>2) understand</th>
<th>3) heard of</th>
<th>4) unfamiliar</th>
</tr>
</thead>
<tbody>
<tr>
<td>river</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2) Background: dog is a kind of animal, BUT an animal is NOT a kind of dog. Animals include dogs, BUT dogs do NOT include animals. A cat is NOT a kind of dog AND a dog is NOT a kind of cat.

For each of the following pairs of watercourse terms, please indicate if either can be considered a kind of the other. In each case tell me which is a kind of which – or tell me “neither” if the situation is like cats and dogs. For each of the following pairs, write the MORE inclusive term in the blank; e.g., if, for you, a stream is a
kind of river, then write in riv, or if, for you, a river is a kind of stream then write in str; write n in the blank if neither is a kind of the other. If you don’t know, simply write in a question mark.

river - stream _____ river - creek _____ etc.

3) Background: Cat is opposed in meaning to dog; that is, represents an alternative (or opposed or opposite) kind of animal in at least some context. If cat is opposed in meaning to dog then it automatically follows that dog is opposed in meaning to cat. Dog is NOT opposed in meaning to animal – since it is a kind of animal. Dog is NOT opposed in meaning to adult or to brown – since these are just logically unrelated entities (even if they may possibly apply to dogs). Kitten and kitty are largely synonymous; that is, they mean pretty much the same thing.

For each of the following pairs of watercourse, please indicate whether or not (“yes” or “no”) the two terms are opposed in meaning to each other. Write “same” if they seem to you basically synonymous. Again, a question mark if you don’t know.

river - stream _____ river - creek _____ etc.

Form 2b

1) For each of the following pairs of watercourse terms, please indicate briefly what you understand to be the difference in meaning (or usage) between the first term and the second term. If they refer to different kinds of watercourses, describe the difference; if they are used in different regions or dialects, describe that difference. If you think you know the answer for some pair, but are unsure, then answer the question for that pair but add “unsure” to your answer. If you absolutely don’t know, then just leave it blank.

river - stream ______________________________________
river - creek ______________________________________
etc.
Appendix B

Examples of data analysis

Form 1

Who the respondents are, where from – native languages plus other places and languages.

1. Calif-English
2. Calif-English plus Costa Rica, Italy-Spanish, Italian, French
3. Calif-English
4. Calif-English plus Spanish
5. US-English
6. US-English
7. US/Washington-English plus Washington, California, Hawaii

1) Terms

Notes: The numbers refer to the respondents. n = 7, these are a collection of mixed ages from Riverside, in suburban Southern California.

<table>
<thead>
<tr>
<th>Actual natural watercourses</th>
<th>total #</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>River</em></td>
<td>1,2,5,7</td>
</tr>
<tr>
<td><em>Stream</em></td>
<td>1,2,5,7</td>
</tr>
<tr>
<td><em>Creek</em></td>
<td>1,2,7</td>
</tr>
<tr>
<td><em>(Crook – as variant of Creek)</em></td>
<td>7</td>
</tr>
<tr>
<td><em>Brook</em></td>
<td>2</td>
</tr>
<tr>
<td><em>Flow (as a verb)</em></td>
<td>6</td>
</tr>
<tr>
<td><em>Bayou</em></td>
<td>7</td>
</tr>
</tbody>
</table>

Ambiguous – either ad hoc stream or small watercourse

| *Trickle*                  | 2       | 1   |

River features – systemic

| *Tributary*                | 2,7     | 2   |
| *Drainage*                 | 2       | 1   |
Stream features – places on the stream

<table>
<thead>
<tr>
<th>Feature</th>
<th>Column 1</th>
<th>Column 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>1,7</td>
<td>2</td>
</tr>
<tr>
<td>Artesian spring</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Wash</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Waterfall</td>
<td>2,7</td>
<td>2</td>
</tr>
<tr>
<td>Cascade</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Rapids</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Riffles</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wailele (‘waterfall’ – Hawaiian)</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

2) Sentences

Respondent 1

_____ flow from the mountains to the sea.
_____ overflowed its banks after the heavy rain.
_____ that runs by my house sounds melodious.
_____ changed its location after the earthquake.

Respondent 5

_____ moved very swiftly.
The tranquil _____ flowed by the rock.

Respondent 7

Let’s go fishing down by the _____.
Let’s catch crawfish in the _____.
Cross the _____. to get to the other side.
Swim in the _____.
The _____. flow into the river.

Form 2a

Data from Professor Robert Moore – collected from students at Rollins College, Winter Park, Florida. Semi-urban, South Florida.

Question 1 / Table 1

The number under each column is the number of students who claimed that degree of knowledge for the row term. n = 30.


Some consistency measures:

1) River and Stream: which was listed as better known than the other

<table>
<thead>
<tr>
<th></th>
<th>1) active</th>
<th>2) understand</th>
<th>3) heard of</th>
<th>4) unfamiliar</th>
</tr>
</thead>
<tbody>
<tr>
<td>River</td>
<td>29</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream</td>
<td>27</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creek</td>
<td>23</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brook</td>
<td>12</td>
<td>16</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Branch</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Bourn</td>
<td>1</td>
<td>3</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Rill</td>
<td>1</td>
<td>1</td>
<td></td>
<td>28</td>
</tr>
</tbody>
</table>

The number indicates the number of students (based on an n of 30, but only using those cases where the student claimed either active knowledge or “understanding” of both terms) who said the column term was a “kind of” the row term. (*) marks the diagonal.

Question 2 / Table 2a

The number indicates the number of students (based on an n of 30, but only using those cases where the student claimed either active knowledge or “understanding” of both terms) who said the column term was a “kind of” the row term. (*) marks the diagonal.)
Question 2 / Table 2b

The numbers indicate the number of students (based on an n of 30, but only using those cases where the student claimed either active knowledge or “understanding” of both terms) who said neither was a “kind of” the other. * marks the diagonal.

<table>
<thead>
<tr>
<th></th>
<th>River</th>
<th>Stream</th>
<th>Creek</th>
<th>Brook</th>
<th>Bourn</th>
<th>Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>River</td>
<td>*</td>
<td>27</td>
<td>24</td>
<td>20</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Stream</td>
<td>*</td>
<td>19</td>
<td>16</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Creek</td>
<td>2</td>
<td>*</td>
<td>15</td>
<td>1</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Brook</td>
<td>3</td>
<td>2</td>
<td>*</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Bourn</td>
<td></td>
<td>1</td>
<td>*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The numbers indicate the number of students (based on an n of 30, but only using those cases where the student claimed either active knowledge or “understanding” of both terms) who said they did not know.

<table>
<thead>
<tr>
<th></th>
<th>River</th>
<th>Stream</th>
<th>Creek</th>
<th>Brook</th>
<th>Bourn</th>
<th>Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>River</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream</td>
<td>3</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creek</td>
<td>6</td>
<td>9</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brook</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bourn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Branch</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Based on their great unfamiliarity to these students, rill and bourn were dropped from the remainder of the analysis.
The number above the diagonal indicates the number of students who said “yes” the terms were opposed, while the number below the diagonal indicates the number of students who said “no”. The n for each cell can be arrived at by adding the yes’s and the no’s for each pair and then adding to that the number of “same’s” and non-answers (as shown in Table 3b). Because of the unequal n’s, the relevant insights have to be gleaned from the ratio of yes to no answers, the ratio of “same” to yes or no answers, and the ratio of “?” to anything definite.

<table>
<thead>
<tr>
<th></th>
<th>River</th>
<th>Stream</th>
<th>Creek</th>
<th>Brook</th>
<th>Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>River</td>
<td>*</td>
<td>4</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Stream</td>
<td>18</td>
<td>*</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Creek</td>
<td>14</td>
<td>15</td>
<td>*</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Brook</td>
<td>15</td>
<td>7</td>
<td>5</td>
<td>*</td>
<td>5</td>
</tr>
<tr>
<td>Branch</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>*</td>
</tr>
</tbody>
</table>

The number above the diagonal indicates the number of students who said the two terms were the “same”, while the number below the diagonal indicates the number who did not know (“?”).

<table>
<thead>
<tr>
<th></th>
<th>River</th>
<th>Stream</th>
<th>Creek</th>
<th>Brook</th>
<th>Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>River</td>
<td>*</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Stream</td>
<td>*</td>
<td>11</td>
<td>12</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Creek</td>
<td>*</td>
<td>15</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brook</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>*</td>
<td>2</td>
</tr>
<tr>
<td>Branch</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>*</td>
</tr>
</tbody>
</table>

Form 2b

Class in Riverside. n = 5. Multiple answers by a single respondent are each counted separately, so number may exceed n.
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Sub-Comparisons</th>
<th>Total #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>River - stream</strong></td>
<td>River larger than stream</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>A stream is small [er than river]</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total river larger/stream smaller</strong></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>River deeper than stream</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>River is wider than stream</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>River runs faster than stream</td>
<td>1</td>
</tr>
<tr>
<td><strong>River - creek</strong></td>
<td>River larger than creek</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Creek much smaller than river</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total river larger/creek smaller</strong></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>River runs faster than creek</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Creek small body of moving water</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>A creek has banks</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>A creek is more harky (?) [than river]</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>A creek can be residential</td>
<td>1</td>
</tr>
<tr>
<td><strong>River - brook</strong></td>
<td>River larger than brook</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Brook is much smaller than river</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total river larger/brook smaller</strong></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Brook runs slowly [more than river]</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Brook similar to creek</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Brook could be branch of a river</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>A clear standing water (??)</td>
<td>1</td>
</tr>
<tr>
<td><strong>Stream - creek</strong></td>
<td>Pretty much the same</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Similar terms</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total quasi identity</strong></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Stream larger than creek</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Stream may be larger than creek</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total stream larger/creek smaller</strong></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Creek moves slower than stream</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Stream is moving fairly fast, creek moving slow</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total stream fast/creek slow</strong></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Stream has fish</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Stream is small [not vs. creek ?]</td>
<td>1</td>
</tr>
</tbody>
</table>
Stream - brook
- Brook is much smaller than stream 1
- Stream is moving faster than brook 1
- Stream is dirtier 1
- Brook is isolated area of stream 1

Creek - brook
- Brook is smaller than creek 1
- A brook is in the country 1
- No difference 1

River - branch
- A branch is a part of a river when it splits 1
- Branch is where river splits 1
  - Total branch is ‘river split’ 2
- A branch is a separate part of a river 1
- Branch = channel of river 1
  - Overall total branch is ‘part of river’ 4

Stream - branch
- Branch is part of a river 1
- Stream splits 1
- Branch of stream 1
  - Total branch is ‘part of stream’ 3
- Stream is not part of a river 1

Creek - branch
- Branch is part of a river 1
- Channel of creek = branch 1
  - Total branch is ‘part of creek’ 2
- Creek possibly man-made 1

Brook - branch
- Branch is part of a river 1
- Brook is smaller than branch 1
- No difference 1
Appendix C

Tables

Table 1. The occurrence of the nouns in the corpora given in the form of number of instances as well as in percentages. American English = italics; English English = underlined; and common English = bold. Only nouns that were found in any of the corpora were included in the table.

<table>
<thead>
<tr>
<th>word/corpus</th>
<th>Brown inst. %</th>
<th>Frown inst. %</th>
<th>SED inst. %</th>
<th>BNC inst. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>beck</td>
<td>0 0</td>
<td>0 0</td>
<td>10 8.1</td>
<td>97 1.2</td>
</tr>
<tr>
<td>bourne</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0</td>
</tr>
<tr>
<td>branch</td>
<td>0 0</td>
<td>1 1.4</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>brook</td>
<td>1 0.9</td>
<td>1 1.4</td>
<td>34 27.4</td>
<td>140 1.8</td>
</tr>
<tr>
<td>brooklet</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0</td>
</tr>
<tr>
<td>burn</td>
<td>1 0</td>
<td>0 0</td>
<td>3 2.4</td>
<td>72 0.9</td>
</tr>
<tr>
<td>creek</td>
<td>8 7.0</td>
<td>10 13.9</td>
<td>2 1.6</td>
<td>33 0.4</td>
</tr>
<tr>
<td>dike</td>
<td>0 0</td>
<td>0 0</td>
<td>9 7.3</td>
<td>0 0</td>
</tr>
<tr>
<td>ditch</td>
<td>0 0</td>
<td>0 0</td>
<td>4 3.2</td>
<td>2 0</td>
</tr>
<tr>
<td>drain</td>
<td>0 0</td>
<td>0 0</td>
<td>2 1.6</td>
<td>1 0</td>
</tr>
<tr>
<td>drill</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>dyke</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0</td>
</tr>
<tr>
<td>gill</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>gote</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>holl</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>leat</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>nailbourn</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>pinnock</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>prill</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>rill</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>3 0</td>
</tr>
<tr>
<td>rindle</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>ripplet</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>rivelet</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>rivulet</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>15 0.1</td>
</tr>
<tr>
<td>sike</td>
<td>0 0</td>
<td>0 0</td>
<td>6 4.8</td>
<td>0 0</td>
</tr>
<tr>
<td>slade</td>
<td>0 0</td>
<td>0 0</td>
<td>1 0.8</td>
<td>0 0</td>
</tr>
<tr>
<td>stall</td>
<td>0 0</td>
<td>0 0</td>
<td>5 4.0</td>
<td>0 0</td>
</tr>
</tbody>
</table>
Table 1 - Continuation. The occurrence of the nouns in the corpora given in the form of number of instances as well as in percentages.

<table>
<thead>
<tr>
<th></th>
<th>21</th>
<th>18.4</th>
<th>11</th>
<th>15.3</th>
<th>30</th>
<th>24.2</th>
<th>1260</th>
<th>15.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>stream</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>streamlet</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>trench</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vein</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>water</td>
<td>1</td>
<td>0.9</td>
<td>2</td>
<td>2.8</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>0.3</td>
</tr>
<tr>
<td>watercourse</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>56</td>
<td>0.7</td>
</tr>
<tr>
<td>waterway</td>
<td>1</td>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 2. The frequency rates (in percentages) for the various terms during the Old and Middle English periods (based on Rundblad 1998). The 7th, 8th, 9th, 12th, and 19th centuries are omitted due to lack of instances.

<table>
<thead>
<tr>
<th>century/ (no of instances)</th>
<th>10th/ (26)</th>
<th>11th/ (15)</th>
<th>13th/ (86)</th>
<th>14th/ (43)</th>
<th>15th/ (60)</th>
<th>16th/ (79)</th>
<th>17th/ (51)</th>
<th>18th/ (18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>become</td>
<td>0.0</td>
<td>0.0</td>
<td>2.3</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>bliDe</td>
<td>3.8</td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>bróc</td>
<td>23.1</td>
<td>13.3</td>
<td>31.4</td>
<td>16.3</td>
<td>28.3</td>
<td>32.9</td>
<td>31.4</td>
<td>44.4</td>
</tr>
<tr>
<td>burna</td>
<td>34.6</td>
<td>20.0</td>
<td>18.6</td>
<td>9.3</td>
<td>13.3</td>
<td>10.1</td>
<td>5.9</td>
<td>0.0</td>
</tr>
<tr>
<td>dic</td>
<td>0.0</td>
<td>6.7</td>
<td>4.7</td>
<td>0.0</td>
<td>5.0</td>
<td>3.8</td>
<td>9.8</td>
<td>0.0</td>
</tr>
<tr>
<td>drain</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>7.8</td>
<td>0.0</td>
</tr>
<tr>
<td>éa</td>
<td>15.4</td>
<td>13.3</td>
<td>10.5</td>
<td>18.5</td>
<td>10.0</td>
<td>6.3</td>
<td>9.8</td>
<td>16.7</td>
</tr>
<tr>
<td>*fléama</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>fléd, flód, flóde, *flódr</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>flést, féote, *fltete</td>
<td>0.0</td>
<td>6.7</td>
<td>1.2</td>
<td>4.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*gotu, gyte</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*grendel, *grinde</td>
<td>3.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>hlúð(e), * hlúding, hlýde</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*læd, (ge-)låd</td>
<td>0.0</td>
<td>0.0</td>
<td>4.7</td>
<td>0.0</td>
<td>1.7</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 2 - Continuation. The frequency rates (in percentages) for the various terms during the Old and Middle English periods (based on Rundblad 1998).

<table>
<thead>
<tr>
<th>term</th>
<th>0.0</th>
<th>0.0</th>
<th>1.2</th>
<th>0.0</th>
<th>0.0</th>
<th>0.0</th>
<th>0.0</th>
<th>0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>must</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>pól</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*ræc, racu, *ric</td>
<td>0.0</td>
<td>6.7</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>riD,(ge-)riDe, riDig</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.3</td>
<td>1.7</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*sæging</td>
<td>0.0</td>
<td>0.0</td>
<td>2.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>seolfor</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.6</td>
</tr>
<tr>
<td>sic, *sicel</td>
<td>0.0</td>
<td>0.0</td>
<td>2.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>sleæd</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*smitre</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>*swifte</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>stræam</td>
<td>15.4</td>
<td>20.0</td>
<td>1.2</td>
<td>7.0</td>
<td>6.7</td>
<td>7.6</td>
<td>3.9</td>
<td>5.6</td>
</tr>
<tr>
<td>wæter</td>
<td>0.0</td>
<td>0.0</td>
<td>11.6</td>
<td>25.6</td>
<td>21.7</td>
<td>27.8</td>
<td>17.6</td>
<td>17.0</td>
</tr>
<tr>
<td>*wende, *wendel, (ge-)wind</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.3</td>
<td>0.0</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>wielle</td>
<td>0.0</td>
<td>6.7</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.6</td>
</tr>
<tr>
<td>*wise, *wisoc</td>
<td>0.0</td>
<td>6.7</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 3. Comments taken from the SED; those who are unfamiliar with the SED should note that the comments are spontaneous responses from informants that should not be treated or seen as semantic definitions. It should be stressed that not all informants commented upon their use; hence the “lack” of comments for some terms and/or shires. The table shows how word 1 is compared to word 2 (or not compared at all, just commented upon); hence in Cumberland a beck is wider than a sike according to the informant.

<table>
<thead>
<tr>
<th>word 1</th>
<th>feature/use</th>
<th>word 2</th>
<th>shire</th>
</tr>
</thead>
<tbody>
<tr>
<td>beck</td>
<td>wider</td>
<td>sike</td>
<td>Cumberland</td>
</tr>
<tr>
<td>beck</td>
<td>unusual</td>
<td></td>
<td>Lancashire</td>
</tr>
<tr>
<td>beck</td>
<td>bigger</td>
<td>dike</td>
<td>Lincolnshire</td>
</tr>
<tr>
<td>beck</td>
<td>bigger</td>
<td>drain</td>
<td>Lincolnshire</td>
</tr>
<tr>
<td>beck</td>
<td>bigger</td>
<td>ea</td>
<td>Lincolnshire</td>
</tr>
<tr>
<td>beck</td>
<td>bigger</td>
<td>stream</td>
<td>Lincolnshire</td>
</tr>
</tbody>
</table>
Table 3 - Continuation. Comments taken from the SED; those who are unfamiliar with the SED should note that the comments are spontaneous responses from informants that should not be treated or seen as semantic definitions.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$beck$</td>
<td>$larger$</td>
<td>$stream$</td>
<td>$Norfolk$</td>
</tr>
<tr>
<td>$brook$</td>
<td>$wider$</td>
<td>$ditch$</td>
<td>$Berkshire$</td>
</tr>
<tr>
<td>$brook$</td>
<td>$bigger$</td>
<td>$stream$</td>
<td>$Essex$</td>
</tr>
<tr>
<td>$brook$</td>
<td>$wider$</td>
<td>$stream$</td>
<td>$Gloucestershire$</td>
</tr>
<tr>
<td>$brook$</td>
<td>$more$</td>
<td>$usual$</td>
<td>$stream$</td>
</tr>
<tr>
<td>$brook$</td>
<td>$wider$</td>
<td>$stream$</td>
<td>$Lancashire$</td>
</tr>
<tr>
<td>$brook$</td>
<td>(tributary)</td>
<td>$stream$</td>
<td>Wiltshire</td>
</tr>
<tr>
<td>$brook$</td>
<td>$larger$</td>
<td>$stream$</td>
<td>Warwickshire</td>
</tr>
<tr>
<td>$creek$</td>
<td>$tributary$</td>
<td>$river$</td>
<td>Essex</td>
</tr>
<tr>
<td>$dike$</td>
<td>$bigger$</td>
<td>$drill$</td>
<td>Kent</td>
</tr>
<tr>
<td>$dike$</td>
<td>$smaller$</td>
<td>$brook$</td>
<td>Oxfordshire</td>
</tr>
<tr>
<td>$dike$</td>
<td>$smaller$</td>
<td>$stream$</td>
<td>Oxfordshire</td>
</tr>
<tr>
<td>$dingle$</td>
<td>$smaller$</td>
<td>$brook$</td>
<td>Herefordshire</td>
</tr>
<tr>
<td>$dingle$</td>
<td>$smaller$</td>
<td>$stream$</td>
<td>Herefordshire</td>
</tr>
<tr>
<td>$ditch$</td>
<td>artificial/natural</td>
<td></td>
<td>Essex</td>
</tr>
<tr>
<td>$gill$</td>
<td>$smaller$</td>
<td>$beck$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$gill$</td>
<td>$smaller$</td>
<td>$stream$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$gill$</td>
<td>$smaller$</td>
<td>$brook$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$gote$</td>
<td>$tributary$</td>
<td>$beck$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$gote$</td>
<td>$tributary$</td>
<td>$stream$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$gote$</td>
<td>$tributary$</td>
<td>$brook$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$sike$</td>
<td>$narrow$</td>
<td></td>
<td>Cumberland</td>
</tr>
<tr>
<td>$sike$</td>
<td>$tributary$</td>
<td>$beck$</td>
<td>Durham</td>
</tr>
<tr>
<td>$sike$</td>
<td>narrower</td>
<td>$beck$</td>
<td>Westmoreland</td>
</tr>
<tr>
<td>$sike$</td>
<td>$smaller$</td>
<td>$gill$</td>
<td>Yorkshire</td>
</tr>
<tr>
<td>$stream$</td>
<td>$smaller$</td>
<td>$brook$</td>
<td>Essex</td>
</tr>
<tr>
<td>$stream$</td>
<td>$smaller$</td>
<td>$brook$</td>
<td>Lancashire</td>
</tr>
<tr>
<td>$stream$</td>
<td>$smaller$</td>
<td>$brook$</td>
<td>Warwickshire</td>
</tr>
</tbody>
</table>
Table 4. The hierarchical structure of the lexical field ‘a natural stream’ in the various periods of English in England. The hierarchy is focused on middling to small size and hence river is not included. It should be noted that Tables 4 and 5 are meant to give an overview of the changes; to get an accurate and detailed picture of the changes Tables 1–3 as well as the etymological discussion in Rundblad 1998 need to be consulted and incorporated. Terms labeled 1 were the default of the period, the other terms being dependent to a varying degree on the default. Terms at level 2 were/are less frequently used than those at level 1, and terms at level 3 were/are less used than those at level 2, etc. In general, the terms on levels 2–4 are smaller in size than the default ones. The information within brackets refers to the most conspicuous and significant part of the semantic definition of the term, especially when compared to the default term; it is not an exhaustive definition (cf. Rundblad 1998). That bekkr/bek/beck is within brackets in the Standard English English row is meant to stress that the term is highly dialectal. The table shows the replacement of burna by bróc in the Midlands, South England, and in Standard English, and the increasing reliance on stream. The clear preference for beck in the North is also evident from an early stage. The table is based on data from Rundblad (1998).

<table>
<thead>
<tr>
<th>region</th>
<th>Old English</th>
<th>Middle English</th>
<th>Modern English English</th>
</tr>
</thead>
<tbody>
<tr>
<td>North England</td>
<td>1. bekkar (Old Norse) 1. a/ea (general) 2. dik/dic (poss artificial) 2. gote (poss artificial) 2. sik/sic (smaller)</td>
<td>1. bek (esp mountain stream) 2. broke (general) 2. sike/sitch 2. water (general) 3. dike/ditch (poss artificial) 3. ea (general) 3. gote (poss artificial)</td>
<td>1. beck (esp mountain stream) 2. brook (general) 2. burn (esp flatland stream) 2. stream (general) 3. sike (smaller (than gill)) 4. dike (poss artificial) 4. drain (poss artificial) 4. e (smaller) 4. gill (small) 4. gote (side-stream, poss artificial) 4. stell (open drain)</td>
</tr>
</tbody>
</table>
Table 4 - Continuation. The hierarchical structure of the lexical field ‘a natural stream’ in the various periods of English in England.

<table>
<thead>
<tr>
<th>The Midlands</th>
<th>South England</th>
<th>Standard English English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. bróc (marshy)</td>
<td>1. burna (clear, well)</td>
<td>1. burna (clear, well)</td>
</tr>
<tr>
<td>2. burna (clear, well)</td>
<td>2. bróca (marshy)</td>
<td>1. bró (marshy)</td>
</tr>
<tr>
<td>3. dik/dic (pos artificial)</td>
<td>2. stréam (general)</td>
<td>(2. bekkr)</td>
</tr>
<tr>
<td>3. stréam (general)</td>
<td>2. stréam (general)</td>
<td>2. stréam (general)</td>
</tr>
<tr>
<td>3. wæter (general)</td>
<td>3. wæter (general)</td>
<td>2. wæter (general)</td>
</tr>
<tr>
<td>3. éa (general)</td>
<td>3. éa (general)</td>
<td>3. éa (general)</td>
</tr>
<tr>
<td>3. dícc (pos artificial)</td>
<td>3. dícc (pos artificial)</td>
<td>3. dícc (pos artificial)</td>
</tr>
<tr>
<td>4. dícc (pos artificial)</td>
<td>4. dícc (pos artificial)</td>
<td>4. dícc (pos artificial)</td>
</tr>
<tr>
<td>4. gote (pos artificial)</td>
<td>4. gote (pos artificial)</td>
<td>4. gote (pos artificial)</td>
</tr>
<tr>
<td>4. wæter (general)</td>
<td>4. wæter (general)</td>
<td>5. gote (pos artificial)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
<tr>
<td>4. break (general)</td>
<td>4. break (general)</td>
<td>5. break (general)</td>
</tr>
</tbody>
</table>
Table 5. The lexical field ‘a natural stream’ during the various periods of English in England showing the structure of the specifications/features of the terms included in Table 4. The specifications include: QL (quality), QT (quantity), SP (speed), and T (type). Specifications within brackets indicate that these specifications are plausible but not certain. The table shows clearly the move towards a quantity (size) specific field, where the number of specifications and the use of other specifications apart from QT increases, the further down the term stands on the scale, which generally are terms used for small(er) watercourses as opposed to middling watercourses. The table is based on data from Rundblad 1998.

<table>
<thead>
<tr>
<th>region</th>
<th>Old English</th>
<th>Middle English</th>
<th>Modern English England</th>
</tr>
</thead>
<tbody>
<tr>
<td>North England</td>
<td>1. QT T (SP)</td>
<td>1. QT T (SP)</td>
<td>1. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>1. (QT)</td>
<td>2. QT</td>
<td>2. QT</td>
</tr>
<tr>
<td></td>
<td>2. (QL) (QT) T</td>
<td>2. QT T (T)</td>
<td>2. QT</td>
</tr>
<tr>
<td></td>
<td>2. QT</td>
<td>3. (QL) (QT) T</td>
<td>3. QT</td>
</tr>
<tr>
<td></td>
<td>2. QL QT</td>
<td>3. (QT)</td>
<td>4. (QL) (QT) T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. QT T (T)</td>
<td>4. QT T SP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. QT T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. (QT) T</td>
</tr>
<tr>
<td>The Midlands</td>
<td>1. QL QT T (SP)</td>
<td>1. QT</td>
<td>1. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>2. QT</td>
<td>1. QT</td>
<td>1. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>2. QT T (SP)</td>
<td>2. QT</td>
<td>2. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>2. (QT)</td>
<td>3. QT T (T)</td>
<td>3. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>2. (QL) (QT) T</td>
<td>3. QT</td>
<td>3. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>2. QT</td>
<td>4. (Q) (QT) T</td>
<td>4. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>2. QT T (T)</td>
<td>4. QT T (T)</td>
<td>4. QT T (SP)</td>
</tr>
<tr>
<td></td>
<td>4. QT</td>
<td>4. QT T (T)</td>
<td>4. QT T (SP)</td>
</tr>
<tr>
<td>South England</td>
<td>1. QL QT T (SP)</td>
<td>1. QT</td>
<td>1. QT</td>
</tr>
<tr>
<td></td>
<td>2. QT T (SP)</td>
<td>1. QT</td>
<td>1. QT</td>
</tr>
<tr>
<td></td>
<td>2. QT</td>
<td>2. QT</td>
<td>2. QT</td>
</tr>
<tr>
<td></td>
<td>3. (QT)</td>
<td>3. (QT)</td>
<td>3. (QT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT T (T)</td>
<td>4. QT T (T)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT</td>
<td>4. QT T (T)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT T (T)</td>
<td>4. QT T (T)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT SP</td>
<td>4. QT T (T)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT SP</td>
<td>4. QT T (T)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT T</td>
<td>4. QT T (T)</td>
</tr>
</tbody>
</table>
Table 5 - Continuation. The lexical field ‘a natural stream’ during the various periods of English in England showing the structure of the specifications/features of the terms included in Table 4.

<table>
<thead>
<tr>
<th>Standard English</th>
<th>1. QL QT T (SP)</th>
<th>1. QT</th>
<th>1. QT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. QL QT T (SP)</td>
<td>2. QT</td>
<td>2. QT</td>
</tr>
<tr>
<td></td>
<td>2. QT (2. QT T (SP))</td>
<td>(3. QT T (SP))</td>
<td>(3. QT T (SP))</td>
</tr>
<tr>
<td></td>
<td>2. QT</td>
<td>3. QT</td>
<td>3. QT</td>
</tr>
<tr>
<td></td>
<td>3. (QL) (QT) T</td>
<td>4. (QT)</td>
<td>4. (QT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. QT</td>
<td>4. QT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. (QL) (QT) T</td>
<td>4. (QL) (QT) T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. QT</td>
<td>1. QT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. QT</td>
<td>2. QT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2. QT T (SP))</td>
<td>(2. QT T (SP))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. (QL) (QT) T</td>
<td>3. (QL) (QT) T</td>
</tr>
</tbody>
</table>
References

Andersen, Henning

Berlin, O. Brent

Berlin, O. Brent and Paul Kay

BNC

Bosworth, Joseph and T. Northcote Toller (eds.)

Boyes-Braem, Penny (ed.)

Clarke, David D. and Brigitte Nerlich

Colby, Benjamin N.

Colby, Benjamin N., James W. Fernandez, and David B. Kronenfeld

Cole, Ann

D’Andrade, Roy G.

Edlund, Lars-Erik
1987 Änman och binann – en idéskiss kring “friare” ord- och namnbildning. In Nionde Nordiska Namnforskarkongressen, Göran Hall-

Ekwall, Eilert

Frake, Charles O.

Friedrich, Paul

Gelling, Margaret

Greenberg, Joseph H.

Holthausen, Ferdinand

Jonsson, Hans

Kempton, Willett

Kristensson, Gillis

Kronenfeld, David B.

Kronenfeld, David B. and Henry Decker
Lakoff, George

Langacker, Ronald W.

Lehrer, Adrienne

Löfvenberg, Mattia T.
1942 Studies on Middle English Local Surnames. Lund: Gleerup.

Lounsbury, Floyd G.

Lyons, John

Nerlich, Brigitte and David D. Clarke

Orton, Harold (ed.)
Romney, A. Kimball and Roy G. D’Andrade

Rosch, Eleanor H., Carolyn B. Mervis, Wayne D. Gran, David M. Johnson, and Penny Boyes-Braem (eds.)

Rundblad, Gabriella

de Saussure, Ferdinand

Simpson, John A. and Edmund S.C. Weiner (eds.)

Smith, Albert H. (ed.)

Soukhanov, Anne H. et al. (eds.)

Tomić, Olga M. (ed.)

Trier, Jost
Tyler, Stephen A. (ed.)

Ullmann, Stephen

Wallace, Anthony F. C. and John Atkins

Wright, Joseph (ed.)
khalîfa – A word study*

David J. Wasserstein

1. Introduction

In this study I propose to consider the history of the Arabic word khalîfa ‘caliph’. Despite what might be considered some limitations on its usefulness, the term offers a good example, or case study, of word history. It shares a common type of background with countless other words in Arabic. It emerges at a definable point in history, out of historical events which, even if they are controversial, can nonetheless be followed in some detail. Over the last fourteen centuries it has acquired a range of meanings, some of which have remained while others have disappeared, and others again have undergone internal development.

The limitations derive from the special character of the word itself. First, it refers to the holder of an office, rather than, say, to a material object, or something else of more tangible, or at least easily grasped, general nature. Secondly, that office is unique, belonging to the cultural world of Islam, and has few if any real parallels outside that world (though the term itself exists, as a loan-word, also in other languages, in all of them however as a name for the same holder of the same institution). Thirdly, there is a sense in which the history of the word is actually something else, the history of the office which it labels and identifies; while this is true in many other cases too, here there is the additional difficulty that the entire previous history of the word’s meanings seems to affect the meaning of the word at any particular moment in its history. Fourthly, word histories are generally far rarer in Arabic than in European languages, old and new. While we do come across remarks, on occasion even short disquisitions, on word histories in classical Arabic authors, these are uncommon and largely unsystematic, with a tendency to anecdote and over-credulity concerning folk

* We would like to thank Eugene Anderson and Robert Moore for their help in distributing and collecting various research forms, and all the students for participating in our research (see Appendices A and B).
etymologies and the like. Fifthly, in the Arabic linguistic tradition, meaning, however that notion may be defined, can easily be imagined to inhere in a word, fully, both diachronically and synchronically, from the time of Adam or, at the least, from the time of the first speaker of the Arabic language, often asserted to have been Ismā‘īl (the biblical Ishmael) (Nadim 1970: 6–9).

In this specific case, moreover, the birth in time of the meaning, or the use, of this word in the sense of ‘caliph’ has a very easily defined cause or origin: as will be seen, that origin post-dates the rise of Islam, even the death of the Prophet Muhammad, in 632 (CE – all dates here will be CE). Despite this, there is a sense, religious as well as lexicographical, in which, from the point of view of Muslims living in the cultural world of classical Islam, and even perhaps in the modern period, the meaning ‘caliph’ may be reasonably assumed to have inhere in the word khalīfa from the beginning of the history of the Arabic language, long centuries before the time of the Prophet.

For all that these are serious limitations, it seems worth trying to draw the main lines of the history of the word, both for its own sake and as a case-study of a sort that has not so far been undertaken very often for Arabic. In its character as a term drawn from, but not quite confined to, institutional history, rather than from certain other areas, moreover, we do seem to have some considerable materials available for the study of its history, in our documents and other sources for the history of the institution of the caliphate.

2. Etymology and meanings

The term khalīfa, like most words in classical Arabic, is built of two principal elements. First, there is a consonantal root, in this case as very often triliteral. Combined with this we have formative elements which give the word shape and, to some degree, often by analogy with other words from other roots which have similar shape, help to determine the direction of its meaning. A base meaning is often understood to inhere in the root itself, though that root does not have any formal existence as a word, and the concept of a base meaning is itself also not without difficulties. Thus the root k.t.b. has a base meaning of ‘write’: it gives us a verb ‘to write’, together with all its associated verbal forms; a noun kitāb, meaning ‘something written’, a ‘book’, a ‘text’, a ‘document’, even ‘agreement’, ‘treaty’, etc.; by
extension, \textit{al-kitāb} (i.e., with the definite article) also means ‘The Book’, \textit{viz.}, the \textit{Qur'ān} (or sometimes the Bible), much in the manner of ‘Holy Writ’ or ‘the Good Book’ in English; a participial noun, \textit{kātib}, meaning ‘scribe’; a verbal noun \textit{kitāba} ‘writing’, ‘the art of writing’, etc.; another verbal noun \textit{mukātaba} meaning ‘exchange of letters’, ‘correspondence’; a noun of place, \textit{maktab}, meaning ‘office’ (i.e., a place where writing is done); and much more besides (Lane 1863–93, sub radice \textit{k.t.b.}; Wehr 1971, sub radice \textit{k.t.b.}). Similarly the root \textit{q.t.l.}, with a base meaning of ‘kill’, ‘slay’, ‘assassinate’, ‘murder’, etc., generates such words as \textit{qatl} ‘murder’, ‘assassination’, ‘killing’; \textit{qitl} ‘enemy’, ‘foe’; \textit{qatala} ‘to kill’, \textit{qāṭala} ‘to combat’, ‘fight against someone’; \textit{qāṭil} ‘casualty’, ‘one killed in battle’; \textit{qāṭal} ‘murderous’; \textit{maqṭal} ‘death’, ‘murder’; \textit{maqtala} ‘carnage’, ‘slaughter’; \textit{muqāṭala} ‘fight’, ‘struggle’; \textit{muqāṭil} ‘fighter’; and so on (Lane 1863–93, sub radice \textit{q.t.l.}; Wehr 1971, sub radice \textit{q.t.l.}). Neither the forms nor their meanings, it should be noted, are necessarily determined in any specific case automatically or obviously.

In the case of \textit{khalīfa} the root letters are \textit{kh.l.f.} (\textit{kh} pronounced as \textit{ch} in German \textit{doch}, or Scottish \textit{loch}). The base meaning of this root appears, on the basis of the occurrences of the root in such languages as Aramaic, Ethiopic, and Hebrew, as well as Arabic, to be something like ‘to pass by’, ‘to go past’, ‘to change’. This last meaning leads us, without any difficulty, in one direction, to such meanings as ‘to follow’, ‘to be behind’, and, in another direction, to meanings like ‘to be different’, ‘to be contradictory’, etc. From this root we have in Arabic a verb \textit{khalafa}, meaning ‘to follow’, ‘to succeed’, ‘to take the place of’, ‘to lag behind’, ‘to stay away’; another verb, or verb-form, \textit{khālafa} ‘to be contradictory’, ‘to oppose’; another, \textit{akhlafa} ‘to leave’, ‘to leave behind’ (offspring, etc.); another, \textit{takhallafa} ‘to stay behind’; yet another, \textit{ikhtalafa} ‘to differ’, ‘to vary’, ‘to come or go frequently’, ‘to visit frequently’; and even \textit{istikhlafa} ‘to appoint as successor or replacement’; a noun \textit{khalīf} ‘rear’; \textit{khulīf} ‘dissimilarity’, ‘contrast’; \textit{khalaf} ‘substitute’, ‘descendant’; \textit{kilīfa} ‘disparity’; \textit{khlīf} ‘difference’, ‘disparity’, ‘divergence’; \textit{kīlīfī} ‘controversial’; \textit{mukhālafa} ‘contrast’, ‘contradiction’, ‘contravention’, ‘violation’; and so on, including also the words \textit{khalīfa} ‘caliph’, and \textit{khlīfa} ‘caliphate’ (both as the name of the office and as the term for a period of office in that institution) (Lane 1863–93, sub radice \textit{kh.l.f.}; Wehr 1971, sub radice \textit{kh.l.f.}).

The particular word from the root \textit{kh.l.f.} whose meaning is of interest here is \textit{khalīfa}, which can be translated as ‘successor’, ‘replacement’, ‘substitute’, ‘deputy’. The simple translation, or explanation, of the meaning of
this word in English, as in other languages, is ‘caliph’, but of course ‘caliph’ itself is nothing more than a loan-transcription of the Arabic, reflecting thus the uniqueness to the Arabic-Islamic context of the notion which the word expresses. We find the same word, in various forms, in most non-Arabic languages that have a name for caliph: thus French has *caliphe*, Spanish *califa*, Italian also *califa*, German *Chalif* or *Kalif*, Dutch *kalief*, Hebrew *khalifa*). And so on. We do find other translations from time to time, preeminently ‘king’, reflecting one of the more obvious of the characteristics of the caliph, his quality as a ruler, or sovereign. But such versions ignore the most significant of the characteristics of the caliph.

Arabic has a number of words for ‘ruler’: *malik* ‘king’, *sultân* ‘sultan’; *sâhîb* ‘lord’; *amîr* ‘commander,’ ‘prince’; *sayyid* ‘lord’; and others. All of these words may be designations, formal or informal, of a ruler; they may be (part of) titles, or merely descriptions. I shall return below to the fact that there is an overlap between description (e.g., ‘ruler’, ‘sovereign’) and title (e.g., ‘king’) in Arabic. That overlap is perhaps greater in Arabic than in certain other languages; the greater overlap seems to me to derive in part from the morphology of the language and the way in which it creates or generates such terms as ‘king’, etc.

The term ‘caliph’ has a meaning different from ‘king’, or ‘lord’; it is a title reserved, in some sense, for a particular sort of sovereign. It is also only one among a number of titles which that sort of sovereign holds. Although it is the one which we in non-Islamic societies use most commonly to designate these particular sovereigns, *khalîfa* is not in fact the term used most commonly for that purpose in Arabic. The principal formal titles of this ruler are three: *khalîfa*, *imâm*, and *amîr al-mu’minîn*. (I disregard here both titles like *malik*, etc. and other formal titles which occur from time to time, as these are generally not titles of the post so much as decorations additional to the title of the office. For example in English the British sovereign is termed ‘Defender of the Faith’: that title does not designate the king of England; it is descriptive, rather than designative. So is it also with titles like ‘Shadow of God on Earth’, and ‘Guardian of the Two Holy Cities’ of Mecca and Madina, titles used by the Ottoman Sultans.)

A remark on the form of the word *khalîfa* is in place here too. It has the ending -*a* which is normally characteristic of feminine nouns in Arabic. Grammatically, however, essentially because of its meaning, the word is masculine. Such apparent crossovers of gender are not unknown in Arabic, as in other languages (those who have learned German as a foreign language will think inevitably of the neutral *Mädchen*). ‘ajûz is masculine in
form, but grammatically feminine, with the meaning ‘old woman’; similarly, *hâmil*, literally ‘carrying’, which is masculine in form, is grammatically feminine, when it has the meaning ‘pregnant’. It does not take a feminine ending with this meaning. *khalîfa* is similar in shape to certain other Arabic nouns which acquire by the addition of the same ending *-a* an intensified meaning; but it differs from them in that it seems not to have a form with similar meaning without this ending, as well as having nothing about it that suggests intensification (Wright 1896–98: I, 139, §304 XX; see also Barth 1894: §85 a, with §54 b). Although there are a couple of exceptions, these are very old and rare (Lane 1863–93: 798, s.v.).

The word *khalîfa* has a variety of meanings in Arabic. The dictionaries give such standard meanings as ‘successor’, ‘deputy’, ‘representative’, ‘substitute’, ‘vicar’. But there are others too: we hear of it meaning ‘commander of fifty men’ (Fagnan 1923: 49); and we also find it with the more specialized meaning of ‘slave in the service of the palace’, in al-Andalus, Islamic Spain (Dozy 1881: I, 397). From this the term comes to be applied also to such slaves when they were granted freedom, on the death of the caliph. In early eleventh-century Spain, we even find two very minor and short-lived rulers who were called *khalîfa*, possibly as a sort of title deriving from this (Wasserstein 1985: 158 n.8). We also find the expression used as a proper name of individuals. A well-known example of this is the early Islamic historian *Khalîfa ibn Khayyât* (Encyclopaedia of Islam, sub nomine). In other languages, heavily influenced by Arabic and by Islam, we find still more meanings. Thus, to take Turkish, we find there the meanings ‘junior clerk in a public office’, ‘an assistant teacher in a school or college’, ‘a builder, a past apprentice’ and even ‘a superior female servant’ (Redhouse 1890: 863, s.v. ‘khalife’). In the region of Khorazm, certain female Quranic teachers have the name *xalpa*, a version of the same word *khalîfa*, and it occurs in Bukhara and Samarkand in the forms *bibi xalpa* and *bi-xalfa*; in Afghan Turkestan, *xalipa* or *xalifa*, with the stress on the first syllable, denotes, among other things, a low spiritual rank, something like the devotee of a holy man. And in Afghanistan, lastly, *xalifa* means ‘truckdriver’ (see Fathi 1997, 1998). But the most important meaning, and the one that is of concern here, is that rendered in other languages by the loan-word ‘caliph’.
3. The origin of the use of the title *khalîfa*

The term *khalîfa* came into use very soon after the death of the Prophet Muhammad in 632, as one of the principal titles of the rulers who succeeded him at the head of the new polity of Islam. It was for a long time understood and explained as meaning here ‘successor’, and as being a shortened form of the expression ‘*khalîfa* *rasūl Allâh*, ‘successor of the Messenger of God’. However, as has been forcefully argued recently by Crone and Hinds (1986), this explanation seems to be unacceptable. What the sources attest from as little as just over a decade after the death of the Prophet is the title ‘*khalîfat Allâh*, ‘the *khalîfa* of God’. As God stands in no need of a successor, the evident meaning of this is ‘deputy of God’ (scil. on earth). We find this title already in the time of the caliph ‘Uthmân (reigned 644–56), and it continues in this form for a very long time thereafter. We also, it is true, find apparent attestations of the title *khalîfat rasūl Allâh* ‘successor of the Messenger of God’, for an earlier period, for the reign of Abu Bakr (632–34), but these are rejected by Crone and Cook (1986: 21) on arguments which are highly persuasive. These apparent attestations look very much like retrojections from a later period, when the implications of the title *khalîfat Allâh* may have begun to cause worry in certain circles. It looks very much, from the evidence which Crone and Hinds have brought together, as though the title came into use at some stage not very long after the death of the Prophet, and indicated ‘deputy’, viz., of God. We find it in the form ‘*khalîfat Allâh*’ on coins, in documents and in literary texts from an early period, with the coins confirming that occurrence in texts which have been copied many times over the succeeding centuries is not a result of later alteration.

The use of this title continues for centuries, through the period of the early caliphs, the entirety of the Umayyad period (661–750), and it is found as late as the penultimate caliph of the Abbasid line in Baghdad, in the middle of the thirteenth century. In all these cases it is one of the formal titles of the leader of the Islamic polity, along with the titles *imâm* and *amîr al-mu'minîn*. All three of these titles are exclusive to caliphs. The first indicates a form of deputy, the second, from a root indicating ‘in front of’, refers to a religious role of the caliph as prayer-leader of the Islamic community, and the third, literally ‘commander of the believers’, ‘commander of the faithful’, has perhaps a more military or political connotation. But these are not alternative roles. All are subsumed in the same person.
(It is relevant here, in terms of the status and meaning at different times of the titles employed by caliphs, to note the title amīr al-muslimīn. This title, with its deliberate similarity to, and echo of, the caliphal title amīr al-mu’minīn, was used by Almoravid rulers in the extreme west of north Africa and in Islamic Spain, in the twelfth century, formally loyal to the Abbasid caliphs in Baghdad.)

The early Islamic period is difficult to study. Much of our material is opaque and recent research has tended to reduce further what little appeared securely known. Thus there is much debate at present over the claims and actual role of the caliphs in the early period, particularly over how far their office possessed merely political or also religious characteristics. However, it is clear that the holder of these titles laid claim to leadership of the Islamic community as a state. Supreme leadership of the community in its turn implied universality and exclusivity. Recognition of that status in the various provinces of the growing world-empire of Islam was given by a variety of means. Governors, prayer-leaders, and others were involved in this: the caliph was mentioned formally in the public prayers in the mosques on Friday, his name and titles were, though not at the start, engraved on coins, and they appeared in other formal acts such as documents issued from or received by state offices.

The exclusivity of the power of the caliph led, less directly, to the desire for exclusivity in the holding of the office. Gradually, especially under the Abbasids, the dynasty that replaced the Umayyads in the middle of the eighth century, there developed ideological support for the idea that the caliphate belonged to a single family, the family of the Prophet Muhammad. The idea of male primogeniture never acquired much strength in the Sunni Islamic theory of caliphal, or for that matter royal, succession (as the examples of Jordan and Saudi Arabia demonstrate to this day), and in some sense the office was always held to be elective, but the notion of blood-relationship with the family of the Prophet became under the Abbasids (who were related to the Prophet) one of the qualifications necessary for succession to the office.

The claim that there can only be a single leader of the Islamic community reflects the idea that there can, or should, exist only a single Islamic state. By the middle of the eighth century, however, the single unitary state of early Islam was already a thing of the past. A new dynasty, the Abbasids, ruling in Baghdad, had supplanted the Umayyads of Damascus, and on the fringes of the Islamic world-empire small states admitting no allegiance to the Abbasids were beginning to emerge. One of these was in
Spain, where a survivor of the Umayyads of Damascus succeeded in creating a state under his own rule. At the very start of his rule there he was faced with a problem: as a survivor of a family supplanted and largely murdered by the Abbasids, was he to signal allegiance, even at a formal level, to these same Abbasids, as leaders of the Islamic world, by mentioning them in the prayers and on his coins? Such a course seemed to be indicated by current practice. Alternatively, as a survivor of the Umayyad dynasty that had ruled the Islamic world as caliphs for nearly a century, should he lay claim to the caliphal title himself? From the isolated backwater that was Islamic Spain in the eighth century, such a course would clearly have been ridiculous.

At the start, current practice dictated the result: the Abbasid caliph, even though he was trying to destroy the renascent Umayyad state, was mentioned in the Friday prayers. But after a short time, another surviving Umayyad prince who made his way to Spain pointed to the absurdity involved in the situation and all mention of the Abbasids was suppressed. For a century and a half thereafter, until early in the tenth century, the Islamic state in al-Andalus was in effect a separate Islamic community, or a part of the universal Islamic community, without a caliph (Wasserstein 1993a: 8–33).

The Abbasids tried, naturally, to rid themselves of these last Umayyads and to restore the overall unity of the Islamic world-empire. However, failure in this and a greater interest in the eastern parts of the Islamic world meant greater freedom for the western areas to develop independently of Baghdad. By the tenth century, political and social changes in the Abbasid empire had made much of it largely autonomous or even wholly independent, with allegiance to Baghdad no more than a formula. On one side this formality of allegiance to Baghdad reflected a desire to stress identity with the greater world of Islam and to derive a form of legitimacy from recognition by the caliph which such formal allegiance to him brought about. On the other it reflected a desire by the caliphs themselves to rescue what could be rescued from the general decline of their strength. From being world rulers, invested by God as His deputies with exclusive rule over the whole of the Islamic community, the caliphs had become no more than puppets of military adventurers in Baghdad, and mere symbols of an illusory unity outside.

At this stage we can discern already the traces of a new meaning for their title too. While ‘caliph’ formerly had indicated a world-ruler, forming part of a collection of titles with real political content, now it began to indi-
cate something rather weaker and looser, the leader of the community of Muslims. The formal claim to political leadership of the entire community remained, and is visible in the formal documents of investiture issued by caliphs, in effect on demand, to local rulers in different areas, often very far removed from Baghdad, and even from areas where no genuine caliphal writ had ever run in the past. But the reality was very different. On one hand, the caliph offered a focus for Muslim identity and unity, a source of legitimation for rulers who lacked, in the Islamic context, any other adequate source of legitimacy. On the other, local rulers who asked for formal investiture from the caliphs and sent them gifts, or what could be regarded as tribute, left the caliphs some elements of their former status and gave their institution new types of meaning. This reality was recognized by both sides. The hollowness of caliphal power was seen for what it was by contemporaries, and it came to be justified in legal and constitutional terms by writers and others who saw in the institution and in its preservation something that it was worth while, or necessary, to preserve (see Mawerdi 1915, composed by a contemporary of such developments, in order to describe, but also to explain and even justify the changes that had occurred up to his time).

Such preservation, at least in the terms of the past, came to be increasingly difficult. Despite the theoretical view that the caliphate is unique and that there can be only one caliph, in a single caliphal institution, the decline of the prestige and the power of the Abbasids meant that alongside challenges to their secular, political power, there came also other challenges, to their caliphal rights as such. These came from different sources. Two are of significance here.

The first is that of the Fatimids. This dynasty, descended from the Prophet Muhammad through his daughter Fátima, ruled parts of North Africa and the Near East for most of the tenth to the twelfth centuries (between 909 and 1171). They were not Sunni Muslims, but Shiʿis, claiming the right to rule in their quality as descendants of ʿAlī, the cousin and son-in-law of the Prophet (he was the husband of Fátima). He had been the third caliph, but his descendants had not been seen by the majority of Muslims as entitled by virtue of their descent to the caliphal rôle. As rulers, the Fatimids used the title caliph too – though among Shiʿi Muslims the title, in its historical context, has a slightly strange ring.

The title ‘caliph’, as used by the Fatimids, offers two principal aspects of interest here: first, it was used as a clear response and counterpoise to the Abbasid use of the title. It was a reflex of that use, and pointed to a set of
essentially political claims, founded in religion, made by the Fatimids over against the Abbasids. Where the Abbasids had originally ruled, and still claimed to rule, over the whole world of Islam, so now the Fatimids also, using the same title, claimed the right to be rulers of the entire Islamic world. The difference was substantial. The second point, though, is that they never managed to expand the boundaries of their empire, or the persuasiveness of their political-religious doctrine, anywhere beyond the frontiers of north Africa, Egypt, and Syria-Palestine for very long. The consequence is that the Fatimid use of the title changes, from our point of view, the meanings which we can attach to the term.

Now the term no longer means simply the exclusive ruler of the entire Islamic world, or the formal claimant to such rule. It has acquired a broader and at the same time a looser set of meanings: now we can see it as a title for a ruler of an Islamic state with claims to rule broader areas than those actually submitted to his dominion. Such a ruler needs in some sense and in some degree to have a claim to relationship to the Prophet, but such claims, in the Islamic world as outside it, are of course easily created – or ignored (Kassis 1988; Wasserstein 1992). The significant point here is that the Fatimids, like the Abbasids in their time, used the title with claims that went far beyond the reality, and far beyond any realistic possibility.

This was a new situation: on one side the realization that the title could serve to demonstrate political-religious claims going beyond the political reality precisely because founded in religious claims; on the other the existence of two men claiming simultaneously to be caliphs, wielding political power, and being recognised as caliphs by their followers. This new situation opened the way to others.

Now also the Umayyads of al-Andalus, Islamic Spain, saw their opportunity. For nearly two centuries they had named no caliph in their formal acts, but they had nonetheless been in most (non-political) ways part of the Islamic world dominated by, and looking towards, the Abbasids of Baghdad. Now the Abbasids were weak; and the Fatimids, politically and militarily strong in the central Mediterranean, had shown a way forward out of the ideological confusion. The Umayyads too could lay claim to a family connection with the Prophet, though it was more diluted than those of the other two dynasties; however, they could bring forward no genuine religious claim to the caliphal title. What they did, in 929, was simply to make use of the idea of political legitimacy and marry it to their own dynastic concept: they had been caliphs for a century, until the Abbasids swept them away in 750, so now they were entitled to regain both the title and the po-
itical power that was thought to inhere in it. Although the Umayyads never seriously entertained the idea of expanding their own state to the east and regaining their old empire, their new claims did coincide with a period of successful political expansion for them in the western Mediterranean, and in this sense served to underpin the claims which they asserted as rulers somehow different from others in their region.

But what did the caliphal titles, as used by these Iberian Umayyads, really mean? These Umayyads never ruled over more than the area of the Iberian peninsula and bits and pieces of north-western Africa. They never attempted to regain their ancestors’ dominions in the east; they never ruled over the holy cities in the Arabian peninsula; they never incarnated or gave birth to any specifically religious notion of Islam as a state. The only thing that tied them, as caliphs, to the notions implicit in the title (and in the other associated titles which they took up along with that of caliph), was their descent from the Umayyad caliphs of Damascus of two centuries earlier. Their use of the title should be seen in this context, for while, on the ground, it served as no more than a title of essentially local rulership, it also served, somewhat less definitely, to refer to an element in the legitimating basis of the Umayyad regime and, rather less definitely still, to broadcast ideas of Islamic universality and unity.

In the event, Umayyad ambitions proved too large for the realities: by the start of the eleventh century, the dynasty was weakened and in the political convulsions that followed they were swept away completely, the last of them being deposed in 1031. Political legitimation, for the rulers who succeeded them in different city-states in the Iberian peninsula, however, continued to matter, and we witness during the eleventh century a succession of different attempts to deal with the problem that this represented. For some, the solution was simple: in the absence of a living Umayyad caliph, to name a dead one (curiously, not the last of the line) on coins and, presumably, in other public acts. This pretence was maintained by two sorts of rulers: on one hand, we find those who did so merely as a vague and essentially meaningless way of stressing the aspiration for unity in the fissiparous political situation of the period; on the other, we find rulers who sought to offer a limited amount of recognition to a pseudo-caliph set up by one of these minor rulers, in Seville, under the pretence that he was actually a survivor of the political troubles of the preceding decades. And a third path that was followed by such rulers was to pay notional obedience to the idea of a caliphate but no more than that. This was done by offering allegiance to a caliph “abd allâh”. “Abd Allâh” is both an Arab name and a
pair of Arabic words meaning ‘servant of God’; in this latter sense it is one of the additional titles used by caliphs, and used in this way by such rulers it served to disguise the absence of a specific, living caliph from the actual seat of the caliphate in al-Andalus (Wasserstein 1993a).

But this situation raises once again the question of the meaning to be attached to the term caliph in this period. The title certainly did not mean, or refer to, the ruler of the Islamic world. Nor can we suppose that “caliphs” whose realms extended sometimes no further than the modern municipal frontiers of a town like Malaga expected anyone to take such ideas seriously. We have to assume that the title by this time, in such contexts, meant no more than ‘ruler’, with perhaps an echo of something of the past. What makes it possible, even necessary, to include this echo in any such definition is the fact that only two sorts of living people in al-Andalus were able to use the title with any success at all: on the one hand, members of the Umayyad house, until their final deposition in 1031; on the other, members of another family (curiously, the family of ʿAlī, the Prophet’s cousin) who made the claim that they had somehow been assigned the caliphal title as an inheritance by one of the last of these Umayyads. In both cases, it is worth noting, the claimants had the force of history and blood, if nothing else, on their side. Others who tried, or who may have tried, to make use of the title, as an extra glorification for their power, without any real claim to a blood-tie to the Prophet or to earlier caliphs, never succeeded in making the claim hold (Wasserstein 1993a: 21–25). This may, in some degree, suggest that even as late as the eleventh century, and even after the title had undergone such extreme degradation, there were still those to whom the origins and the significance of the title and its associated privileges meant something. However, this remains in the realm of speculation.

As has been seen, the Almoravids, a little later, adopted a title (amīr al-muslimīn) which echoed another caliphal title, amīr al-nuʿminīn, but they did not go so far as to claim caliphal titles in the fullest sense. Members of other later dynasties in the extreme west occasionally did, and we can see this phenomenon as being in part a reflex of the physical distance of the far west from Baghdad and of Abbasid impotence. In the far west the Abbasids had never ruled or had ruled only very briefly and weakly. Any reluctance to take on such titles was therefore the weaker, and the Umayyads were always available as an example.

In the east things developed rather differently. By the early tenth century, the Abbasids had lost virtually all of their political significance. But, as has been seen, both for them and for many of the local rulers who came
to replace them in different places it remained convenient to retain the fiction of the caliphate, and its holders, as standing at the head of a united Islamic polity. In different ways and with varying effect, this remained the situation until the Mongol capture and sack of Baghdad in 1258, an event with which the disappearance of the caliphate (and not just of the Abbasid dynasty in Baghdad) is generally associated.

Once again, however, as in al-Andalus earlier, there were those who saw potential for themselves in the new situation. The Mamluks, new rulers in Egypt, saw the potential offered by this vacancy, and, profiting both by the availability of Abbasid refugees from Baghdad and from the example of a short-lived experiment in Aleppo (Heidemann 1994), set up a minor official, of Abbasid descent, at their court, with the title of caliph. This official post, always filled by a member of the Abbasid house, remained in existence, with some unimportant and purely formal and ceremonial duties, until the conquest of Egypt in 1517, and its incorporation in the Ottoman empire. By this stage, the word caliph had lost all of its political content: it no longer referred to a ruler, nor suggested the idea of political power in the Islamic world; it had lost the religious meaning that it had had at the start, and the quasi-religious universalizing implications that later Abbasids had tried to substitute for their lost political influence; and, if it had any positive meaning at all, it was that associated with the desire of the regime that used it to identify, in a period when this no longer mattered very much, a distant source in religion for the legitimacy that they may otherwise have felt themselves to lack.

The Ottomans, it is often claimed, asserted an entitlement to the title ‘caliph’, as being in some sense heirs to these minor court officials following their conquest of Egypt. This is mistaken. What they did do, from a later stage, was to make use of the title caliph, along with a host of other titles of different types, as part of an elaborate and variegated titulature which gathered up materials of disparate origins in order to create a new and comprehensive dress for the leader of the Islamic world. In this context the use of the title caliph was merely a way of appropriating one title among many. As before, while it echoed the great days of the early caliphs in Islam, it carried no intrinsic meaning beyond that.

In echoing those great days, however, it did imply a universal role for its holder in the world of Islam. It was in that sense, as seeking to shed precisely that rôle, that the Turkish Republic abolished the caliphal institution in 1925 (Wasserstein 1993a: 156, with further references). In the nineteenth and early twentieth centuries the Ottoman empire declined, and non-
Muslim states, especially Russia and Britain, came to rule over large numbers of Muslims, in Central Asia and India. In this situation the notion of the caliphate came to enjoy a faint new life, as an institution offering less political and more religious sanction to the idea of Islamic universalism. In the first half of the twentieth century, we find the term Khilafat (from the Arabic term for the caliphal institution, khilāfa) serving as the name for just such a movement offering to provide supra-national religious unity for Muslims everywhere via political action in the Indian sub-continent (Kramer 1986). By this time, the term khilāfa has come to resemble the notions contained in the word papacy.

4. Parallel with papacy

It is common in discussions of caliphs and of their institution, the caliphate, to stress that the caliph is not a pope, and the caliphate not a papacy. Yet the similarities are very strong. The titles of caliph and pope both refer to the idea of deputizing for God: Khalīfat Allâh and Vicarius Christi. Both call themselves by titles involving the idea of enslavement to God: ‘abd allâh and servus servorum dei. More importantly than this, the two offices seem to have developed in similar ways: both began with a distinct and characteristic religious element; each acquired (though at different stages) political content, which gave new meaning also to the titles themselves. Later on, as the two institutions came to lose real political significance, their holders were able to cling to vestiges of that significance by reinventing their political role in religious terms, standing above political divisions and offering something around which all could, in theory at least, unite.

5. Conclusion

The developments that we have seen in the history and the meaning of the word khalīfa have many sources. One of the sources that help to make these changes possible lies in the nature of the Arabic language itself. The term khalīfa was understood from a very early stage as a shortened form of a longer expression; while the content and nature of that longer expression offered matter for discussion, the term khalīfa itself, in the senses of ‘deputy’, ‘successor’, was able to serve as a catch-all term for the person holding the office. It is only the missing words, allâh or rasûl allâh, with their
religious implications, which give the word its specialized political meaning here. As a title, it was understood to have content going beyond the literal meaning of the word itself. Unlike such a word as ‘king’, it does not have a single narrow meaning. Indeed, even the word ‘king’, in one of its forms in Arabic, can be seen as meaning no more than ‘possessor’.

Perhaps the most interesting feature of this word-history is the durability of the word, or the institution which it labels – or which, more precisely, it has so often been thought to label. The word has acquired, and lost, a large number and a wide variety of meanings over the last millennium and a half without ever losing entirely the flavor of its origins. Even in such a set of societies as those which use Arabic, a written language with a continuous history for this period, the ability of a word like this to retain the central core of its meaning over so long a period as this is unusual. In this particular case, of course, the religious element has played an important, even a decisive, role, but the purely semantic, lexicographical aspect of the term, shorn of the religious element, remains impressively durable.

References

(This bibliography includes, deliberately, not only items cited in the text but also others which are of help in situating the history of the term and the institution of caliph in their proper historical contexts.)

Ayalon, Ami

Ayalon, David

Azmeh, Aziz

Barth, Jakob

Bosworth, Clifford E. et al.

Classen, Peter
130  David J. Wasserstein


Crone, Patricia and Martin Hinds


Dozy, Reinhart


Fagnan, Edmond


Fathi, Habiba


Gibb, Sir Hamilton Alexander Rosskeen


Heidemann, Stefan


Ibn-Manzûr


Jackson, Sherman A.


Kassis, Hanna E.

Kleinmichel, Sigrid  
2001  

Kramer, Martin  
1986  

Lambton, Ann K.S.  
1977  

Lane, Edward W.  
1863–93  

Lewis, Bernard  
1961  

1988  

Maccarone, Michele  
1952  

Madelung, Wilferd  
1971  

Mawardi, Aboû ’l-Hasan Ali  
1915  

Nadim  
1970  

Niemeijer, Albert C.  
1979  

al-Qâdî, Wadåd  
1992  
Redhouse, Sir James
1890 A Turkish and English Lexicon, showing in English the significations of the Turkish terms. Constantinople: Printed for the American Mission by A.H. Boyajian.

Schaller, Hans Martin

Shaked, Haim

Sourdel, Dominique

Stern, Samuel M.

Tabari

Walker, John

Wasserstein, David J.
<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Title</th>
<th>Notes</th>
</tr>
</thead>
</table>
Words in discourse – On the diachronic lexical semantics of *discours*

*Judith Meinschaefer*

This study presents a lexical semantic investigation of the French noun *discours*, one of the most important cultural keywords in the twentieth century, on the basis of its usage in historical literary texts. First, this paper sketches the semantic evolution of *discours* in the preceding five centuries. Second, it analyzes the usage of *discours* in the writings of two French authors, Michel de Montaigne (1533–1592) and Jean-Jacques Rousseau (1712–1778). It is shown that the semantic evolution of *discours* is characterized by decreasing polysemy and increasing specialization. The results of this study are discussed in the light of Putnam’s (1975) conception of the “division of linguistic labor”.

1. Introduction

Linguistic changes, that is, changes in the structural and symbolic descriptions speakers assign to linguistic expressions, arise not only from processes of reanalysis in oral communication and spontaneous speech, but also come about in written language. Semantic change, one of the most obvious aspects of language change, should be particularly visible in literary and philosophical writings, where speakers are careful and conscious in their choice of words. In such texts, we can expect to find characteristic, innovative and elaborated patterns in the semantic variation of a word. Semantic change is often considered to stand in close relation to socio-cultural and

*This work was supported by the DFG through its Sonderforschungsbereich 471. I would like to thank Peter Kohlhaas for providing me with the database of occurrences of *discours* in the writings of Montaigne and Rousseau which are analyzed here. I would like to thank the editors of this volume, as well as two anonymous reviewers, for helpful comments on an earlier version of this paper. Thanks also to Bruce Mayo for correcting my English. Needless to say, all remaining errors are mine.*
intellectual as well as to technological development. Therefore, many investigations of the history of words in literary or non-literary language use have concentrated not only on the evolution of socio-cultural keywords and philosophical terms, but also on words for objects of functional relevance, e.g., tools or clothing. Such studies have been conducted in various theoretical frameworks, such as history of cultures (Matoré 1985, 1988; Wandler 1970), structuralist semantics (Trier 1931), cognitive or prototype linguistics (Geeraerts 1997; Blank 1997), discourse analysis (Busse 1991) and pragmatic semantics (Fritz 1997). In some studies, the history of socio-cultural keywords, like nation, prolétarier, or responsabilité, is treated as being equivalent to the history of the ideas standing behind these words (for a discussion see Busse 1987). Accordingly, the history of words, like the history of ideas, is often explained with recourse to ideological and social changes that take place at certain points in time, an aspect which is certainly essential to a theory of change in the cultural lexicon of a language.

From this approach to the explanation of the history of words follows that the semantic aspects of the words treated in these investigations belong to the level of encyclopedic knowledge, defined as the sum of everything a speaker knows about natural, mental, or cultural entities, rather than to the level of lexical knowledge, defined as what a speaker knows about lexical entities, i.e., words or phrases. Like the history of words, cognitive linguistics assumes that the meaning of a word is encyclopedic, comprising everything a speaker knows about the referent of a word (e.g., Croft 1993). This is an adequate assumption for a research program that proposes to identify the conceptual relations between different meanings of a word, conceived of as symbolic representations in the minds of the speakers, and the principles governing possible patterns of linguistic polysemy and semantic change. But even though the distinction between encyclopedia and lexicon is highly problematic, if it can be drawn at all, the distinction between how speakers conceive of the world, and how they use words, is crucial for a characterization of semantic change, since the reasons for the evolution of the mental representations and ideas standing behind the words are undoubtedly different in nature from the restrictions on changes in the distributional properties of the words of a given language. What is more important, both aspects of the history of words require distinct methodologies.

The present study takes the latter aspect of word meaning as its point of departure, i.e., the word's combinatorial properties, which so far has received little attention in historical semantics. It will be shown that the combinatorial possibilities of words can be used to gain insights into some as-
pects of their meaning, and that they represent, as we propose to demonstrate, a useful methodological tool for the reconstruction of word meaning in historical texts. The present study analyzes the lexical semantic development of the French noun *discours*, one of the most frequent sociocultural terms in contemporary scientific discourse. This study emerged from an interdisciplinary collaboration between Peter Kohlhaas, who has studied the historical semantics of *discours* from a historic-cultural perspective (Kohlhaas 2000), and the present author, who is a linguist. By looking at data taken from the writings of two French authors of the sixteenth and eighteenth century, this investigation reconstructs some aspects of the semantic evolution of the term *discours*, a history which is, as will be demonstrated, characterized by a reduction in the range of conceptual domains (e.g., space, time, communication) within which a term is used, and an increasing specialization of its usage.

A lexical semantic investigation of the history of words which stresses a linguistic classification of different usages of a word found in historical literary texts presupposes a theoretical approach that distinguishes a word’s distributional and denotational properties from the conceptual entities standing behind the word’s meaning. Therefore, a few remarks are in order about the concept of “division of linguistic labor”, as introduced by the philosopher Hilary Putnam (1975), and the related concepts of “average speaker” and “expert speaker”. According to Putnam (1975: 144–45), not every member of a linguistic community who uses a term like water, aluminium, or elm needs to be able to make a distinction between elms and beeches or between aluminium and molybdenum. Likewise, not every speaker who uses the word water needs to be able to distinguish between H2O and other liquids superficially similar to water. Rather, the “average speaker” can rely on a distinct subclass of speakers within his or her linguistic community, the “expert speakers”, who possess criteria for recognizing elms, aluminium, and water. While average speakers need no more than a minimum level of competence to correctly form and understand sentences containing words like beech and elm, they may leave it to the expert speakers to distinguish the objects to which the sentences can refer, namely elms from beeches. Putnam uses the notion of “division of linguistic labor” to refer to this distribution of competence between the average speaker and the expert.

Seen from a methodological viewpoint, Putnam’s conception allows us to distinguish those semantic aspects of a word that are related to the specialist’s knowledge about the extension of the word from the more strictly se-
mantic aspects of a word that are reflected in its combinatorial properties. While the former pertain exclusively to the specialist’s scientific or philosophical perspective, the latter must be mastered equally by the specialist and the average speaker. To be sure, Putnam’s conception of the division of linguistic labor aims exclusively at a characterization of meaning and reference with regard to natural kinds like tigers or water. This study proposes to show that Putnam’s concept is in principle also applicable to cultural kinds.

The two authors analyzed in this study, Michel de Montaigne and Jean-Jacques Rousseau, can be seen as “experts” whose philosophical systems describe human thought and communication in ways far more specific than those of average speakers. They enjoyed, however, a high reputation among their contemporaries and among the generations of speakers to follow. According to Putnam, we could expect the average speaker to rely on them, as specialists, adopting their criteria for the determination of the extension of discours, just as the average modern speaker relies on the chemist to delimit the extension of molybdenum. This hypothesis will be further illustrated by the findings described here, taking as an example the set of cultural concepts referred to by the term discours.

An additional aspect of Putnam’s conception is of relevance to the present study. According to Putnam, the line dividing expert knowledge from the average speaker’s knowledge is socio-linguistic in nature, and it can therefore shift with scientific progress or with socio-cultural changes. As Putnam argues, scientific progress is likely to lead to an increase in the division of linguistic labor, since the accumulation of knowledge about the necessary and sufficient criteria associated with a natural kind gives more and more authority to specialist speakers in determining the extension of such terms. In this study, we will see that Putnam’s conception of an ever-increasing linguistic division of labor is not only fruitful for the characterization of the semantics of natural kind terms, but can also be applied to the evolution of cultural kind terms like discours.

One restricting remark is in order, though, as to the empirical grounding of this study. It is proposed that the authors analyzed here act as specialist speakers, and that the average speakers adopt their criteria for determining the extension of the term discours. To prove this strong hypothesis, however, would require an extensive historical study, for the following reasons. First, one needs to show that the author in question displays usages of the term which are also found at later stages of its history, but which are in principle not found at earlier stages. Thus, the author has to be innovative
rather than conservative. Second, one needs to show that an author was the first to use the term with a particular meaning, instead of adopting an already existing usage from another specialist or linguistic community. Still, these are only necessary conditions for the thesis, since an innovative usage of a particular term can arise in the language of any speaker, but this usage will vanish from the language if it is not adopted by a larger community of speakers. Note that this usage can be reinvented later, or at the same time, by another speaker. A sufficient condition for the thesis to hold would be fulfilled only if it could be demonstrated that the author in question was not only the first to adopt a particular usage, but also that those speakers who are claimed to rely on the author as a specialist speaker did indeed read his works and took over a certain usage for this reason, instead of creating themselves a novel usage which might have been conceptually closely related to some usage current at their time. This, however, is a philological enterprise rather than a linguistic one, and goes far beyond the scope of this study. Thus, we restrict ourselves here to a detailed linguistic description of the usages of *discours* found in two authors, relating this description to the history of the term as represented in a number of dictionaries, and interpret our findings in the light of Putnam’s conception of the “division of linguistic labor” between the specialist and the average speaker.

The object of the present study, the French noun *discours*, belongs to a class of words denoting oral or written linguistic actions. However, it stands apart from other words in this domain because of its far-reaching socio-cultural and terminological implications, as exemplified on one hand by the concept of *discours* developed by Foucault and others (Foucault 1971; for a review see Maingueneau 1991), which for the last forty years has deeply influenced philosophical reflection on the evolution of knowledge. On the other hand, *discours* belongs to the basic vocabulary of contemporary linguistic terminology, referring, broadly speaking, to linguistic utterances beyond the sentence. In the framework of historical semantics two subclasses of the domain of words denoting oral or written linguistic actions have been studied: both the development of the meaning of speech-act verbs and linguistic action verbs (Schlieben-Lange 1987; Anscombe, Létoublon, and Pierrot 1987; Traugott 1991) and of the vocabulary of argumentation (von Polenz 1988) have been investigated from a pragmalinguistic viewpoint, focussing on the types of speech acts that can be referred

---

1 See Geeraerts 1997: 62–68 for a discussion of what he terms “semantic polygenesis”, i.e., recurrent semantic innovations that appear to remain incidental.
to by different verbs. In contrast to previous studies, this investigation stresses the linguistic classification of the usages of a single word at various points in time, based on the analysis of its combinatorial properties. By looking at a variety of linguistic data taken from the writings of two French authors (more than 150 occurrences of \textit{discours} drawn from the \textit{Essais} by Montaigne, and more than 400 occurrences drawn from different texts by Rousseau), in addition to considering information on the usage of the term \textit{discours} that can be found in historical dictionaries, the present study aims at developing a methodology for the linguistic classification of different usages of a word in literary texts of the past.

The presentation is structured as follows. First, a characterization of the meaning of \textit{discours} in contemporary standard French is given, followed by an overview of the history of this term from early modern French to the language of our time. Then the variation in the use of \textit{discours} is described as it appears in the writings of Michel de Montaigne (sixteenth century) and Jean-Jacques Rousseau (eighteenth century). For each author, a classification of the different usages found is developed. In the concluding section, both authors' usage of the term is compared and discussed in the context of a history of the term \textit{discours} from the sixteenth century to contemporary French. Finally, the findings of this study are related to Putnam's (1975: 145) thesis that with the rise of science the division of labor between average speakers and expert speakers continuously increases.

2. On the history of \textit{discours}

This section gives an overview of the semantics of \textit{discours} in modern French, as well as a short history of its evolution in the course of the last five centuries. The data presented in this section are taken from a variety of dictionaries (see section 7), already containing classifications of different usages of the term \textit{discours}. A schematic overview of the semantic evolution of \textit{discours} is given in Figure 1. In contemporary standard French, the noun \textit{discours} has three closely related readings. First, it can refer to a rhetorical performance, i.e., a “speech” (see Figure 1, 1. rhetorical performance), as in expressions like \textit{discours inaugural}, \textit{de clôture}, \textit{de réception} ‘opening, closing, reception speech’, or \textit{faire, lire, prononcer un discours} ‘hold, read, speak a discours’. Second, it denotes a written text representing such a rhetorical performance, as in \textit{écrire, rédiger un discours} ‘write, edit a speech’.
Third, it can be used, with a pejorative connotation, to denote the linguistic expression of one’s thoughts in conversation (see Figure 1, 2.), implying that the speaker’s expositions are lengthy and redundant, as shown in (1).

(1) *Il aime à parler, et avec tant de vêhémence et si longtemps, que l’on a soif pour lui – mais gardez-vous d’abonder dans son sens, car il en profitera pour réprendre son discours à l’envers sous prétexte que vous n’avez rien compris.* [Bianciotti 1995, FT2]

In addition to its use in the standard register, *discours* has readings that pertain to specialist terminologies (see Figure 1, 3.). Among others, its specialist usages comprise the following. In linguistic terminology it denotes a particular register or style of speaking (*discours soutenu, familier* ‘elaborated, familiar speech’). Furthermore, it is used to refer to the techniques of

---

2 In citations, abbreviations (see section 7) are used which refer to the dictionaries from which the examples are drawn.

3 ‘He likes to talk, and with such vehemence and so long, that one thirsts with him – but avoid agreeing with him, because he will take advantage of this to recommence his *discours* backwards on the pretext that you have understood nothing.’ [my gloss]
reported utterance (discours direct, indirect ‘direct, indirect speech’). In traditional linguistic terminology, the phrase parties du discours is used to make reference to the parts of speech, i.e., lexeme classes. In semantics, the univers de discours means the ‘universe of discourse’. The linguistic uses of the term all relate to the abstract meaning of “observable linguistic utterance (beyond the sentence)”. In philosophical terminology, the noun discours, relating to the adjective discursif ‘discursive’ (as in pensée discursive ‘discursive thought’; discursif < lat. discursivus, adj., derived from the Latin noun discursus), refers to a way of thinking that reaches its object via a sequence of propositions organized according to the laws of logic. Finally, in sociological, historical, and literary terminology, discours refers to communicative action and argumentative practices beyond the level of the individual. As suggested in the introduction, a precise explication of the meaning of discours within each of these specialist terminologies would require an exposition of the relevant theories rather than a lexical-semantic investigation.

The semantic variation present in contemporary French usage of the term discours has emerged during a history going back to early modern French. The first citations of the noun discours date from the early sixteenth century. The noun discours is derived from the Latin noun discursus, which is a nominalization of the Lat. verb discurrere > old Fr. discourre > modern Fr. discourir.

While the history of the noun discours begins only in early modern French, the verb discourre can be traced back to the 12th century, originally meaning ‘to run here and there’. In this meaning, discourre is closely related to other movement verbs like aller ‘go’, vaguer ‘wander’ and naviguer ‘navigate’, and, like these, it combines with non-directional prepositional phrases denoting unbounded paths or regions. In old French, discourre could also denote the action of understanding, and of reporting a sequence of facts. It is perhaps interesting that the Alfranzösisches Wörterbuch (Tobler and Lommatzsch 1925–76) claims that in its spatial meaning old French discourre is derived from Lat. discurrere ‘run here and there’, while in the intellectual and communicative meanings it derives from Lat. discutere ‘discuss’, an assumption not shared by etymological dictionaries of French, like e.g., the Dictionnaire historique de la langue française (Rey 1994), which restrict its origin to the Latin verb discurrere.

In the sixteenth century the noun discours, like its base, the verb discourir, could be used to denote an action of moving, as in faire des discours et circuits, or the path of movement, as in the expression connaître tous les dis-
cours. This semantic shift from the denotation of the movement to the path of movement is also shown by other nominalizations like trajet ‘trajectory’ or voyage ‘voyage’, to cite only a few.

As early as in the sixteenth century, i.e., shortly after its borrowing into French, the term discours shows a remarkable polysemy. The Dictionnaire de la langue française du seizième siècle (Huguet 1928–67) gives nine readings of the term, the first two belonging to the spatial, the third to the temporal domain, the next two to the domain of speaking and writing, and the last four to the mental domain. In its meaning of ‘movement in space’ (see Figure 1, 4.), discours can in particular be combined with a complement referring to a celestial body, as in discours du soleil, de la lune ‘path of the sun, of the moon’. According to the same dictionary, discours may refer not only to the progress of a movement or voyage in space, but, when combined with an event noun, also to the progress of an event in time (see Figure 1, 5.), as in the expression discours du tournoy, du bal, du procès, discours du mois, des ans, de la vie ‘course of the tournament, ball, trial, month, year, life’. In these constructions, the noun discours serves to focus upon the progression of the event denoted by the prepositional phrase, as opposed to its completion.

However, discours refers not only to a movement in space or the progress of an event in time, but may also denote a report thereof (see Figure 1, 6.), as exemplified in (2).

(2) Le discours de ceste bataille est publié en tant de lieux ...
[Monluc, DLFSS]4

Here, the construction discours de + event noun is combined with a predicate (être publié ‘be published’) that selects for nouns denoting (linguistic) descriptions, but not movements or other events. Hence, the phrase discours de + event noun is interpreted as referring to the report of an event, not to the event itself. Apart from referring to movement in space, progress in time or linguistic descriptions thereof, discours is used in the sixteenth century to refer to mental phenomena. More precisely, discours denotes the faculty of reason (see Figure 1, 7.), but refers also to intellectual education, as in homme de discours ‘man of discourse’. In the latter meaning, it is used only in (syntactic) compounds, therefore this meaning is not separately considered here. Finally, the term discours refers to the action of rea-

4 ‘The discours of this battle is published in many places ...’ [my gloss]
The citation in (3) gives an explicit definition of the term *discours* as referring to the action of reasoning as understood by the author Charron, from which it can be inferred that *discours* firstly belongs to the same semantic type as *action, office, exercice* ‘action, office, exercise’, i.e., it is an action noun. Secondly, a *discours* consists in a sequence of actions, i.e., collecting things, dividing them, and adding more. Thirdly, *discours* can be synonymous to *ratiocination* ‘ratiocination’.

(3) *L’action et l’office ou exercice de cette force et puissance qui est d’assembler, conjointre, separer, diviser les choses receues et y en adjouter encore d’autres, c’est discours, ratiocination.*

[Charron, DLFSS]5

While for the sixteenth century there is ample evidence for the term *discours* as referring to movement in space and progress in time, in the seventeenth century the spatial and temporal senses fade, giving way to those meanings related to the domain of reporting, speaking, and writing, and the meanings related to reasoning become marginal. The *Dictionnaire universel* (Furetière 1978) lists two meanings of *discours*, one referring to the action of oral linguistic expression (see Figure 1, 9.), by which are understood “tant des *discours* oratoires, tant des entretiens familiers” ‘rhetorical performances or familiar conversations’, the other to a written text (see Figure 1, 10.). These readings of *discours* have been passed down to contemporary French, and are also found in the usage of Montaigne and Rousseau, which will be considered more closely later on. Note that a usage referring to linguistic conversation is already attested for the Latin noun *discursus* (Rey 1994). The *Dictionnaire François* (Richelet 1970) adds another usage of the term *discours*, referring to a person’s style of speaking (see Figure 1, 11.), as exemplified in (4).

(4) *Sans cesse en écrivant variez vos discours: un style trop egal et toujours uniforme en vain brille a nos yeux, il faut qu’il nous endorme.*

[Boileau, PGCF]6

5 ‘The action and the office or exercise of this force or power which is to assemble, join, separate, and divide things received and to add to them others as well, this is *discours*, ratiocination.’ [my gloss]

6 ‘In writing, continuously vary your *discours* (pl.): a style too even and always uniform vainly glitters in our eyes, causing us to fall asleep.’ [my gloss]
The article on discours in the Encyclopédie of Diderot and D’Alembert (1751–65) points out two prominent senses in the eighteenth century. In one sense, discours is defined as “tout ce qui part de la faculté de la parole” ‘everything emitted by the faculty of speech’. This definition refers to the diction of the “average speaker” of their time, the Enlightenment. Interestingly, a false etymology is given for the noun, as it is claimed to derive from the verb dicere in the meaning of “dire, parler” ‘say, speak’. In a second sense, discours refers to a genre, synonymous to harangue ‘harangue’ or oraison ‘oration’. As becomes clear in the article, the second definition refers not to the diction of the average speaker, but to the knowledge the expert in rhetoric has about the concept of discours. While the article does not elaborate on the first usage, i.e., discours as denoting entities and actions belonging to the domain of speaking and writing, the second usage is described in detail, concentrating on the various subtypes of the genre, and on the structuring and composition of a discours (see Figure 1, 12.). The information contained in the Encyclopédie would thus be of interest for reconstructing the specialist’s conception of discours in the era of Enlightenment, but is of restricted relevance to the object of the present study. Finally, as to the usage of discours in the nineteenth century, Mots et dictionnaires (Journet, Petit and Robert 1966–78) lists only two meanings, one relating to the oral expression of ideas, the other to the written expression of one’s thoughts on a particular subject.

The semantic evolution of discours can be summarized as follows. We find first a set of readings that appear only in texts of the sixteenth century. These readings extend “discursive motion” to a broad range of spatial, temporal and mental conceptual domains. In newer texts, dating from the seventeenth to twentieth centuries, we find a second group of readings in which the polysemic range of discours is restricted to various kinds of “rhetorical performance”, while outside this range we find a set of specialist meanings, interpretable only in the conceptual frameworks of sciences like grammar and sociology. A third group of readings, relating more generally to the oral or written expression of one’s thoughts, is found in both periods.

In the sixteenth century discours displays its most extensive polysemy (that is, the range of the conceptual domains its different usages pertain to). Later on, the morphological transparency of discours, relating it to the base verb discourir ‘to run here and there’, as well as to the verb courir ‘run’
and noun *course* ‘action of running’, fades. As a result, the non-linguistic, i.e., spatial, temporal and mental domains are lost, with only those usages remaining which relate to the linguistic expression of ideas. With the beginning of the seventeenth century, the available polysemy of the noun *discours* diminishes significantly. From the seventeenth century onwards, *discours* refers exclusively to the conceptual domain of communication. Third, the usage of *discours* in contemporary French, in contrast to earlier periods, is marked by its extensive use in various specialist terminologies, and by the negative connotation it can have in stylistically neutral contexts.

In the following, we examine more closely the usage of *discours* in the diction of two influential authors of the sixteenth and eighteenth century, i.e., Montaigne and Rousseau. While the preceding section has drawn exclusively on information from dictionaries, in the two following sections we will examine a number of citations containing the noun *discours*, taken from various texts of the two authors. Here, the question of how information about the meaning of a word can be inferred from the analysis of historical texts will be addressed explicitly, with the aim of developing a methodology for the semantic analysis and classification of different usages of a word in literary writings.

### 3. Discours in the writings of Montaigne

Michel de Montaigne was born in 1533 to a wealthy merchant family in Périgord, who recently had become members of the aristocracy. After studying law, he became a *conseiller* at the *Parlement de Bordeaux*. In 1571, he retired to his country estate, where he began to write the *Essais*, three books which were written and published between 1572 and 1588. Previously, he had translated and published a treatise of Raymond Sebond, a Spanish theologian of the fifteenth century. In 1581, he became mayor of Bordeaux, following his father, who had held this office from 1554 to 1556. Montaigne died in 1592, and a posthumous edition of the *Essais* with the author’s corrections and additions appeared in 1595. All citations analyzed here are taken from the *Essais*.

---

7 Note that the same loss of morphological transparency has occurred for the term’s predecessors, the Latin terms *discurrere* and *discursus*, which could be used to refer to a linguistic action (Rey 1994, *discourir*). Likewise, for the French verb *discorre*, this sense is found from the thirteenth century on (Rey 1994, *discourir*).
3.1. *Discours* as process of reasoning

When one has to explain the meaning of a word like *dog*, *table*, or *spoon* to a person unfamiliar with these words, the quickest (but possibly not the most adequate) method may consist in pointing to a dog, table, or spoon, and informing the person that entities of this kind are in general called *dog*, *table*, or *spoon*. When concerned with the analysis of historical texts, one can proceed likewise, not pointing to real dogs, tables, or spoons, but to linguistic descriptions thereof in the text. Therefore, the first example considered here is a citation in which the noun *discours* is used to refer to the description of a process of reasoning which Montaigne presents in the text (see also Figure 2, 1.).

Figure 2. The semantics of the noun *discours* in the writings of Michel de Montaigne. The usages are numbered according to their occurrence in the text.

In the citation given below, he describes first the behavior of a dog, then the mental process he believes to take place in the dog’s mind. According to the author, the mental process which underlies the dog’s behavior must be understood as a kind of *discours*. 
(5) ... considering the movements of the dog which, finding itself at an intersection of three paths, ... seeking his master whom he has lost, ... goes trying one path after the other, and, after assuring itself with respect to two of them, not having found the sought scent, runs into the third without hesitation, he [the observer, J.M.] is forced to admit that such a discours takes place in this dog: I have followed my master just up to this intersection; it is necessary that he continued by one of these three paths; it is not this one nor that one; therefore incontrovertibly he has taken the other path; ...' [my gloss]
se passe en ce chien” ‘such a discours goes on in the head of this dog’. In contemporary French, the verbal predicate se passer may be combined only with words denoting situations with a certain temporal extension [PR, passer]. In another passage from the Essais, discours is used in the slightly different construction “NP + passer + par + NP”, as in “ce discours lui passe par la tête” ‘this discours goes [for him] through the head’ (Essais II, Chapitre 12, 460, A). In contrast to reflexive passer, in contemporary French the non-reflexive verb passer can be combined with words denoting entities and with words denoting situations [PR, passer]. Hence, in modern French the reflexive verb se passer in the first construction can refer only to the course of an event in time, and can only be combined with an event noun, while the non-reflexive verb passer in the second construction can both refer to the movement of an entity in space, and to the course of an event in time. To be sure, we lack comprehensive lexicographical information on the selectional restrictions of passer in the sixteenth century. Assuming that these are identical to modern French, we may still conclude that while in the first construction discours must denote an event, i.e., an event of inner reasoning, in the second case it might also be understood as denoting an abstract or concrete object, i.e., an ordered collection of ideas, concepts, or thoughts, moving through some mental space.

In most cases discours refers, to be sure, to the reasoning and argumentation of intelligent and even educated human beings. However, in the Essais Montaigne explicitly argues in various passages for the view that not only human beings, but also animals like dogs, foxes, elephants, or birds may be said to show the kind of reasonable behavior denoted by discours. When referring to the inner reasoning of an individual, discours is often used in the prepositional phrase par discours, functioning as the adverbial adjunct of a causative (and agentive) verb, as in appuyer, fortifier, augmenter, encroûter, épaissir par discours, where the direct object of the verb refers to a mental state or property, as exemplified in (6).

(6) Je me deffais tous les jours par discours de cette humeur puerile et inhumaine, qui faict que nous desirons d’esmouvoir par nos maux la compassion et le deuil en nos amis.
(M: Essais III, Chapitre 9, 979, B)9

9 ‘Every day I free myself by discours of this puerile and inhuman humor, which forces us to want to awaken compassion and grief in our friends by our misfortunes.’ [my gloss]
In this meaning, *discours* stands in coordination with nouns derived from agentive verbs like *étude* (*par étude et par discours*), and in contrast to nouns derived from non-agentive verbs like *sentiment* (“je ne me juge que par vrai sentiment, non par discours” (*Essais* III, Chapitre 13, 1095, B)), indicating that *discours* refers to a mental activity that is controlled by an agent, i.e., to a deliberate intellectual effort eventually showing causal effects within the mind. Hence, in the usage described in this section, *discours* can denote a process of reasoning, which can also be viewed as an object, but it can also be interpreted as denoting an intellectual activity.

3.2. *Discours* as the faculty for reason

In Montaigne’s usage, the term *discours* not only refers to the process or activity of reasoning itself, but also to the underlying ability of reasoning (see Figure 2, 2.). The activity presupposes the ability; hence, the relation between these two usages of *discours* may be termed metonymic. By “metonymy” we understand the relation between two words referring to entities which are contiguous in some sense (Ullmann 1957: 232; Geeraerts 1997: 71ff.). Sometimes it is argued that the meaning of a word corresponds to its position in a network of lexical semantic relations, like synonymy, hyperonymy, or antonymy. Under this assumption, the meaning of a word can be reconstructed by analyzing how its lexical semantic relations are embedded in the network of meanings in a text. In the following we will therefore look at words which stand in a relation of semantic similarity or contrast to *discours*. In various passages, the noun *discours* stands in opposition to the noun *corps* ‘body’, but also to the noun *âme* ‘soul’, thus creating a contrast between the emotional (âme), mental (*discours*), and physical (*corps*) aspects of the human being, as in (7).

(7) ... j’estoy blessé à mort par la teste. C’eust esté sans mentir une mort bien heureuse, car la foiblesse de mon discours me gardoit d’en rien juger, et celle du corps d’en rien sentir.
(M: *Essais* II, Chapitre 2, 376/377, A)\(^{10}\)

\(^{10}\) ... *my head were mortally injured. It really would have been a fortunate death, because the weakness of my *discours* would prevent me from judging anything, and the weakness of my body, from sensing anything.* [my gloss]
The first thing to note is that here *discours* denotes not an event, but an object, since it can be qualified by the noun *faiblesses* ‘weakness’. Second, the above citation implies that *discours* and *corps* are of the same semantic type, both are assigned the same property, and are characterized by a certain function. While it is the function of the *discours* to understand, *juger*, the function of the body is to feel, *sentir*. Both terms show a certain contrast in meaning. This sort of semantic contrast, however, does not necessarily point to a lexical relation; it may well be the expression of some theological or ontological conception of animate beings in the world.

In this meaning of *discours*, the determiner of the noun refers to the possessor of the faculty, as in *mon discours*, *notre discours*, and the noun cannot be pluralized. A more precise definition of *discours* as referring to a faculty is given in (8).

(8) *Certes, nous avons estrangement surpaié ce beau discours dequoy nous nous glorifions, et cette capacité de juger et connoistre ...* (M: *Essais II*, Chapitre 12, 486, A)\(^ {11} \)

Here, the two noun phrases *ce beau discours* ‘this beautiful *discours*’ and *cette capacité de juger et de connaître* ‘this ability to judge and understand’ are coordinated, with an intersecting relative clause. Both noun phrases appear as objects to the verb *surpayer*, where *discours* appears as the definiendum, and the second noun phrase as the definiens, giving first the general term, i.e., *capacité* ‘faculty’, and then the specific difference, i.e., *de juger et de connaître* ‘for judgement and knowledge’. In other words, *discours* is a hyponym of *capacité*; hence, a *discours* is a faculty. This usage of *discours* is attested for other writers of the sixteenth century (see section 2; see also see Figure 1, 7.); however, it remains marginal and is quickly lost.

3.3. *Discours* as an oral or written linguistic utterance

The mental process referred to by *discours* stands not only in a close conceptual relation to the underlying faculty, but also to the linguistic, overtly observable action of reasoning, and its product, oral or written utterance.

\(^{11}\) ‘Certainly, we have strangely overpaid for this beautiful *discours* for which we praise ourselves, and this capacity to judge and to know ...’ [my gloss]
The third usage of the term *discours* in the writings of Montaigne, which is described in this section, refers thus to the linguistic expression of one’s thoughts, or to the underlying set of propositions (see Figure 2, 3.). From a synchronic viewpoint, the relation between inner reasoning and its linguistic expression may be termed metonymic, since the linguistic expression of one’s thoughts presupposes the existence of the inner process of reasoning. Note, however, that historically, the relations between the two readings of *discours* are not so clear, since the usage referring to linguistic expression arose already for Latin *discurrere* and French *discourir* (thirteenth century, Rey 1994), and it is only for the French noun *discours* that the meaning ‘process of reasoning’ occurs prior to the meaning ‘linguistic action’. Since this study is not primarily concerned with an account of the conceptual links between different usages of a word and the possible directions of semantic change, we will leave this problem aside.

In this meaning, *discours* is often used as the direct object of the agentive verb *tenir*, a construction which belongs to a paradigm of verb-noun constructions denoting linguistic actions controlled by an agent, such as *tenir des propos* and *tenir un langage* ‘make a speech’. Accordingly, *discours* can also be coordinated with action nouns like *enhortements* ‘encouragement’, and it can be the argument of verbs that presuppose an agent, like *s’appliquer* ‘undertake’. In this usage, a possessive determiner of the noun *discours*, as in *ses discours*, is interpreted as referring to the agent of the action denoted by the head noun of the phrase. However, *discours* can be coordinated not only with action nouns like *enhortements*, but also with nouns referring to objects, as with *sentence* ‘utterance’ in (9).

(9)  ... recueillant leurs sentences et leurs discours comme des oracles  
     ...\(^{12}\) (M: *Essais* II, Chapitre 12, 438–439, A)

Here, the nouns *sentence* and *discours* refer to propositions, with the possessive determiner referring to the author of the proposition, a participant represented in some conceptual frame or script which is linked to the (lexical as well as encyclopedic) meaning of the noun in question. As represented in this frame or script, an utterance referred to by *discours* has not only an author, but also a topic, which can either be expressed by means of a prepositional phrase headed by *sur*, as in “quelque discours sur la considération de la vaillance et différence de cette vertu aux autres” ‘some dis-
**Words in discourse** 153

cours on the consideration of the worth and difference of this virtue relative to others’ (Essais II, Chapitre 7, 383, A), or by an adjective of very general meaning, as in discours politiques ‘political discours (pl.)’, discours éthiques ‘ethical discours (pl.)’. In the second case however, the adjective not only refers to the topic of the utterance, but more generally to the prototypical situation in which it is uttered, implying a number of additional properties typical of particular types of texts or utterances.

In the citation above, (9), both nouns (discours and sentence) refer to countable objects, can be pluralized, and thus are suitable objects for the verb recueillir ‘collect’. Likewise, in this usage, i.e., as referring to countable objects, discours is compatible with predicates expressing a local relation, where it can be locatum, as with the verb se trouver (les discours qui se trouvent dans les livres), or with adjectives like farci ‘filled’ and plein ‘full’ (“des lettres farcies” and “pleines de beaux discours” (Essais I, Chapitre 40, 252, A)). Like other countable nouns, it can also itself identify a location, as for the verb découvrir (“la faute qu’il descouvvrira en son propre discours” (Essais I, Chapitre 26, 155)).

As a product of purposeful action, a discours can contain mistakes, and it can be in error. The type of text Montaigne calls discours is thus nonfictional in nature: it contains propositions, the truth of which can be determined with respect to some non-linguistic reality. Montaigne in fact often uses the noun discours to refer to portions of his own reflections, by means of a prepositional phrase denoting the final phase of some process (au bout de ce discours, pour la fin de ce discours ‘at the end of this discours’). Hence, for Montaigne his own writings belong to the category of discours.

In stylistic analyses of the Essais, it has often been argued that Montaigne’s text constitutes an intermediate form of expression, incorporating both properties of written texts and of spoken utterances (Starobinski 1986). Similarly, in Montaigne’s usage discours can refer equally to oral text or written text, and no commitment with respect to the modality of expression is made. Still, in numerous instances, Montaigne’s formulation allows inferences about the stylistic dimension of discours. This usage of discours anticipates a semantic restriction that becomes much stronger in the following centuries.
3.4. Discours as a genre

In a few cases, Montaigne uses the term discours in a sense (see Figure 2, 4.) parallel to that of other words denoting literary genres, like conte, vers, and poésie. A generic quality seems implied, since in this meaning discours appears either in the plural with the definite article (les discours de la philosophie), or without a determiner (en discours). Grammatically speaking, these noun phrases are used for generic reference. However, this usage of discours is of marginal relevance in Montaigne’s diction, in contrast, as we shall later see, to the usage of this word by Rousseau. To be sure, the precise way in which discours refers to a literary genre must rely on expertise in literary or genre theory, and lies beyond the scope of the present investigation.

3.5. The course of discourse

As shown in section 2, in the sixteenth century the term discours was used to denote an action of moving, or the path of movement (see Figure 1). Montaigne no longer uses discours in this sense. However, he certainly is familiar with the word’s etymology, and the spectrum of its usages in the French of his century, as well as with its use in Latin and Italian. Taking a closer look at Montaigne’s diction, it becomes clear that he sometimes uses the term in the context of words denoting causation of movement, such as tirer, suivre, and path of movement, such as voie and chemin. An example is given in (10).

(10) Mais je laisse ce discours, qui me tireroit plus loin que je ne vou[drois suivre. (M: Essais II, Chapitre 12, 488, A)\(^{13}\)

Here, discours refers to what the narrator has said before, that is, to the sequence of propositions preceding the citation in (10). The combination of discours with the verbs tirer ‘pull’ and suivre ‘follow’ evokes a spatial configuration with a moving object, the speaker, moving along an indefinite path, the discours or the text, where the movement is brought about partly by the flow of text and partly by the speaker, a configuration which may be

\(^{13}\)‘But I abandon this discours, which would take me farther than I would want to continue.’ [my gloss]
understood as a metaphor for the process of text production. Hence, even though in Montaigne’s diction *discours* in its literal meaning never refers to a movement in space or time, he uses this term in the context of other spatial terms, where the whole construction is intended to receive a metaphorical interpretation (see Figure 2, 5.). Thus the etymological underlying semantic structure ‘move purposely, but in a not necessarily direct way’, is still very often the basis from which his particular, specialist usage arises.

3.6. Summary

In this section, we have distinguished five readings of the term *discours* as used by Montaigne (see also Figure 2). First, *discours* refers to the mental process or activity of reasoning, and second to the faculty presupposed by this process. Third, *discours* denotes the action of linguistic expression, but also its product, an oral or written utterance, and its propositional content. This usage is restricted to utterances which can be true or false, and which stand in a close relationship to reason and intellect. Fourth, it was said that in some contexts *discours* denotes a particular type of written text, i.e., a literary genre.

To conclude, Montaigne uses the word *discours* to discuss his concept of thought and reason. To be sure, he no longer uses the term in its spatial or temporal meaning, as his contemporaries still do, although his usage occasionally shows remnants of this older meaning. Montaigne uses only a subset of the variety of readings which were in use in the sixteenth century. Thus, relative to the use of the term by his contemporaries, it appears as if he reduces the domain of application of the term *discours*, explicitly and systematically restricting the class of possible referents of this term. It is exactly these restrictions that constitute the usage of a specialist speaker, the distinctive criteria that determine the extension of the noun, and that thus define the specific place that *discours* holds in Montaigne’s philosophy. Starting in the sixteenth century the semantics of *discours* has undergone considerable changes, which in the following will be exemplified with regard to the usage of the term by the eighteenth century author Jean-Jacques Rousseau.
4. Discours in the writings of Rousseau

Jean-Jacques Rousseau was born in Geneva in 1712. Starting in 1728, he lived mainly at the house of Mme de Warens at Chambéry, interrupted by stays in Switzerland and Italy, and went to Paris in 1742. Rousseau’s career as a man of letters began in 1750, when he was awarded a prize by the Académie de Dijon for an essay on whether “le rétablissement des sciences et des arts a contribué à épurer les mœurs”. This essay made him famous overnight, but he failed to gain an award with another essay, *Sur l’origine et les fondements de l’inégalité* in 1754, which was however a success after publication. In the same year he returned to Geneva, where he wrote, among other works, the *Lettre à M. d’Alembert* (1758), in which he argues against the establishment of a theatre in Geneva, *Julie ou la nouvelle Héloïse* (1761), a novel on passionate love and rigorous duty, *Émile ou de l’éducation*, a novel propounding new principles of education, and *Le contrat social* (1762), a treatise of political theory. Threatened with arrest on account of his views, in 1767 he went to England on the invitation of David Hume, and returned to France in the following year, where he completed and published the *Confessions*, an autobiographic work. Rousseau died in 1778 at Ermenonville. The occurrences of the term *discours* analyzed in the present study have been taken from a number of writings, including those mentioned above.

4.1. Discours as oral linguistic expression

For the historian of ideas, *discours* certainly has a meaning in the writings of Rousseau that is intimately tied to the Enlightenment. But as we shall see, Rousseau actually uses the word in ways that are not so different from those of his contemporaries. As in the previous discussion, we take as our point of departure a citation in which Rousseau refers to a linguistic representation of a rhetorical performance as an instance of *discours*. In (11), the referent of the term *discours* is a passage of direct discourse, that is, a quoted sequence of propositions (see Figure 3, 1.).
“Voyez donc”, continua-t-il, “combien tout ce que vous pouvez me dire est à présent hors de propos. Voyez si des préférences que la pudeur désavoue et quelque feu passager de jeunesse peuvent jamais être mis en balance avec le devoir d’une fille et l’honneur compromis d’un père. S’il n’était question pour l’un des deux que d’immoler son bonheur à l’autre, ma tendresse vous disputeroit un si doux sacrifice; mais mon enfant, l’honneur a parlé et dans le sang dont tu sors, c’est toujours lui qui décide.” Je ne manquais pas de bonne réponse à ce discours; ...\[my gloss\]

As can be inferred from an analysis of the structuring of the sequence of propositions given in (11), the referent of discours here is a monologue, a

---

14 “But look”, he continued, “how everything you are able to tell me is now beside the point. See whether those inclinations which shame disavows and the ephemeral passions of youth can ever hold the balance with the obligations of a young woman and the compromised honor of a father. If it were only a question for one of the two to sacrifice his happiness to the other, my sympathy would contend such a sweet sacrifice; but, my child, honor has spoken, and for those of your blood, it is always honor which decides.” I was not at a lack for a good response for this discours; ...’ [my gloss]
well structured and elaborated piece of argumentation directly addressing the hearer, with the aim of convincing by a particular argumentative strategy. In (11), the noun *discours* refers either to the utterance itself, i.e., the sequence of sentences, or to its informational content, three facets of the meaning of a word that Bierwisch (1983) has discussed as instances of conceptual shifts. The content of an utterance can be communicated to a third person, hence in Rousseau’s usage the noun *discours* can occur in constructions like *rendre compte d’un discours* ‘report a speech’, *un discours parvient à qn* ‘someone hears about a discours’ or *rapporter un discours* ‘report a speech’. The sequence of propositions constituting a *discours* may be true or false, and thus can be denied (*démentir un discours*) or refuted (*réfuter un discours*). Apart from its truth value, the content can be qualified with regard to additional dimensions; hence in this usage a *discours* can more generally be judged (*juger un discours*), condemned (*censurer un discours*), or weighed (*peser les discours de qn*). Note that *discours* can be not only the complement of the verb *démentir*, as in “leurs actions démentent leurs discours” ‘their actions contradict their discours’ (*La nouvelle Héloïse*, II, 468), but also the subject of this verb, as in “ce discours démentit ... les brillantes espérances” ‘this discours contradicted the brilliant hopes’ (*Les confessions*, I, 92). In both cases, *discours* denotes what is said, not the action of speaking. Note also that in the usage of Rousseau, *discours* refers exclusively to oral expression, in contrast to what was said with respect to Montaigne. Still, as will be shown in section 4.3., Rousseau also uses *discours* for reference to written texts, but only to texts which pertain to a literary genre, the genre of *discours*.

However, in Rousseau’s diction *discours* can refer not only to the content of utterances, i.e., to propositions, but also to actions that have a particular spatial and temporal extension. Therefore, the noun *discours* is a possible object of the aspectual verbs *commencer* ‘begin’, *continuer* ‘continue’, and *reprendre* ‘take up again’. Furthermore, since *discours* can be the complement of constructions like *au milieu de* ‘in the middle of’ and *à la fin de* ‘at the end of’, we can infer that the structure of the action denoted by *discours* has a medial phase and a final phase which may temporarily and spatially coincide with a second event. Likewise, PPs like *après ce discours* ‘after this discours’ are used as time adverbials.

From the fact that the referent of *discours* has a temporal and spatial extension it follows that the action can take place in the presence of non-involved third persons (*assister à un discours* ‘attend a discours’) or be watched by an observer (*témoin de ses discours* ‘witness of his discours’).
Since in this usage, *discours* refers to an action, it can occur in the agentive verb-noun construction *tenir un discours à NP* ‘make a speech to NP’, a type of construction which in the corpus analyzed here appears more than 25 times. Likewise, when followed by a prepositional phrase introduced by *de*, the noun phrase contained in the prepositional phrase is interpreted as referring to the agent of the action denoted by the superordinate noun *discours*, as in *discours des plus sages* ‘discours of the wisest’, *discours d’un menteur* ‘discours of a liar’. The same holds for the combination of *discours* with a possessive determiner, such as *ses discours*.

In many passages, Rousseau uses *discours* in constructions where coordinated phrases have strictly parallel syntactic structures, but contain semantically contrasting lexemes. In these passages, a general opposition is created between the noun *discours* and other nouns denoting the concept of ‘action’ on a fairly high level of abstraction, as exemplified in (12).

(12) *En général il y a beaucoup de discours et peu d’action sur la scène française; peut-être est-ce qu’en effet le François parle encore plus qu’il n’agit...*  
(R: *La nouvelle Héloïse*, seconde partie, lettre XVII, *Œuvres* II, 253)\(^{15}\)

In (12), *discours* contrasts with the noun *action* ‘action’, just as the verb *parler* ‘talk’ contrasts with the verb *agir* ‘act’, implying that *discours* is of the same superordinate semantic class as *action*, and that both nouns belong to the same level of abstraction. Note that *action* (sg.) is clearly used as a mass noun here, while *discours* is ambiguous between the reading as mass noun (sg.) and count noun (pl.), for lack of a distinctive plural marker and determiner. While in (12), *discours* refers to an action of speaking, located in space and time (*sur la scène française*), it can also stand in contrast to the noun *œuvres* ‘works’, denoting the product of an action, as in “les œuvres des hommes ne ressemblent guère à leurs discours” ‘men’s works hardly resemble their words’ (*La nouvelle Héloïse*, II, 254-5). In previous sections, it was argued that the analysis of the network of lexical semantic relations is not always useful for the reconstruction of the strictly lexical aspects of meaning. Here, in contrast, the analysis of contrasting terms allows for the reconstruction of one aspect of the denotation of the term *dis-

\(^{15}\) ‘In general there is a lot of discours and little action on the French scene; it may be that in fact the French talk more than they act...’ [my gloss]
cours, i.e., the level of abstraction on which it refers to entities. In Rousseau’s usage this level of abstraction is much higher than was the case for Montaigne, who uses discours mainly for reference to linguistic utterances centered around the concept of reasoning and thought.

Since the article on discours in the Encyclopédie defines discours as “tout ce qui part de la faculté de la parole” ‘all that stems from the faculty of speech’, it may be concluded that the diction of Rousseau does not differ much from that of his contemporaries. It should also be noted that the un-specific meaning discours has in (12), denoting communicative action in general, is comparable to its usage in modern socio-cultural terminology, where it refers to communicative action and argumentative practices beyond the level of the individual.

4.2. Discours as a rhetorical performance

In the preceding section it was claimed that discours can refer to an oral linguistic utterance directly addressing a specific hearer, with a certain argumentative structure and with the aim of persuading the addressee, but that it can also denote the concept of communication on a fairly high level of abstraction. But for Rousseau, the term discours can also have a more specialized usage. As in contemporary French, for Rousseau discours refers to a deliberately formed rhetorical performance adapted to the needs of a chosen goal (see Figure 3, 2.). In this meaning, discours appears in a number of constructions which together may be said to form a semantic frame or script centered around the action which is denoted by the noun in question, such as méditer, écrire, faire, travailler, réciter, prononcer un discours ‘plan, write, make, work out, recite, pronounce a discours’, to cite only a few.

The example in (13) illustrates the particular kind of rhetorical performance Rousseau often has in mind. In (13), discours again refers to the linguistic representation of a discours, i.e., to the representation of the words and sentences spoken.

(13) Supposons, Monseigneur, que quelqu’un vint tenir ce discours aux hommes. “Vous vous tourmentez beaucoup pour chercher des Gouvernements équitables et pour vous donner de bonnes loix. Je vais premièrement vous prouver ... Je vous prouverai, de plus, ... et je
vrais vous montrer ensuite ...” Supposons qu’il expliquât après cela son système et proposât son moyen prétendu.
(R: Lettre à Christophe de Beaumont, Œuvres IV, 941–942)\(^\text{16}\)

The point of departure for this reported *discours* is a question of general concern, namely how to find the proper form of government, and the *discours* consists in a sequence of argumentative “steps”, leading to the speaker’s proposal of an alternative solution. In (13), *discours* denotes a sequence of propositions. Still, in this usage of *discours*, a second interpretation is available, namely that of *discours* as referring to a concrete object, to the written record of a rhetorical performance, (14).

(14)  *Au reste, on a cherché inutilement parmi tous les papiers de feu M' de Bernex le discours qu’il prononça en cette occasion ...*  
(R: Fragments sur Dieu et sur la Révélation, Mémoire, Œuvres IV, 1041–1042)\(^\text{17}\)

Interestingly, as the object of the superordinate clause *on a cherché le discours* ‘one looked for the speech’ the noun phrase *le discours* refers to a concrete object, an interpretation which is imposed by the localizing prepositional phrase *parmi ses papiers* ‘among his papers’, while as the subject of the subsequent relative clause the NP *le discours* resp. the relative pronoun *que* refers to a sequence of propositions. Here, the possible co-occurrence of both interpretations within one sentence is taken as evidence for the assumption that *discours* as an utterance and *discours* as a concrete object do not constitute two separate readings, but rather two possible interpretations within one reading (Cruse 1995). Both interpretations stand in a systematic relation which may be termed “conceptual shift” (Bierwisch 1983), or the two interpretations may be denoted as two “facets” of one lexeme (Cruse 1995).

\(^{16}\) ‘Let us suppose, Monseigneur, that someone came to hold this *discours* to the people. “You trouble yourselves with searching for an equitable government and with giving youselves good laws. I shall first prove to you ... I shall prove further ..., and I shall show you following ...” Suppose that afterward he would explain his system and present his proposed means.’ [my gloss]

\(^{17}\) ‘For the rest, one looked in vain among the papers of the deceased M' de Bernex for the *discours* he had uttered on this occasion ...’ [my gloss]
4.3. Discours as a genre

In the diction of Rousseau, the noun discours refers not only to the written record of a rhetorical performance; it can also be used to denote an instance of a written text of a particular sort, namely a literary genre (see Figure 3, 3.), a usage attested also for Montaigne. This usage is exemplified in (15).

(15) ... les trois principaux de mes écrits, savoir ce premier discours, celui sur l'inégalité, et le traité de l'éducation, lesquels trois ouvrages sont inseparables et forment ensemble un même tout.
(R: Fragments autobiographiques IV, Lettres à Malesherbes, Œuvres I, 1136)

In (15), the noun phrase ce discours is coordinated with a second noun phrase, le traité ‘the treatise’, both subordinated to the noun phrase trois de mes écrits ‘three of my writings’, and superordinated to the phrase lesquels ouvrages ‘the works which’. The concepts referred to by discours and traité stand in a relation of co-hyponymy to ‘written text’, while the concepts referred to by écrit and ouvrage are hyperonyms to discours and traité. By explicitly specifying the network of lexical semantic relations discours pertains to in this usage, in (15) the context restricts the intended interpretation of discours as referring to a written, literary text. In the texts analyzed in this study there are numerous examples of discours as referring to a literary genre. In particular, Rousseau makes extensive use of discours in titles and subtitles, indicating that in his vocabulary, the noun has become a specialized literary term. It is noteworthy that discours, when denoting a literary genre, cannot be interpreted as an action, unlike its reading of ‘spoken utterance’. Therefore, when discours is combined with a possessive determiner, as in ses discours, or a prepositional phrase as in discours de Diderot, the person referred to by the determiner or noun phrase is interpreted not as an agent of an action, but as the author of the text referred to by discours.

18 ‘... my three most important writings, that is to say the first discours, the one on inequality, and the treatise on education, these being three works which are inseparable and which together are a whole.’ [my gloss]
4.4. Discours and spoken language

In some contexts, Rousseau draws on a semantic contrast between discours on one hand, and words referring to written language on the other. In this usage, discours serves to focus upon a specific bundle of properties of spoken language, as exemplified in (16).

(16) Je doute que la même équivoque se trouvât originairement dans la prononciation latine. Car la langue latine étant surtout dans ses commencements beaucoup plus parlée qu’écrite, il n’étoit pas naturel qu’on y laissât dans le discours des équivoques qui ne fussent levées que par l’orthographe.
(R: Mélanges de Littérature et de Morale, Prononciation, Œuvres II, 1249)\(^{19}\)

In this passage, discours has to be interpreted as referring to ‘what is said’, i.e., to an utterance, not to the action of speaking. However, the definite article is used generically, and the generic interpretation of the concept of ‘utterance’ can only be taken to denote general properties of utterances, e.g., to denote the way of speaking or pronunciation (see Figure 3, 4.). A related, but clearly different, usage is found in French of the seventeenth century, where discours could refer to the style of writing, as in the citation given in section 2 above, repeated here: “en écrivant sans cesse variez vos discours” ‘in writing, continuously vary your discours’ (Boileau, PGCF). Both usages, discours as referring to spoken language, and discours denoting a style of writing, focus upon a “quality” of a linguistic utterance: on the one hand, the quality of being spoken, on the other, the quality of having particular stylistic properties.

---

\(^{19}\) ‘I doubt whether the same ambiguity was found originally in Latin pronunciation. For the Latin language being in its origins more spoken than written, it would not have been natural for people to retain in discours ambiguities which could only be resolved by orthography.’ [my gloss]
4.5. A negative connotation

In modern French, when used in the meaning of ‘conversation’, the term *discours* can have a pejorative connotation (see Figure 1, 2.). In the diction of Montaigne, a negative connotation can not yet be detected. On the contrary, for him *discours* carries a consistently positive connotation, referring to the intellectual capacities of man. In the writings of Rousseau, however, numerous examples can be found where *discours* is conjoined with negatively colored words. On one hand, a negative nuance is introduced by certain adjectives, such as *vain* ‘vain’, *superflu* ‘superfluous’, *déshonnête* ‘dishonest’, or *beau* ‘beautiful’. On the other hand, the negative connotation is discernible in constructions where *discours* refers to utterances or texts produced by certain groups of speakers who are qualified by lexically negative adjectives, such as *les gens oisifs*, *les méchants* ‘the idle, wicked people’. How can a term like *discours*, which in its first occurrences (as cited in section 2) has a clearly neutral connotation, and which for Montaigne denotes a rather positively connotated concept, acquire in the diction of Rousseau the negative meaning it still has in contemporary French usage? As a first approximation, one might suspect that in the usage of Rousseau, the negative component is brought in by modifying expressions, i.e., adjectives or noun phrases, but not by the term itself. Hence, it could be said that Rousseau speaks about the referents of the term *discours* by using negatively connotated predicates, without *discours* itself yet displaying a negative connotation, which, metaphorically speaking, only later percolates into the meaning of *discours*. Since this aspect of pejorization is central to the history of *discours*, the emergence of this negative connotation will be further discussed in the conclusion.

4.6. Summary

As shown in the citations analyzed in this section, Rousseau uses *discours* in five more or less distinct senses (see also Figure 3). First, it was shown that Rousseau uses the term *discours* to refer to the action of speaking, but also to the product of this action. Secondly, *discours* is used for reference to a rhetorical performance, and thirdly to a literary genre. Fourth, *discours* denotes spoken language, in contrast to written language. Fifth, when referring to oral linguistic expression, *discours* can have a pejorative connotation.
To conclude, Rousseau’s usage of the term *discours* belongs exclusively to the conceptual domain of linguistic expression or communication. In particular, *discours* is no longer used for reference to the spatial, temporal or mental domains, as it was in the sixteenth century. In this respect, the language of Rousseau shows hardly any difference to contemporary French. An important difference, however, is that in contemporary language *discours* is no longer used as a neutral term for reference to linguistic communication, as in the first usage considered here; only the more specialized usages have been preserved, in particular the reading of ‘rhetorical performance’, and the usage carrying a negative connotation. Like Montaigne, Rousseau appears to have acted, linguistically, as an expert speaker, as a specialist developing a systematic conception of the referential domain of a *discours*, which presumably is later on taken over by the linguistic community. Thus, the usage of *discours* as described in 4.1., i.e., as denoting linguistic action in contrast to non-linguistic action, is closely related to the meaning of *discours* in contemporary socio-cultural terminology, and the usage described in 4.4., i.e., as denoting spoken language in contrast to written language, stands in close relation to the concept of discourse in modern linguistic terminology, as used in the recent coining of the term for the linguistic subdiscipline of discourse analysis.

5. Conclusion

The previous sections first gave an overview of the semantic evolution of the term *discours* from the sixteenth century to contemporary French. Following sections analyzed the usage of this term by two French authors of the sixteenth and eighteenth century. In this analysis, a classification of the usages was developed for each author. Here, the emphasis was on the analysis of the combinatorial properties of the word in question, of its position within a network of lexico-semantic relations, its relation to semantic frames or scripts of complex situations, and its use for reference to linguistic representations of a *discours* in the text, with the aim of developing a methodology for the linguistic analysis of different usages of a word in historical texts.

To characterize its distribution, we considered the term’s combinability with predicates ascribable to concrete objects, like spatial prepositions, predicates ascribable to propositions, or predicates ascribable to events, like aspectual verbs or temporal prepositions. We took into account also the
type of event the word can refer to, as could be inferred from its combi-
ability with agentive verbs or non-agentive verbs. This aspect of the tech-
nique helped to develop a classification of usages, and it lead for each
author to a set of lexical semantic properties of discours that could be de-
scribed separately from the systematic, “expert” meaning discours has in
the authors’ philosophical thoughts.

With regard to the position of the term discours in a network of seman-
tic relations, a number of passages were considered in which discours was
subsumed under a more general term or hyperonym, but also passages in
which it was combined with words pertaining to the same level of abstrac-
tion, standing in a relation of semantic contrast or similarity, with the
paired terms standing in a relation of co-hyponymy, or even near-
synonymy or antonymy. Sometimes, the network of semantic relations of a
word is identified with its lexical meaning. However, it could also be ar-
gued that semantic aspects which can be inferred from a network of rela-
tions of synonymy, antonymy, or hyperonymy relate also to the way the
world is conceived of by the speaker. Similarly, the semantic aspects of
discours inferred from the analysis of definitions of discours found in the
texts and of descriptions of conceptual frames or scripts related to discours,
are dependent in varying degrees upon specific conceptual knowledge. Fi-
nally, in various citations the term discours referred to linguistic representa-
tions of an oral or written utterance. By looking at linguistic representa-
tions of utterances referred to by discours, it was claimed, we can analyze
the properties of the referent, but not always reconstruct the lexical semantic
properties of the referring expression.

To conclude, this investigation has aimed at an analysis of the lexical
semantic properties of the term discours by looking at its combinatorial
restrictions, while trying to separate the conceptual aspects of a word that
belong to the specialist knowledge about the word’s extension, which tradi-
tionally have formed the object of studies on the history of words. It was
shown that the proposed methodology provides a useful means for isolating
central lexical semantic properties, as opposed to meanings given to words
by specialists. The method developed here allowed us to reconstruct a clas-
sification of different usages according to the combinatorial properties
which are specific of different readings of the word. This classification
could serve as the basis for a historical semantic investigation, elaborating
on the specialist’s perspective of the socio-cultural or philosophical impli-
cations of the conception standing behind discours.
Taking the conception of the “specialist speaker” as a point of departure, it could be demonstrated that both authors, Montaigne and Rousseau, function linguistically as experts, in Putnam’s sense, in two respects. First, metaphorically speaking, they select from the set of meanings of discours in common use in the sixteenth and eighteenth century, and develop their own systematic conceptions of the referential domain of the term discours. Second, their criteria for determining the extension of discours appear to be gradually adopted by the French linguistic community. This assumption was inferred from the fact that Montaigne’s and Rousseau’s usage of discours may be said to anticipate the further semantic evolution of this term, as reconstructed from the information contained in a number of French dictionaries. Thus, both authors are innovative rather than conservative. The exact nature of their influence on the evolution of French however awaits further socio-linguistic and philological research.

The general line of development of discours is characterized by a reduction of its denotational domains and an increasing specialization, as a result, we assumed, of an increasing “division of linguistic labor”. In the sixteenth century, discours referred to a variety of conceptual domains, but from the seventeenth century onwards it referred exclusively to the domain of communication. In the seventeenth century it could have a variety of meanings related to communication, i.e., within the domain of communication it showed a considerable polysemy, but in the twentieth century its usage is restricted to specialist terminologies, apart from two specialized readings in standard French. In this respect, the findings of this study are in accordance with Putnam’s more general hypothesis that scientific progress leads to an increase in the division of linguistic labor. Putnam’s hypothesis, however, refers exclusively to the semantics of natural kind terms. The development of discours demonstrates that this assumption may be equally true for the semantics of cultural kind terms.

Still, this study shows that cultural kind terms are in some respects different from natural kind terms. In the case of discours, in the twentieth century its semantic evolution has come to a point where most concepts referred to by the term are no longer relevant for the average speaker. In the case of natural kind terms, in contrast, one would expect that with scientific progress these terms are of increasing relevance to the average speaker, as is the case with words like silicon or asbestos, even though the average speaker has to rely more and more on the expert for exact extensional characterization of these terms. In standard French of Putnam’s average speaker, discours can refer to a rhetorical performance, a concept of mar-
ginal relevance to the average speaker, and it is used more frequently to denote a linguistic utterance with a pejorative connotation. The usages it has in linguistic or philosophical terminology are by and large references to objects which have no well-defined place in the conceptual system of a non-specialist.

It could be said that *discours* displays its most extensive polysemy in the century in which it is borrowed from Latin. For Montaigne, the readings *discours* can have are centered around the concept of reason, while for Rousseau they are more closely grouped around the concept of communication. Thus, in the eighteenth century, *discours* comes to be understood as “tout ce qui part de la faculté de la parole”. This very general usage, however, has not been preserved in contemporary French. Instead, the term has acquired a number of functions in specialist terminologies. Its usage in standard French is restricted to a small number of narrowly defined readings. To conclude, the term *discours*, whose diachronic lexical semantics is investigated in this study, shows an evolution characterized by decreasing polysemy, i.e., by a restriction of the set of the conceptual domains it denotes, such as space, time, mind, and text, and increasing specialization. Its semantic evolution begins in the spatial domain, via the temporal leads to the mental domain, and ends up in culturally defined specialist terminologies relating to the domain of language and communication.

**References**


Busse, Dietrich (ed.)

Croft, William

Cruse, D. Alan

Foucault, Michel

Fritz, Gerd (ed.)

Geeraerts, Dirk

Kohlhaas, Peter

Maingueneau, Dominique

Matoré, Georges

von Polenz, Peter
Judith Meinschaefer

Putnam, Hilary

Starobinski, Jean

Schlieben-Lange, Brigitte

Traugott, Elizabeth Closs

Trier, Jost

Ullmann, Stephen
1957 The principles of semantics. 2nd ed. Glasgow: Jackson.

Wandruszka, Mario

Primary Texts and Dictionaries

Diderot, Denis and Jean LeRond D’Alembert

Dubois, Jean, René Lagane and Alain Lerond

Frantext


Furetière, Antoine
Huguet, Edmond


Imbs, Paul

Journet, René, Jacques Petit, and Guy Robert

de Montaigne, Michel

Richelet, Pierre

Rey, Alain

Robert, Paul, Josette Rey-Debove, and Alain Rey

Rousseau, Jean-Jacques

Tobler, Adolf and Erhard Lommatzsch
Theoretical concepts in flux: Conceptual knowledge and theory change*

Hans Rott

1. Introduction

Take gold, for instance. Is gold a yellow metal? For Kant this was an analytic truth1, while Locke and Leibniz agreed that gold in itself had no color at all – as can be seen, for instance, from the fact that gold changes its color in contact with mercury.2 Putnam said that chemically pure gold was nearly

---

* Earlier versions of this paper were given at the Philosophy Departments of Bochum, Essen, and Hannover, as well as at the Workshop on Belief and Meaning held in Regensburg in May 2000. I want to thank the audiences there, the participants of my advanced seminar in Regensburg, and in particular Regine Eckardt, Wolfram Hinzen, Paul Hoyningen-Huene, and Isaac Levi for helpful comments and discussions.

“For the predicate of an affirmative analytic judgment is already contained in the concept of the subject, of which it cannot be denied without contradiction. ... For this very reason all analytic judgments are a priori even when the concepts are empirical, as, for example, gold is a yellow metal; for to know this I require no experience beyond my concept of gold as a yellow metal: it is, in fact, the very concept, and I need only analyse it, without looking beyond it elsewhere.”

2 In his Essay Concerning Human Understanding (Book 4, Chapter 6, Section 11; ed. P.H. Niddich, Clarendon Press, Oxford 1975: 585–586), Locke writes: “Put a piece of gold ..., separate from the reach and influence of all other bodies, it will immediately lose all its Colour and Weight, and perhaps Malleableness too. ...” Also see Book 2, Chapter 31, Section 6, p.379 where Locke observes that gold even changes its colour “upon a slight touch of Mercury”.

“Philalethes [for Locke]. ’A piece of gold ..., separate from the reach and influence of all other bodies, [would] immediately lose all its [yellow] colour and weight,
white and that its yellow appearance was only due to the presence of copper in the samples of gold we typically see in jewellery.\(^3\) But if gold is yellow, what is the source of our knowledge of this? Is gold a heavy substance which is not consumed by fire, fusible, ductible, malleable, and soluble in Aqua Regia? If so, how do we know? What in fact is gold?

In this essay, I will consider theoretical terms, terms that “come from”\(^4\) the theory in the sense that their correct application or determination of their values is not possible independently of the theory they are part of. In the following, I will use the term “theory“ also, but not exclusively, to refer to theories as produced by an established science; my considerations are also meant to apply to belief systems that might be termed naive or folk theories.

Theoretical terms are the linguistic correlates of theoretical concepts. Theoretical concepts are distinguished from those concepts that are (a) innate, (b) fixed by the subject’s Lebenswelt, or determined on the basis (c) of direct perception, or (d) theories that occupy a lower position in a logico-methodological hierarchy of theories. It has been claimed by proponents of various radically holistic positions that there aren’t any non-theoretical terms in the sense of (a)–(d). Since we are interested in theoretical terms, however, we need not take a stand on this question here.

Theoretical concepts are characterized by the role they play in their respective theories. We can gain a (possibly restricted) understanding of this role by considering what the theory says about these concepts, or more precisely, by considering the sentences of the theory in which the relevant term occurs. I assume in the following that there is a one-to-one correspondence between theoretical concepts (mental entities) and the terms (linguistic entities) that are used in the formulation of the theory.

and perhaps malleableness too’, becoming friable. We know how much the vegetables and animals depend on earth, air and sun; and who knows but that even the most remote fixed stars have some influence on us?

_Theophilus [for Leibniz]._ This is a very good point ..."

It looks as if Locke and Leibniz took as a real possibility what in Kripke’s blue gold example (1980: 118–119) is only conceived as a thought experiment.

\(^3\) Putnam (1975: 250). This is repeated by Stegmüller (1979: 67).

\(^4\) This formulation is due to Putnam (1962b: 219) and made widely known as “Putnam’s challenge” by Stegmüller (1973: 30–34); both authors argued rightly, I think, that theoretical terms cannot simply be identified with “non-observational” terms.
Furthermore I assume that every theory is formulated in a certain language. That is, theories are first of all given as linguistic entities (for instance, as a collection of axioms and theorems, or as ordinary pieces or fragments of text in a textbook). That means, we are primarily confronted with a given concrete formulation of a theory, and the language appertaining to it remains yet to be determined. This seems to be a trivial task, as theorists are always at the same time speakers of some natural language like German, French, Dutch, Portuguese, Japanese, Chinese, and most importantly these days – English. This appearance is deceptive, however, since theorists are also at the same time – though less evidently – speakers of a certain expert language, the language of physics, of chemistry, of sociology, of linguistics, etc. If we may trust the most famous German dramatist and poet, then expert language and colloquial language may be as far apart as any two vernaculars can be:

Mathematicians are like Frenchmen; whatever you say to them they translate into their own language and forthwith it is something entirely different.  

Leaving the national language fixed for the moment, we are thus faced with at least one kind of poorly documented affiliation to a specific language community, one that will moreover be complemented by dialectal and idiolectal variation, as well as more specialized subcommunities of experts and their idioms. The complications posed by the latter variations should not be underestimated, as philosophers will know from personal experience: It may be doubtful whether a Leibniz scholar and a devoted Hegelian will ever really understand each other, and nowadays communication between followers of, say, Wittgenstein and Chomsky proceeds on equally shaky grounds – even though they may be concerned with “the same subjects”.

Let us take it that theories (as concrete linguistic products) are our primary givens, and that languages (as abstract systems of rules) are assigned to theories only in a second step. We cannot tell the language in which a theory is couched by simply looking at the latter; I will, however, assume one more thing: that competent participants in linguistic communication (who are always multilingual, as argued above) are capable of deciding on intuitive grounds whether two given theories are formulated in the same or

---

5 Goethe (1972: 662), maxim 1279.
in different languages. I am going to make use of this assumption in the formulation of my main proposal below; an alternative proposal will be offered that can do without the assumption, but is instead based on the equally problematic notion of “genuinely” or “essentially” different theories.

We have noted above that theoretical concepts are (at least partially) determined by what the theory says about them. Usually, this will comprise much more than only one definition or only one relevant proposition of the theory; it will be a collection of definitions and propositions. Nevertheless, not everything in a theory which uses the terms in question will be relevant for the assessment of their role in the theory. Intuitively speaking, we would like to say that only those propositions of the theory that encode “knowledge of meanings” or “conceptual knowledge” help determining the meaning of the theoretical terms. Meaning is only partially determined in this way of course, because such sentences will not tell us anything about the terms’ reference to objects in the world.

Which are the propositions that express conceptual knowledge? Analytic sentences or judgments, well known from the history of philosophy, seem to be prime candidates. Following Gergonne (1818/19) and Schlick (1918: 30–37), one can also call sentences that partially determine the meanings of (some of) the terms occurring in them “implicit definitions”. This term is connected with one of the more widely discussed approaches to analyticity. Boghossian (1996: 368, 375–387) attributes it to Carnap and the middle Wittgenstein and identifies Conceptual Role Semantics as one of its successor theories. Bealer (1998) closes with a section saying that whatever is interesting or valuable about analyticity is “really about the nature of definitions”, of implicit definitions in particular. We will go on using the label “analytic judgments” in the following.

2. Analytic judgments

Whatever one has to say about analytic judgments, the starting point is clear – clearer than for virtually any other philosophical subject. It is Kant.

The present essay proceeds in a historical order, but it does so with a systematic aim in mind. We will touch upon three central figures: Kant (1781–90; 1800), Frege (1884) and Quine (1951, 1974, with further developments by Putnam 1962–1979). We begin with Kant, who took synthetic judgments a priori, judgments claimed to be non-existent by the empiri-
cists, as the pivot of his epistemology and metaphysics; we move on to Frege’s logicist program; and we end with Quine, a dominating figure in analytic philosophy up to the present date, who denied any sensible distinction between analytic and synthetic judgments and thus removed the basis for both the empiricists’ and Kant’s position. It will strike the eye that the philosophy of mathematics takes an important part in our guided tour.

2.1. Kant

Kant’s critical philosophy lives crucially on two contrasting dichotomies: the one between the a priori and the a posteriori, and the one between the analytic and the synthetic. While the former is an epistemological dichotomy (and consequently, the predicates “a priori” and “a posteriori” apply to bits of knowledge), the latter distinction pertains to the philosophy of language or to semantics (and, correspondingly, “analytic” and “synthetic” are predicates that apply to judgments formulated in some language). As the two dichotomies do not coincide, the synthetic a priori is conceivable; Kant’s metaphysics explores the ways the synthetic a priori is real.

Let us briefly recapitulate the terminology. Knowledge a priori is knowledge which is independent of (and in some way “precedes”) empirical experience. This can in turn be spelled out in at least three different ways. First of all along the specifically Kantian lines, according to which this independence can be demonstrated by transcendental arguments (concerning the conditions of the possibility of experience). We will not be concerned with this idea, however, nor with Kant’s criteria according to which knowledge a priori is both necessary and strictly universal. The dependency of knowledge on empirical experience can be spelt out in two different ways: as the capability of being verified, confirmed, or supported (“positively affected”) by empirical evidence, or as the capability of being falsified, weakened, or undermined by empirical evidence (“negatively affected”). I refer to these respectively as the positive and the negative interpretation of dependence on empirical evidence. Correspondingly, knowledge a priori will be knowledge that – the positive interpretation – cannot be supported by any kind of empirical evidence, or that – the negative interpretation – cannot be undermined by any kind of empirical evidence.

6 In Kant, the metaphysical term “necessary” is still linked to epistemology, and it was only much later (Kripke) that this link was called in question.
These two latter interpretations of knowledge a priori are not to be found in the writings of Kant, but we will come across them again in the explication of the analytic (not of the a priori!) in Frege and Quine. Knowledge a posteriori (also known as empirical knowledge) is knowledge that is not a priori.

Now consider the second of the two basic Kantian dichotomies. Kant calls a judgment analytic if the predicate term is “contained in” the subject term. Due to this containment, analytic judgments do not give us any new information; according to Kant they are only elucidating or analyzing concepts (zergliedert), they are not ampliative (erweitert) with respect to our knowledge. The metaphorical phrase “contained in” is in need of further interpretation. As we shall presently see, Kant intends it to express that the predicate concept is “actually thought” in thinking the subject concept. As a criterion for recognizing analytic judgments, Kant tells us that they are valid according to the law of identity or contradiction (which Kant takes to be one and the same law).

---

7 A few times Kant actually speaks of the confirmation of knowledge a priori, for instance in the Prolegomena, §16: “we are here ... concerned ... only with cognition of nature, the actuality of which can be confirmed by experience, though it (the cognition of nature) is possible a priori and precedes all experience.” or §42: “All pure cognitions of the understanding have this feature, that their concepts present themselves in experience, and their principles can be confirmed by it.” See also the Preface to the second edition of the Critique of Pure Reason, footnote to BXVIII (transl. Paul Guyer and Allen W. Wood, Cambridge UP 1997): “This method, imitated from the method of those who study nature, thus consists in this: to seek the elements of pure reason in that which admits of being confirmed or refuted through an experiment.” All this is surprising at first sight; my interpretation is that confirmation of the Kantian a priori lies in the possibility of experience at large, rather than in the local confirmation of particular principles by particular pieces of evidence.

8 Kant’s remarks about the informational content of analytic judgments are not without ambiguities. On the one hand he fervently denies that such judgments can ever convey any new information, on the other hand there are passages where he admits that “analyses of the concepts” (i.e. analytic judgments) may well “afford us a multitude of cognitions” or “yield a real a priori cognition, which makes secure and useful progress” (Critique of Pure Reason, B9–10).

9 See Critique of Pure Reason B193f, B622, Prolegomena, §3, and Logik Jäsche, A XV, A75.
Theoretical concepts in flux 181

As Kant made explicit in the *Prolegomena* (more so than in the *Critique of Pure Reason*), the starting point of his metaphysics and its core question “How are synthetic judgments a priori possible?” is the claim that mathematical propositions are perfect examples of synthetic judgments a priori. Therefore, so Kant, such judgments do exist. But how then do we know that mathematics, i.e., arithmetic and geometry, are synthetic? Consider the following passage which, to my mind, lays down Kant’s clearest idea of how to define the notions of analyticity and syntheticity.

What usually makes us believe [in mathematics\(^{10}\), HR] that the predicate of such apodeictic judgments is already contained in our concept, and that the judgment is therefore analytical, is the duplicity of the expression, *requesting us* to think a certain predicate as of necessity implied in the thought of a given concept, which necessity attaches to the concept. But the question is not what we are *requested to join in thought* to the given concept, but what we *actually think* together with and in it, though obscurely; and so it appears that the predicate belongs to these concepts necessarily indeed, yet not directly but indirectly by an added visualisation (*Anschauung*).\(^{11}\)

The phrases in italics seem to suggest that Kant wants to draw a distinction between “what one should think” in a normative sense and “what one actually thinks” in an empirical, psychological sense. If this is true – and the literal reading says it is\(^{12}\) – then this reference to “actual thinking” links up analyticity to subjective or psychological processes.

\(^{10}\) Paul Hoyningen-Huene (1998) argues that this passage does not refer to mathematics in general but only to arithmetic, and that it should accordingly be shifted in the text of the *Prolegomena* (and in the second edition of the first *Critique*). According to Hoyningen-Huene’s interpretation of Kant (personal communication), no one would ever be tempted to think that geometry is analytic.

\(^{11}\) *Prolegomena* (1783; §2c), emphasis in the German original but not in Logan’s 1996 edition of the Carus translation. This passage is almost literally repeated in the second edition of the *Critique of Pure Reason*, B17, with the following differences: “directly” is replaced by “as thought in the concept itself” (and “added” is replaced by “added to the concept”).

\(^{12}\) Similarly in *Critique*, B205, B746, B749 and *Prolegomena* §2a. These repeated passages show that Kant posits the link between *wirklich denken* and analyticity as present in all consciousness. Our background knowledge about Kant’s idea of logic and philosophy suggests that he could not have thought of anything psychological. It remains an open question, however, what *wirklich denken* is supposed to mean,
If we furthermore recall that all thinking is expressed and perhaps also
guided by judgments formulated in the language of some language com-
ity, then we might feel tempted to say that Kant’s notion of analyticity
is even rooted in part in social processes. This, however, clearly exceeds
what can be found in Kant’s own writings.\footnote{This hint at language-
dependency also suggests that we consider not particular,
real speakers but an “ideal” speaker of some language. We would then be con-
cerned with linguistic competence, not just with performance. Yet the question is:
Competence with respect to what language? And furthermore: Which competences
must or may competent speakers possess without at the same time being pro-
ponents of a certain theory about the world?}

2.2. Frege

In Frege’s understanding of the analytic and the a priori, the philosophy of
mathematics plays an even more crucial role than in Kant. Frege’s position,
however, differs significantly from Kant’s in that he takes both dichotomies
as purely epistemological distinctions. In the Foundations of Arithmetic,
Frege offers definitions that refer solely to types of proof and premises
(primitive truths, Urwahrheiten). An analytic truth, according to Frege, is a
proposition the derivation of which requires only “general logical laws”
and “definitions”; the derivation of a synthetic truth in contrast requires
propositions that are “not of a general logical nature but belong to the
sphere of some special science”. The derivation of an a priori truth only
requires “general laws which themselves neither need nor admit of proof”;
to obtain an a posteriori truth one needs “facts”, i.e., “truths which cannot
be proved and are not general, since they contain assertions about particular
objects”. (Frege 1884, §3, all transl. from Frege 1959, 4\textsuperscript{e})

Note the similarity between Frege’s definition of the analytic/synthetic
dichotomy and the “positive interpretation” of the Kantian a priori/a poste-
riori dichotomy that I proposed above. But in Frege’s eyes, at least one of

\footnote{But in Frege’s eyes, at least one of the definitions of an analytic truth and a
synthetic truth is somehow “psychological” in nature. This is because Frege
himself maintains that the distinction between analytic and synthetic
propositions is based on the fact that the former can be derived from
premises (primitive truths) while the latter requires the introduction of
new terms. However, if we interpret this distinction as purely
epistemological, then we can say that Frege’s definitions are indeed
“psychological” in nature.}

\footnote{This hint at language-dependency also suggests that we consider not particular,
real speakers but an “ideal” speaker of some language. We would then be con-
cerned with linguistic competence, not just with performance. Yet the question is:
Competence with respect to what language? And furthermore: Which competences
must or may competent speakers possess without at the same time being pro-
ponents of a certain theory about the world?}
Kant’s characteristics of analytic judgments is completely off the point: The logical consequences of analytic truths “extend our knowledge and ought therefore, on Kant’s view, to be regarded as synthetic; and yet they can be proved by purely logical means, and are thus analytic” (Frege 1884, §88, transl. Frege 1959, 104e). Here Frege has evidently more in mind than merely a terminological shift. In the light of the interpretation of Kant I sketched above, we can take Frege’s deviation as signaling that he denied (with good reason I think) the Kantian claim that analytic consequences are always actually thought together with the concepts in question.

Frege’s terminology is in the first instance restricted to mathematical judgments, and it is not clear whether or how it might be transferred to non-mathematical judgments. But tensions arise even within mathematics proper. Why, one might ask, does Frege follow Kant in assuming that geometry is synthetic but refuses adherence when it comes to arithmetic? Why does Frege hold arithmetic to be analytic? Here is an important argument:

For purposes of conceptual thought we can always assume the contrary of some one or other of the geometrical axioms, without involving ourselves in any self-contradictions when we proceed to our deductions, despite the conflict between our assumptions and our in-

---

14 Frege agrees with Kant that geometry is not rooted in conceptual reasoning but in intuition – at least he does so in The Foundations of Arithmetic. In the posthumous manuscript “Logic in mathematics”, however, written 30 years after the publication of Foundations, Frege says that in his (personal!) opinion the axiom of parallels is valid due to the meanings of the words “straight line”, “parallel”, and “intersect”:

Can the axiom of parallels be acknowledged as an axiom in this [the traditional, HR] sense? When a straight line intersects one of two parallel lines, does it always intersect the other? This question, strictly speaking, is one that each person can only answer for himself. I can only say: so long as I understand the words ‘straight line’, ‘parallel’ and ‘intersect’ as I do, I cannot but accept the parallels axiom. If someone else does not accept it, I can only assume that he understands these words differently. Their sense is indissolubly bound up with the axiom of parallels. (Frege 1914, transl. Frege 1979, p. 247).

For an assessment of this passage in the context of Frege’s philosophy see Rott (2000a).
tuition. The fact that this is possible shows that the axioms of geometry are independent of one another and of the primitive laws of logic, and consequently are synthetic. Can the same be said of the fundamental propositions of the science of number? Here, we have only to try denying any one of them, and complete confusion ensues. Even to think at all seems no longer possible. The basis of arithmetic lies deeper, it seems, than that of any of the empirical sciences, and even than that of geometry. The truths of arithmetic govern all that is numerable. This is the widest domain of all; for to it belongs not only the actual, not only the intuitable, but everything thinkable. Should not the laws of number, then, be connected very intimately with the laws of thought? (Frege 1884, end of §14, transl. Frege 1959, p. 20)

Hence, propositions are analytic if one cannot deny them without violating the very foundations of thinking, or to put it less emphatically, if their negation would lead to logical contradictions (this is quite precisely Kant’s criterion). Conceptual thinking – to which logic and arithmetic belong – is more fundamental for Frege than (even the pure forms of) intuition. Logical theorems, as well as the definitions with the help of which Frege intended to build up arithmetic from logic, enjoy this fundamental status.

Of course, the question arises how we can delimit conceptual thinking in general, and thus, how we can designate a particular logic and particular definitions of arithmetical terms as “the right ones”. The above quotation suggests that we approach the task by saying that the distinguished parts cannot be abandoned without undermining the very possibility of thinking. But haven’t even fundamental terms like “not” or “number” in fact been subject to change? Aren’t there in fact alternative, deviant axiom systems for logic and arithmetic on the basis of which thinking can still advance? Consider, for instance, the development of intuitionistic logic where the inference from $\neg\neg A$ to $A$ is forbidden, or the glorious career of “imaginary” numbers that, multiplied by themselves, yield a negative number.\footnote{Such numbers were conceived of as early as in the 16th century (G. Cardano 1545, R. Bombelli 1572); the term “imaginary” was coined by Descartes (La Géométrie, 1637, Book III). See Nahin (1998).}

One might continue Frege’s line of thought by arguing that what is accepted as a reasonable system of logic, arithmetic or geometry is determined by individual decisions or practices of everyday and scientific language use. There actually are communities who prefer intuitionistic to classical logic, and it has by now become a matter of course to use techniques
Theoretical concepts in flux

Involving imaginary and complex numbers – which by themselves are very strange creatures indeed. Developments in mathematical physics, such as relativity theory, quite evidently have an effect on which geometries we are ready to accept or reject.\textsuperscript{16} The history of science offers a wealth of examples illustrating that things which had earlier been considered as intuitively or conceptually impossible may well turn out to be conceivable, amenable to investigation and even useful. We must not close our eyes to our dependence on the current state of individually or socially attained knowledge, a dependence that precludes an everlasting, “objective” fixation of the extension of the analytic.

2.3. Quine

Quine’s background is Carnap, a student of Frege’s. Carnap’s (1950) “external questions” concern the choice of a language or conceptual framework; here it is appropriate to exercise a lot of tolerance.\textsuperscript{17} Only afterwards, in a second step, follow “internal” empirical questions that deal with the factual truth or falsity of sentences within the chosen conceptual framework; no pragmatic freedom of choice is admitted here. For Carnap, the conceptual framework determines clearly which judgments are analytic: the judgments that can be derived from meaning postulates of the conceptual framework. Thus there is a definite demarcation between language (step 1) and theory (step 2), a demarcation that became particularly clear in the light of Carnap’s formalization program.

Quine, for whom the philosophy of science is inseparable from the philosophy of language, objects to Carnap’s position on at least two points.

\textsuperscript{16} This is not to say that mathematics and physics share the same subject. Recall that Einstein (1921) distinguished purely axiomatic from practical geometry, Carnap (1922) formal from physical space and Reichenbach (1928) mathematical from physical geometry.

\textsuperscript{17} Compare Carnap’s (1937: 51) \textit{Principle of tolerance}: “It is not our business to set up prohibitions, but to arrive at conventions.” The concept of a \textit{framework} is central in Carnap (1950), where “framework” is used without any adjective; later (in the reprint in the second edition of \textit{Meaning and Necessity}), Carnap also uses the term “linguistic framework”. Alternative terms of Carnap’s are “linguistic form”, “form of language”, or “system”. The use of the term “conceptual framework” in this context is mine. According to Quine, choosing a framework is choosing an ontology, but Carnap disapproved of this way of expressing things.
Firstly, he denies that folk or scientific theories are ever set up in the described two-stage process, and secondly – a point frequently overlooked – he denies that Carnap’s formalization program was fruitful or even relevant for the project of isolating knowledge of meanings from knowledge of facts.

Let us have a look at Quine’s classical 1951 paper Two dogmas of empiricism. In the first part of the paper, Quine discusses several definitions of analyticity that are based on notions like “definition”, “interchangeability”, “synonymy”, “semantical rule”, or “postulate”. We find for instance a characterization coming very close to Frege’s notion, according to which an analytic judgment is one that can be turned into a logical truth by replacing some words by synonymous words. Quine argues that this attempt does not explain anything because the notion of “synonymy” is just as loaded with problems as the notion of “analyticity”. The same, he claims, is true for all of the above-mentioned terms. After these rather destructive arguments, however, Quine in the second half of Two dogmas offers a constructive proposal of how to make sense of the term “analytic”. According to this proposal, analytic judgments are judgments with an empty “factual component”, i.e., ones which are “vacuously confirmed, ipso facto, come what may”, and are thus confirmed in precisely the same way as logical truths. This “positive reading” requires too much, because the existence of black ravens, for example, should not count as confirming that all bachelors are unmarried. We should be content with the “negative reading” that analytic judgments do not get undermined “come what may”. 18

Thus the concept of analyticity finds a well-defined interpretation even for Quine, but he considers it as a pointless concept that can never actually be applied in (scientific or folk-theoretical) practice:

... it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements, which hold come what may. Any statements can be held true come what may, if we make drastic enough adjustments elsewhere in the system. ... Conversely, by the same token, no statement is immune to revision. ... (Quine 1951, p. 37)

18 Observe the similarity between Quine’s definition of analyticity and the “negative interpretation” of the a priori that was proposed above.
This passage makes clear that Quine does not deny that there is a coherent concept of analyticity; he only insists that – given the explication proposed – there are no sentences qualifying as analytic.

In the final section of *Two dogmas*, Quine outlines a unified process of language and theory formation which simply does not exhibit the two different stages postulated by Carnap.\(^1^9\) There are rather voluntary aspects all the way down to the determination of which theory elements to adopt, which to maintain, and which to jettison in face of recalcitrant experience. Hence, decisions about revisions *would* define which judgments to class as analytic and which as synthetic, *if* there were principled criteria determining which revision strategies are transsubjectively valid and binding within a given language community. But this is not the case, according to Quine’s picture.\(^2^0\)

Why does Quine refuse to consent to a solution along the lines of his teacher and friend Carnap? Like Carnap, Quine excelled in giving formal-logical, “rational” reconstructions and reductions of philosophical problems. It is important to see – and has in my opinion received much too little attention – that Quine resolutely rejected this kind of solution as early as when he wrote *Two dogmas*:

> Appeal to hypothetical languages of an artificially simple kind could conceivably be useful in clarifying analyticity, if the mental or behavioral or cultural factors relevant to analyticity – whatever they may be – were somehow sketched into the simplified model. But a model which takes analyticity merely as an irreducible character is unlikely to throw light in the problem of explicating analyticity. (Quine 1951, p. 34)

Attempts at formalization thus are irrelevant for the elucidation of analyticity, and this is due to the fact that some essential factors of language use are

\(^{19}\) In the year in which *Two dogmas* appeared, shortly before his death, Wittgenstein (1969, §§ 318–319) noted that there is “no sharp boundary between methodological propositions within a method” (between “rules” including the “propositions of logic” and “empirical propositions”). Obviously, this is similar to Quine.

\(^{20}\) At this point an interesting question arises. If we understand by a *theory community* a subgroup of a linguistic community with a common theory about the world, might there be binding revision strategies at least for a given *theory community*? My suggestion below assumes that there is a positive answer to this question (which saves it from Quinean pointlessness).
just filtered out in formalizations. Formal, artificially constructed languages work in a way different from natural, living languages. The former may serve to model the latter, but they will at best approximate natural languages in certain aspects. Quine seems to suggest in the above passage that development and use, in their artificial languages are actually of a completely different kind than natural languages. Insofar as the analytic/synthetic distinction depends on an essential similarity with natural languages, the modelling provided by formal languages is completely inadequate. This view was of course advanced against the background of the state of philosophy in the middle of the 20th century. It is an empirical question, and the possibility cannot be excluded on a priori grounds, that some day we shall find proper means for the formal representation of “the mental or behavioral or cultural factors” and that Carnap’s program can be carried out after all. Yet I guess that the situation has not much changed since 1950.\(^{21}\)

It seems to me that Quine’s picture is by and large correct. I will not follow him, however, in his sceptical conclusion that in view of his diagnosis, we had better give up the fruitless notion of analyticity altogether. Hilary Putnam has made many important contributions to the extension, but also to the criticism, of Quine’s position. I can only discuss one of these here.\(^{22}\)

\(^{21}\)The last quotation may also help furnishing an answer to the Kantian issues as to whether the competent speakers of a natural language really think the predicate when thinking the subject concept of an analytic judgment, and whether denying such a judgment would lead to logical contradictions. Both questions could perhaps be answered in the affirmative if indeed the respective concepts or terms of English, say, were somehow, somewhere officially and bindingly related to each other (as through the axioms or definitions of a formal system). As this is not the case in ordinary language, I think that the Kantian criteria are not applicable.

\(^{22}\)Putnam’s most influential contribution to the discussion of analyticity can perhaps be found in “Meaning of meaning” and/or “Holism” (with the well-known dictum “Meanings must be invariant under belief fixation”). According to Putnam, the synonymy of two terms does not guarantee analytic truth, and not even truth simpliciter. To provide further samples of Putnam’s creative thinking: In “The analytic and the synthetic” (1962) Putnam makes a distinction between planned revisions and spontaneous ad hoc revisions, defines analyticity relative to the set of alternative theories available (an adaptation of Kuhnian ideas; revisions are never occasioned by experience alone, but are also influenced by the theoretical situation), and finally introduces the notion of a “law cluster concept”. In “Analyticity and apriority: Beyond Wittgenstein and Quine” (1979), he offers arguments
Putnam (1976) says – like others before him – that Quine’s critique was really directed against apriority rather than analyticity. Is that true? At the present stage of discussion, we can say that a judgment is a priori if it cannot be revised, which is what we called the negative notion of “independent of experience” above. A judgment is called analytic if it cannot be revised – except by a local or global change of the language, i.e., a change of the meaning of single terms or the language system as a whole. Analyticity is always relative to a given language. A judgment, on the other hand, can only be a priori if the terms themselves that occur in it and are thus applied to experiences are in some objective way “valid a priori”, or at least absolutely cognitively privileged – whatever this may mean. Without any such presupposition, doubtful as it seems to me, a notion of apriority that is not relativized to a given language is idle.

At this point we have severely reversed the relations between the core concepts as originally given by Kant. In the sense just explained, every judgment a priori is analytic (this is even true, according to the rules of reasoning in classical predicate logic, if there are no a priori judgments at all). Conversely, not every (L-)analytic judgment is a priori, because there is always the possibility of a tacit or explicit change of language (to L’).

2.4. Conclusions from the historical review

From the discussion of three classical accounts of analytic judgments supports, I draw the following conclusions which are largely, but not entirely in agreement with Quine.

(1) The revisability of a sentence (qua sentence of a certain theory) offers the best – “negative” – criterion for its analyticity (and also apriority).

(against Wittgenstein) to the effect that no part of one’s beliefs can ever be completely immunized against revision by one’s communal practices.

23 If they are, for instance, part of Fodor’s “language of thought” or express “elite properties” in the sense of D. Lewis.

24 This result is consonant with Carnap’s (1937: 318) statement about the revisability of logical “L-rules” (“Not every analytic judgment is unrevisable.”) and with Kitcher’s (1981: 223–226) conclusion on the basis of his distinction between “weak” and “strong” revisions (“Not every analytic judgment is apriori.”).
(2) The distinction between the two dichotomies analytic/synthetic and a priori/a posteriori has been blurred in much of the contemporary discussion within analytic philosophy. This tendency was recognizable, at the latest, in Frege’s epistemological reinterpretation of the notion of analyticity, and it was continued in Quine’s discussions who, although concerned with epistemological issues, never dwells on the notion of the a priori. Presumably, the blurring of the distinctions is not (only) the result of a somewhat lax way of dealing with our philosophical heritage but, as I will presently argue, it is in part motivated by internal reasons. What will happen to Kant’s cross classification, what to his basic metaphysical question: “How are synthetic judgments a priori possible?” If the reasons I mentioned are convincing, the Kantian project loses its foundation.

In the following, I want to accept Quine’s diagnosis of the problems of demarcating the analytic and the synthetic, as well as the pragmatic solution that he sketched.25

(3) We found reason to interpret the writings of all three protagonists in a way that discloses references to psychological processes or to social practices, even though the authors did not always explicitly mention them.

25 Quine was not at all the first to argue that the analytic/synthetic distinction is pointless. Eisler (1904) reports in his entry Urteil, analytisches: “According to J. G. Fichte, there is ‘no judgment that is purely analytic in its contents.’ (Grundlage der gesamten Wissenschaftslehre, second edition, 1802: 33). ... G. E. Schulze emphasizes: ‘For one person is an analytic judgment what for another one constitutes a synthetic judgment’ (Über die menschliche Erkenntnis, 1832: 196). According to Schleiermacher, too, the difference between analytic and synthetic judgments is vague (Dialektik, ed. Jonas, 1839: 264, 563). ... According to Trendelenburg, every judgment is at the same time analytic and synthetic (Logische Untersuchungen, 2 Vols., 1862, II: 241 ff), an opinion shared by Jodl (Lehrbuch der Psychologie, 1896: 616).” The original contribution of Quine thus consists only in his specific way of arguing, not in the claim itself. – While the analytic/synthetic distinction seems dubious, there is no need to challenge the value of analytic judgments themselves. No need thus to object to Frege’s view (1884, §§17, 88, 91) that analytic judgments need not necessarily be uninformative; or to Ayer’s (1946: 79–80) remark that analytic judgments may “reveal unsuspected implications in our assertions and beliefs”, as well as of linguistic usage in general.
As we have seen, Kant was ultimately after the “actual thinking” of the person who endorses a judgment. This at least admits of a psychological interpretation. But a thinking and judging person is part of a socially defined language community. So we may, qua members of a community of speakers, researchers, or communicators, reasonably ask ourselves in what case we would think of a person as an incompetent speaker (lack of terms); in which case we would think of her power of imagination as too poor (lack of intuitions); and in which case we would think of her as being simply ignorant (lack of world knowledge).

Frege discusses languages and theories in a mathematical context. Here, questions like the following will arise: When do we, qua members of such communities of speakers and theorists, accept a given system of arithmetic or geometry as one that correctly represents our “natural” numbers or our “natural” space? To what extent are such issues dependent on the decisions of individuals, to what extent are they settled by social agreements or conventions?

Quine, finally, makes explicit mention of cultural factors when rejecting purely formal reconstructions à la Carnap. What enters into the “definition” of a term’s meaning in a natural language cannot be fixed through philosophical analysis, but is determined only by the concrete social practices of speakers (and always remains preliminary, open to revisions). According to Quine, Carnap’s voluntarism with respect to the choice of a conceptual framework should be extended to all stages of language acquisition and theory formation. The question that dislodges the traditional problem of analyticity is this: Which parts of a network of theories do we choose to sacrifice when we face “recalcitrant experiences.” Quine elaborates on Duhemian confirmation holism here.

The core question that I want to address now with my own proposal, is the following: When do we – qua competent participants of a linguistic or theoretical practice – speak of a different theory in “a new language”, when do we speak of a “really new” theory formulated in the old language and when do we only speak of a new version of an old theory that has essentially remained the same? To put it more generally: What is the difference between a change in language and a change in beliefs, what the difference between a change in the dictionary and a change in the encyclopaedia? And, finally, are all these issues material questions, or might they simply address a matter of terminology?
3. A new proposal

We start by giving some definitions.

(1) A revision of a theory $T$ in a given language $L$ is called small (or evolutionary or conservative) if the resulting theory $T'$ is still understood as a theory in the language $L$. The revision is called big (or perhaps revolutionary) if $T'$ is not understood as a theory in language $L$ any longer.

The main idea is that in the case of a big revision – and only in such a case – theory change will lead to meaning change.

(2) Analytic sentences of a theory $T$ in the language $L$ are those sentences that are retained across all potential evolutionary revisions of $T$.

Conversely, but equivalently:

(2') He who – for whatever reason – gives up an analytic sentence (of theory $T$ in language $L$), steps out of the evolution of theory $T$ in favour of a theory $T'$ in a new language $L'$.

An alternative proposal is this:

(3) He who – for whatever reason – gives up an analytic sentence (of theory $T$ in language $L$), steps out of the evolution of theory $T$ in favour of a theory $T'$ that is a theory really or essentially different from $T$ (i.e., $T'$ can no longer be regarded as a revised version of theory $T$, but constitutes a genuine break or fresh start).

The variant in (3) avoids the problematic task (discussed at the beginning of the paper) of assigning a language to a theory. It does so, though, at the expense of the equally problematic task of telling “real” or “essential” differences between the old and the new theory. I am not sure which variant is to be preferred. The aim in any case is to systematize correctly the intuitive judgments of competent speakers and theorists. There is an obvious discrepancy between the two proposals that looks unacceptable at first glance – shouldn’t there be a big difference between language change and theory
change? – but I want to postpone this point and address it only in the next section.

Both variants are motivated by the common idea that theory change leads to meaning change in exactly those cases where very central (Quine’s metaphor) or high-level (another metaphor) principles of the original theory are sacrificed, principles that were in some sense meaning constitutive for the terms that occur in them.

Let us recapitulate the course we have taken with this proposal. We said that conceptual knowledge or, more precisely, knowledge of the meanings of theoretical terms, is dependent on a certain class of propositions of the theory within which the terms in question are used. These propositions are the analytic judgments of the theory. They in turn are characterized by the fact that abandoning them leads to a theory in a new language (or alternatively, to an essentially different theory). The question whether such a difference in language (or such an essential difference in theory) is indeed present is to be determined on the basis of intuitions of the participants of the relevant linguistic (or theoretical) practice, intuitions that relate to large units – whole languages or whole theories – rather than to individual propositions.

If this analysis is right, what do we need to flesh it out?

1. We need a fully specified, powerful model of how to revise theories – for whatever reason. A rich spectrum of such accounts is already available and can be taken from the fields of philosophical logic and knowledge representation. In particular, the models used there offer various ways of making sense of the “centrality” or “high rank” of sentences in a given theory. This concerns a technical issue that we need not be concerned with at present.26

---

26 For the theory of theory change and belief revision, see Alchourrón and Gärdenfors, and Makinson (1985), Gärdenfors (1988), Gärdenfors and Rott (1995) and Hansson (1999). For a criticism of the philosophical self-understanding of belief-revision theories, and for their relation to Quine (1951), see Rott (2000b). In order to model the change of scientific theories, one moreover needs to account for “meta-revisions”, i.e., revisions of strategies for theory revision, or revisions of priorities, of degrees of theoretical importance, of epistemic entrenchments, or the like.
2. We need an account of what it means “to speak a different language” or “to hold an essentially different theory”. This certainly is anything but a technical issue, and indeed I think that this issue cannot and should not be reduced to anything else. Answers can in my opinion only be provided by competent speakers, writers, or theoreticians themselves.

Let me finally point out some features that are characteristic for the definition of analyticity that I proposed above:

(a) Since language change or an essential theory revision is always possible, there are no sentences that are analytic simpliciter (and that we have termed “a priori” above). There aren’t even any sentences that are \(L\)-analytic simpliciter, because analyticity is always relative to the currently accepted theory \(T\).

(b) According to the above explication, meaning is holistic insofar as the ascription of meaning proceeds from the theory or the language in which the theory is phrased to the terms. Here, “theory” may well denote our all-encompassing theory about the world, especially if \(L\) is a natural language.

(c) Speakers’ decisions as to when two languages or theories are identical are subject to vagueness and perspectivity and dependence on speakers’ interests. All this will carry over to the notion “analytic in theory \(T\) in language \(L\)”. I do not consider this a disadvantage of the explication but take it simply to reflect facts that will have to be taken care of by any attempt at explicating analyticity.

(d) The problems of vagueness, perspectivity and interest-dependence become less severe if we move from a categorical to a comparative notion of analyticity. In order to make this idea precise, we can make use of so-called “entrenchments” or “priorities” commonly used in belief revision theories which were originally meant to reflect comparative “degrees of retractability” but can be reinterpreted as “degrees of meaning constitutivity”.\(^{27}\)

\(^{27}\) We must, however, not assume that a statement’s retractability and its power to determine meanings always go hand in hand. Observation sentences, for instance,
In judgments about the (essential) identity of languages and theories, more emphasis will be given to either individual, personal aspects, or to social, conventionalized aspects, depending on whether we are interested in the momentary idiolect of a single individual at a particular point of time, or rather in the stable language spoken by a community. A judicious choice of factors will get us identity judgments of the desired granularity.

All definitions have to be taken with a grain of salt, as they all presuppose that the terms in question are of sufficient importance for the identity of the whole language or theory. The well-known bachelor sentences can certainly be classed as analytic or constitutive for the meaning of the term “bachelor”, yet they do not seem to be essential for the identities of either the languages or the all-including folk theories of the speakers of English. The term “bachelor” is simply too marginal.

According to the present proposal, meaning change does not imply reference change. As was already noted by Locke for the case of gold, theories with differing conceptual structures may well speak about the same objects, substances or natural kinds.28

The above proposal is not sensitive to the choice of the philosophical modeling of scientific change: Models that assume that revisions can be forced by experience alone fit the proposal, as do models assuming that revisions can only be accepted if a superior alternative theory is available.

---

28 In this respect there are no differences between my proposal and externalist theories of meaning in the spirit of Kripke and Putnam. Compare especially Putnam’s (1975: 250–257) remarks about stereotypes which are (in his view) empirically wrong, like the stereotype that gold is yellow.
Scientific revolutions in which an old theory is given up altogether in favor of an entirely new theory $T$ in an entirely different language $L'$ can be seen as the extreme case of a “big” revision in the sense of definitions (2) or (3) above. Kuhnian ideas about the development of scientific disciplines can thus be integrated, and perhaps even nicely reconstructed within the approach.\textsuperscript{29}

4. Theory change vs. meaning change

Semantic holism says that the meaning of words (as well as the truth of propositions) is determined by the environment in which they occur, especially by the linguistic environment. That is, one and the same word or sentence might, or must, mean different things in different linguistic contexts. If the theory changes, then the meaning of its terms changes as well. This position has often been criticized, since intuitively it is easily possible to change a theory (or simply, one’s opinions or beliefs) without any variations of meanings. It is even necessary in a certain sense that meanings can remain constant when theories or beliefs change. For if linguistic expressions continually changed their meanings, how could we ever come to know whether the content of a theory (a belief) has changed? Phrases looking exactly alike would mean different things, and contradictions on the linguistic surface could be explained away by appropriately reinterpreting the words used.\textsuperscript{30}

So a modicum of constancy in meanings is a prerequisite for an understanding and communication. This is a very general argument against a holism that has overshot the mark. The generality of the argument ensures its broad range, but is also its weakness. Let us therefore, in order to arm us

\textsuperscript{29} Rott (1994) is an attempt at a reconstruction of Lakatos’ (1970) model of scientific theory change which can be considered as a rationalization of ideas of both Popper and Kuhn. This reconstruction in terms of belief-change theory (see footnote 26) is, however, restricted to evolutionary theory developments and neglects the problem of meaning change altogether.

\textsuperscript{30} In reaction to Kuhn’s \textit{Structure of Scientific Revolutions} it was critically observed that after a revolution, the new theory cannot simultaneously be inconsistent and incommensurable with the old theory (Achinstein 1964, Shapere 1966). Both at the same time are impossible indeed: Inconsistency requires \textit{constant} meanings, while incommensurability involves \textit{changes} of meanings.
against philosophical sophistries, discuss two everyday cases to illustrate the intuitive differences between theory change and meaning change.

Example 1: If a leading politician today calls Perugona an unreliable trade partner while she claimed the opposite some days ago, then we will normally assume that she, on the basis of whatever evidence, has changed her beliefs. It is however conceivable that her current statement can be explained by the fact that the standards for economic reliability have been changed, tacitly or officially.

Example 2: If today I find that my computer at home is slow, but I considered it a fast computer some months ago, then this will probably be due to the fact that what was considered fast when the computer was purchased is perceived as slow today. It is not impossible, however, that my computer does in fact work at a slower pace, due to some software problems, say.

What these examples illustrate is that presented with an utterance that is in conflict with preceding utterances of the same speaker, we cannot read off whether it results from a change of beliefs (change of opinions, change of theory) or from a change of meanings. On an intuitive level, however, we can clearly distinguish two cases. Either the beliefs about the world have changed; perhaps – as in the case of our computer – because the world itself has changed, or perhaps – as in the Perugona case – because our information about the world has changed (while the world itself has remained the same in all relevant respects). Or else, the denial of a previously asserted sentence is due to a difference in the rules that govern the usage of words, as in the examples where the meanings of the words “reliable” and “fast” have changed.\(^{31}\)

\(^{31}\) Semanticists have pointed out to me that in the second example the meaning of “fast” does not really change, it is only the parameter that fixes the relevant threshold value that changes. Against this, one might argue that intuitive meaning is the formal semanticist’s meaning plus a certain value of the parameter. I don’t want to do this, since the point of the two little examples in the text is just that there is an undeniable intuitive difference between belief change and meaning change. No further claims are made at this point.
How can we determine whether the change from a given statement to its negation constitutes a substantial change of theory, or rather a change of the façon de parler? I think that we can give two answers to this question, an objective one and a subjective one.

The objective answer reminds us of the fact that theories, after all, are not merely something enshrined in our heads but will usually have certain empirical parts or consequences – often called observation sentences – that are amenable to a more or less immediate perceptual verification or falsification. If the two different theories or belief systems in question make different predictions about (or admit different ranges of) observations, then we can at least potentially reach a decision between the two theories on empirical grounds. In so far as the two sentences in question contribute individually to the diverging empirical claims, they cannot simply be translation variants of one another.

There are epistemologists and philosophers of science who doubt the very possibility of such experimenta crucis, or generally any direct confirmation or weakening of theories by empirical findings. They view theories in a more coherentist way, as networks of propositions where even allegedly “basic” observation sentences are dependent on theoretical assumptions (like, e.g., assumptions about the presence or absence of imperceptible disturbing factors that would help to reconcile the data with incompatible predictions or explanations). On such an account, there is no way of deciding in favor or against a given theory on empirical grounds alone. We might then try using a subjective method to determine whether we are facing theory change or meaning change, by inquiring if the speaker is inclined to say that he simply was wrong with his previous beliefs, and that things are really different from what he had thought before. If this is the case, then we have a case of genuine belief change; otherwise we can hope to find a suitable non-homophonic translation procedure to explain away the apparent difference between the two theories.

A remark in closing: Differences in theory or meaning can, of course, also be conceived as differences prevailing between persons or cultures, not only as differences along the time axis of a single person or community. Note, however, that the temporal variant has two major advantages:

(a) We may assume that the linguistic and epistemic systems at time $t$ and $t+1$ largely coincide, to a larger extent at least than what we may reasonably expect when we compare languages or belief systems of varying persons or cultures.
(b) The subject has privileged access to both linguistic and epistemic systems at different stages of her own identity. Hence, she is the prime expert in making a comparison of these two stages, and we can talk to her. Any comparison of two persons or cultures will naturally be a much more delicate task.

The *subjective method* of deciding between meaning change and theory change in the interpersonal or intercultural case may perhaps proceed by investigating whether it is necessary to quarrel about “who is right” or by looking for an interpreter *well-acquainted* with both persons or cultures, who is able to judge and resolve all apparent conflicts.

## 5. Summary and related work

I have put forward for discussion a way of making philosophical sense of the meaning of theoretical terms and of the changes thereof. The proposal is based on an explication of analytic judgments that is in turn inspired by Quinean ideas without, however, subscribing to his scepticism about meaning. Inspiration from Quine is drawn in three respects: the use of the revisability of sentences as a vehicle for the explication of *analyticity*; the acknowledgement of substantial difficulties in drawing a line between knowledge of facts (beliefs, theories) and knowledge of language (meaning); and the recommendation of a pragmatic solution to these difficulties that refers to speakers and communities of speakers.

My position differs from Quine’s mainly in that I consider the investigation of the meaning of theoretical terms and analytic sentences a meaningful task – assuming that we can get reasonable intuitive judgments about identities of languages or theories in general, as well as a precise model for theory revision. The explications given in section 3 were an attempt to provide a basis for fruitful further research.

Later in life, Quine partially rehabilitated the concept of analyticity; see, *e.g.*, *The Roots of Reference* (1974, §21): “Here then we may at least have a line on a concept of analyticity: a sentence is analytic if *everybody* learns that it is true by learning its words. Analyticity, like observationality, hinges on social uniformity.” According to this passage, someone who thinks that an analytic truth is false has not learnt his words properly. Quine seems to find this an unproblematic explication; it is however not related to
the many other explications that can be found in *Two dogmas*. It is, on the other hand, fairly close to Locke’s explication of “maxims” which are truths that are “generally assented to, as soon as proposed, and the Terms they are propos’d in, understood.” (*Essay*, Book I, Chapter 2, §17, original emphasis). Similarly, the “epistemological conception of analyticity” in Boghossian (1996: 363) says that “a statement is ‘true by virtue of its meaning’ provided that grasp of its meaning alone suffices for justified belief in its truth.” That means that someone who is not ready to embrace an analytic statement has not grasped its meaning. The present approach is more charitable in interpreting the speaker. If a bona fide analytic statement does not get accepted, the diagnosis is not that the speaker does not understand the words that he is using. Rather the point is that he understands the words differently (as compared with the common use in the relevant speaker community). And a different understanding is here taken to mean that a change in language – or alternatively, an essential change in theory – has taken place.

In a recent paper, Pagin (2001) advances an approach that is on the face of it quite similar to the one presented in this paper. He, too, uses the revisibility of propositions as the central concept for the explication of synonymy and analyticity.\(^{32}\) His basic idea is that two expressions are *synonymous* if and only if substitution of the one for the other in any statement preserves the revisibility of the statement (likelihood of re-evaluation in terms of truth-value in the event of recalcitrant experience). When Pagin (2001: 23) says that “analytic” is a vague predicate admitting borderline cases, he seems to point to the comparative notion of analyticity that I suggested above. There are, however, important differences with my view. According to Pagin (2001: 14), analytic statements are just as hard to revise as logical truths and nothing is harder to revise than a logical truth. In contrast to Pagin (2001: 18–22), however, I do think that

\(^{32}\) Pagin distinguishes between what he calls *analyticity*-s (which is just what Boghossian terms “Frege-analyticity”: transformability into a logical truth by substitution of synonyms) and *analyticity-fl analyticity*-e (corresponding to Boghossian’s “metaphysical”/“epistemic” analyticity: truth/acceptance in virtue of meanings, independently of matters of fact). Pagin argues that the latter has been demolished by Quine while the former is defensible and may indeed be identified with *analyticity simpliciter*. In my opinion, however, the analysis based on comparative revisibility shows how the concept of analyticity can be thought of as a single one after all.
Theoretical concepts in flux

(1) Eye doctors are eye doctors
is harder to revise than

(2) Ophthalmologists are eye doctors

It is indeed hard to imagine that any experience could make one abandon statement (1), but further specialization of ophthalmology might well lead to a new distinction of nomenclature. So even though (2) is analytic, (1) is harder to revise than (2). Unlike Pagin who suggests that minimal revisability is a necessary condition for analyticity, the account presented in this paper suggests that low revisability – revisability on pain of dropping out of the language and theory games other people play – is a necessary and sufficient condition for analyticity (with the qualifications mentioned).

Finally, I do not think that it is illuminating to distinguish “theoretical revision” from “linguistic revision” by simply saying that the former is motivated “by the need to avoid untenable theoretical or observational consequences” while the latter is motivated “by terminological needs” (Pagin 2001: 24). This seems far too casual to me. I have tried to argue in this paper that the recourse to identity judgments about language and theories is a more promising route. But in any case, making sense of the difference between substantial changes of theories and mere changes of ways of speaking is crucial if we want to understand how theoretical concepts can be in flux.

References

Achinstein, Peter

Alchourrón, Carlos, Peter Gärdenfors, and David Makinson

Ayer, Alfred J.

33 Compare Harman’s (1996: 399–400) point that “stipulative definitions” are assumptions that may be abandoned just like any other theoretical principle.
Bealer, George
Boghossian, Paul Artin
Carnap, Rudolf
Einstein, Albert
Eisler, Rudolf
1900 Wörterbuch der philosophischen Begriffe. Berlin.
Frege, Gottlob
Gärdenfors, Peter
Gärdenfors, Peter and Hans Rott
Gergonne, Joseph D.

Goethe, Johann Wolfgang von

Hansson, Sven Ove

Harman, Gilbert

Hoyningen-Huene, Paul

Kitcher, Philip

Kripke, Saul A.

Lakatos, Imre

Nahin, Paul J.

Pagin, Peter

Putnam, Hilary


Quine, Willard Van Orman


1974  The Roots of Reference. La Salle, Ill.: Open Court.

Reichenbach, Hans


Rott, Hans


Schlick, Moritz


Shapere, Dudley


Stegmüller, Wolfgang


Wittgenstein, Ludwig  
Meaning change as character change*

Ulrike Haas-Spohn

1. Introduction

This paper offers a reconstruction of Putnam’s account of natural-kind terms and theoretical terms within David Kaplan’s context theory. It will allow us to represent epistemic as well as metaphysical aspects of the determination of reference within a single notion of meaning, namely, the Kaplanian character. It also entails an account of meaning change which confirms the familiar doctrine of the theory dependence of meaning and at the same time retains Putnam’s insight that the reference of natural-kind terms remains constant across theoretical change.

I shall start with an informal sketch of Putnam’s theory of natural-kind terms, emphasizing its motivation from philosophy of science. Next, I shall propose a format for the meaning of natural-kind terms in terms of Kaplanian characters, a task that will crucially rely on an explication of the notion of usage of words. Finally I shall reconstruct meaning change within the account presented. I shall distinguish three kinds of meaning change, understood as character change, and demonstrate the utility of the approach with some central examples.

2. Putnam’s theory of natural-kind terms and its relation to the Philosophy of science

Putnam’s (1975b) proposal for a semantics of natural-kind terms\(^1\) must be understood against the background of long-standing debates in the philosophy of science. On the one hand, Putnam rejects the verificationist notion

---

* I am indebted to Regine Eckardt and Wolfgang Spohn for translating my German version of the paper into English and to two anonymous referees for valuable proposals to improve the paper.

of meaning and the resulting logical empiricism, as well as the so-called "Californian" semantics\textsuperscript{2} which he sees as still a basically verificationist approach. On the other hand, he argues against the view propagated by Kuhn (1962) and Feyerabend (1962, 1965, 1975), who held that theory changes affect not only the meaning of words but also their reference, to such an extent that we are confronted with a new world with a new ontology after each scientific revolution.

Generally speaking, Putnam objects to an epistemic notion of meaning and promotes a realistic semantics. The basic idea is this: Words are not tied to descriptions or definitions that are constitutive for their meaning. Neither do words acquire their meaning by the role they play within some scientific theory. Either of these views would entail that the meanings of words change along with changes in our knowledge or our theories; and there would be no basis for claiming that some later theory was more adequate than an earlier theory, or that it leads to enhanced knowledge about a certain subject matter: If the terms of the earlier and the later theory differ in meaning, the theories cannot contradict each other (or at least not in the simple way commonly assumed in science). It may even be that a false theory actually refers to nothing at all, because there is nothing that satisfies its descriptions. Thus we have to live with the possibility that our own words do not refer at all. Kuhn escapes this problem by assuming that each theory talks about its own ontology and in particular about its own natural kinds. But this move does no better in enabling us to define or explain knowledge accumulation and scientific progress.

All these views are misguided, according to Putnam, and do not capture what scientists really do or even how we use everyday language. Scientific language, he claims, is based on "transtheoretical notions" and relies on the fact that one and the same property or kind can be dealt with by different theories, possibly replacing wrong statements about it by true ones. The relevant example, used by Putnam several times\textsuperscript{3}, is the word "electron": There exists nothing that would match Bohr’s early description of electrons, simply because he did not yet know the uncertainty relation discovered by Heisenberg in 1927. Yet electrons, as we know them today, follow Bohr’s description at least approximately, e.g., with respect to mass and charge and many observable effects produced by electrons. Hence it is rea-

\textsuperscript{2} Putnam thereby refers to possible worlds semantics in the tradition of Carnap (1947).

\textsuperscript{3} Cf., e.g., Putnam (1975c: 275).
Meaning change as character change  \hspace{1cm} 209

Reasonable to assume that Bohr and his fellow experts actually were talking about these particles when they used the word “electron”.

These observations about language use are not restricted to science. Everyday language contains at least some words, the so-called natural-kind terms, that we faithfully use on the assumption that they refer to one and the same natural kind independently of our current state of knowledge. Let us assume, to use one of Putnam’s examples, that there were pieces of metal which were indistinguishable from gold at the time of Archimedes, but could be identified as a different metal by contemporary chemical tests. It is clear then that these pieces are not in the extension of our word “gold”. But it is also clear that even Archimedes, calling them “gold”, would have made a false assertion – even though he would not have been able to discover its falsity. Similarly, if we traveled to a remote planet, the famous Twin Earth, and discovered a liquid which superficially looked and behaved like our water, but turned out upon closer investigation to consist of XYZ instead of H₂O molecules, we would again conclude that this liquid is not water. And even if we had encountered the liquid on Twin Earth before we mastered the relevant chemical tests, we would have been wrong in calling it “water”.⁴

If we accept Putnam’s claims, then neither a word’s extension in the real world nor its extension in counterfactual worlds (i.e., its intension) can be captured by definitions or descriptions that apply to all and only the intended referents. In a way, Putnam turns the traditional view upside down: it is not the intension which determines the extension, but rather the extension of a word in the real world that determines its intension (or extension in all possible worlds). This is why Putnam refers to natural-kind terms as “hidden indexicals”, and why his semantics is called “realistic”.

In reversing things, Putnam has the following picture in mind. Natural-kind terms have a certain usage in a language community, and this usage crucially depends on the existing paradigms, on standard samples that are taken to be instances of the respective kind. These paradigms will fix the extension of the word both in real and counterfactual situations: other pieces of matter are gold, or water, because they share the nature of the paradigm samples of “water” or “gold”, because they obey the same objective laws. According to current physics, this amounts to sharing the molecular structure of the paradigm samples of gold/water, because this structure is what determines the physical behavior of the respective substance.

⁴ See Putnam (1975b: 233ff, 235) for a discussion of these examples.
Putnam draws a parallel between his proposal and the theory of proper names in Kripke (1972). Kripke’s acts of baptizing associated with proper names correspond to how natural-kind terms are tied to paradigmatic samples. For both Kripke and Putnam, the usage of these terms originates in concrete objects. The reference of natural-kind terms in counterfactual worlds is determined by the physical nature of those objects, and likewise the reference of proper names in counterfactual worlds is determined by the essential properties of the baptized individuals (which according to Kripke lie in the genetic origin of the individual). This view propounded by Kripke, Putnam, and others has gained considerable fame under the label “causal theory of reference”.

By clearly distinguishing between epistemic and reference-fixing aspects of meaning, Putnam enables us to assume that words retain their original reference even though our beliefs and theories about the referents change, at least as long as the old and the new theories are still linked by sufficiently similar paradigmatic samples.

Putnam proposes to capture the epistemic aspects of meaning by his notion of a stereotype. Stereotypes comprise the minimal knowledge required by speakers if they are to be said to master the corresponding words. Stereotypes can be regarded as Putnam’s concession to the descriptionist. The picture offered by the causal theory of reference says only that any user who is properly connected to the usage of a word, say “gold”, is talking about gold. Here the proper connection requires that the user has acquired the word “gold” from another person who herself uses it in the appropriate paradigmatic way or has correctly adopted it in turn, and moreover that the user sincerely intends to talk about the same things as the person from whom she has acquired the word. The user need not know much about gold; concerning the reference of “gold” she may always rely on the gold experts of her community. This is the division of linguistic labor. Intuitively, however, it would be inappropriate to grant someone the mastery of the word “gold” if he points, say, at a snowball or a strawberry and asks “is this gold?”. Such a person should not qualify as an appropriate member of the causal-intentional network that constitutes the usage of the word “gold”. Therefore, Putnam adds the epistemic requirement that speakers should at least be familiar with the corresponding stereotype.

---

5 See, e.g., Devitt and Sterelny (1999).
The question of what should count as the minimal amount of knowledge necessary for the mastery of a given word is vague and depends obviously on the cultural and social context. In everyday conversation the mastery of the word “electricity” will hardly require more than to know that it is what makes lamps give light and lets the kitchen equipment function. In the context of an expert discussion, however, one would certainly go as far as to exclude anyone from the relevant linguistic community who is not sufficiently acquainted with Maxwell’s equations.

Ultimately, Putnam describes meanings as complexes consisting of four different components: syntactic properties, semantic type, associated stereotype, and the extension itself. However, the mechanisms of reference fixing – in the case of scientific terms, the theories that codetermine what counts as a paradigmatic sample – are not part of Putnam’s semantic theory. Only by excluding them from our notion of meaning we seem to be able to guarantee that reference remains constant while theories change.

In what follows I want to reconstruct Putnam’s insights in a different manner, namely in a way which fits them into traditional views of compositional semantics and which also allows the integration of epistemic aspects of meaning in an appropriate way. I think that this task will turn out to be quite simple once Putnam’s claim that natural-kind terms are hidden indexicals, i.e., that their meaning depends on the actual world, is taken literally.

3. A reconstruction of hidden indexicality in terms of Kaplanian characters

In traditional intensional semantics meanings are construed as intensions, i.e., as functions assigning (categorially appropriate) extensions to possible worlds (or to more complex indices). In order to provide a semantic analysis of (overtly) indexical or context-dependent expressions like “I”, “here”, or “that”, Kaplan has developed a formal framework in which the basic notion of meaning is explicated by a two-dimensional construction: a function mapping pairs of a context and a world (or a more complex index) into extensions (which may be alternatively conceived as a function-mapping contexts into intensions in the previous sense. Such functions are called

---

7 Cf. Putnam (1975b: 268ff.).
characters. A word is context dependent (or indexical) if it does not have the same intension in all contexts.

Kaplan argues in favor of the double parametrization of meaning by pointing out that we have to distinguish between (metaphysical) necessity and contingency on the one hand and apriority (= epistemic necessity) and informativity on the other. Sentences like “I am here now” are, in a sense, always or a priori true, since they cannot be uttered falsely (unless we use “here” in a demonstrative or anaphoric way). However, the propositions expressed by the utterances of “I am here now” are never metaphysically necessary, since the speaker could always have been somewhere else at that time. An opposite example is provided by the sentence “I am David Kaplan”, uttered by David Kaplan. In this case, the sentence expresses a necessary truth, namely the self-identity of David Kaplan. Nevertheless the sentence can convey new information to the hearer.

This distinction can be captured in two-dimensional semantics. Apriority or informativity are properties of sentences. A sentence \( \phi \) is a priori true iff it is true in every context. Here, \( \phi \) is true in a context \( c \) iff the character of \( \phi \) assigns truth to the context \( c \) and to the context world \( w_c \) – taken now as the point of evaluation. \( \phi \) is (potentially) informative or a posteriori iff it is true in at least one but not all contexts. Necessity and contingency, by contrast, are properties of utterances. A sentence \( \phi \), uttered in a context \( c \), is necessarily true iff its intension in \( c \) is the set of all possible worlds, and the utterance expresses a contingent proposition iff its intension is neither empty nor the set of all possible worlds.

Hence we can use characters to derive a metaphysical (or realistic) as well as an epistemic notion of meaning. The former results from considering a sentence in a fixed context and computing the intension (i.e., the set of possible worlds in which it is true) determined by the sentence’s character in that context. The latter results from collecting the set of contexts in which the sentence is uttered truly; this set, or rather its characteristic function, is also called the diagonal of the character because it is obtained by diagonalizing the character, i.e., by evaluating the character only at those pairs \( \langle c, w_c \rangle \) of contexts and worlds in which the index world is the same as the context world.\(^8\) Of course, intensions and diagonals of other expres-

\(^8\) The notion of a diagonal is not explicitly used by Kaplan. I have borrowed it rather from Stalnaker, who has introduced the two-dimensional notion of a propositional concept resembling, as well as differing from, Kaplan’s notion of a character. Stalnaker applies his notion to various problems, for instance, to ex-
Meaning change as character change

This framework clearly suggests a way to capture the hidden indexicality of natural-kind terms along the same lines. Natural-kind terms behave like the indexicals and demonstratives considered by Kaplan with respect to metaphysical necessity and potential informativity. Following Kripke and Putnam, we have to claim that a sentence like “water is H\textsubscript{2}O” is necessarily true (assuming that it is true), and at the same time we would not call it an a priori truth.

Of course, the overt context dependence of “I” (on the speaker), or “now” (on the time), or “this” (on the speaker’s pointings) is quite different from the context dependence of natural-kind terms like “gold”. This was emphasized by Putnam (1975b: 234) himself. Whether “water” means H\textsubscript{2}O or XYZ does not depend on where “water” is uttered, and it depends on the speaker only insofar as she belongs to one or the other linguistic community. In using the word “water” the man from earth refers to H\textsubscript{2}O even on Twin Earth. However, the claim that the extension of “water” varies with the linguistic community would not make much sense either. After all, we want to assign meanings to the words of a given language and not to syntactic forms outside the context of a language.

Hence, the contextual aspects that are relevant for a natural-kind term are not local features like speaker or place, but rather global ones, i.e., the nature of the paradigmatic instances of the term in the context world. As Putnam notes, it is the world itself that determines the meaning of words, an assertion that may be reconstructed in Kaplan’s framework by saying that the intension of a natural-kind term depends on the context world. This leads us to the following definition of the character of natural-kind terms:

plain the informativity of identity sentences with two proper names which express necessary truths (cf. Stalnaker 1978). The account I am about to develop is indebted to Stalnaker’s theory at least as much as to Kaplan’s. Presumably, however, both would not agree to my use of their ideas. Stalnaker would not want, to mention just one point of difference, to use propositional concepts as semantic values representing the meanings of the expressions of a language; he rather considers them as parts of a pragmatic theory of utterance meaning (cf., e.g., Stalnaker 1987: 182f., or 1999: 5ff.). For the difference between Kaplan’s and my conception of characters see footnote 11.

9 What I call diagonals and intensions corresponds, respectively, to Chalmers’ primary and secondary intensions and also to Jackson’s A (actual) and C (counterfactual) intensions; cf. Chalmers (1996: sect. 2.4) and Jackson (1998: ch. 2).
Suppose $N$ is a natural kind term of a language $L$ at time $t$, $c$ a possible context in which $L$ exists as it is at $t$, and $w$ a possible world. Then

$$\| N \|^{L,t}(c)(w) = \{ x \in w \mid x \text{ shares in } w \text{ the essential properties of most of those objects in } w, \text{ to which the usage of "} N \text{" in } L \text{ at } t \text{ is paradigmatically applied}\}.$$ 

This is, of course, a very schematic meaning rule. However, it allows us to clearly distinguish the various components that contribute to determining the extension of words and to show the interactions between them.$^{10}$

Obviously, the notion of usage carries the main weight in the above definition. The usage of a word determines what counts as a paradigmatic application of the word. The current criteria for “water” paradigms, for instance, are not confined to their being liquid, transparent, and thirst-quenching, but include also properties like “behaves like $\text{H}_2\text{O}$ in electrolysis”. In current English “water” is unsuited not only for referring to beer or soft ice, but also for referring to quantities of $\text{XYZ}$, since they do not satisfy the more theoretical criteria. Current science and its theories along with all its operationalization procedures, measurement devices, and experiments, etc., will contribute to the usage of a word, at least in its scientific sense. The criteria provided by the usage of some term $N$ in a given language at some time thus consist in the best methods available to the linguistic community, including its experts, at that time for determining whether or not something is an $N$.\textsuperscript{11}

$^{10}$ This meaning rule is discussed in much more detail in Haas-Spohn (1995: sections 3.1–6). I would like to emphasize two points, though. First, like Kaplan (1990), I conceive of languages not as merely formal objects assigning characters to syntactically specified expressions, but as specific historic products (see Haas-Spohn 1995: 136f., for the identity conditions of languages so conceived). Second, one should observe the partiality of the characters just defined: the character of $N$ is defined only for those contexts in which the state of the language $L$ at $t$ and hence the relevant usage of $N$ exist (a restriction which I have explained and defended in Haas-Spohn 1995: sect. 3.3, in particular 101f.).

$^{11}$ At this point it is very clear that the character of “water” as just specified describes the semantic knowledge of a linguistic community as a whole and not the individual knowledge of normally competent speakers. It is this social knowledge which matters in the examples discussed by Putnam and Kripke, since their point is...
Meaning change as character change

By themselves, however, these criteria do not normally fully determine the intension of a word, i.e., to which objects in the actual and in counterfactual worlds it is correctly applied. Within the limits set by the usage, the intension of “water” is finally fixed only by the real world, by the actual nature of the liquid which constitutes the paradigmatic referents of “water”. And precisely because there is still room for variation at this point, as Putnam’s examples demonstrate, “water” is a hidden indexical.

It is this range of variation which is characteristic for natural kind terms. However, this range is something granted by the linguistic community; it is part of its usage of the word “water”. The usage of “water” includes assumptions about what kind of properties may be essential for water (namely, its physical structure), and it leaves it to the world to fix the properties that are actually essential for water. The presupposition that there actually is a unique physical structure underlying the paradigm cases is part of these assumptions about the essence of water.

Thus, the crux of the matter is that there may be essentially different paradigmatic referents of the usage of “water” in different context worlds which cannot be distinguished even by the best identification methods currently available. This opens sufficient epistemic room for varying essences of the paradigms. It finally explains why the sentence “water is H₂O” may be necessarily true, though it is not a priori, i.e., true in all possible contexts.

For clarification, let us go through our example once again. Consider the usage of “water” at the time that modern chemistry started to develop so that one could already consider, but did not yet believe the hypothesis “water is H₂O”. Thus, the best contemporary means to identify water were still traditional, surface-oriented ones. At that time, the diagonal of the sentence “water is H₂O” obviously did not assign truth to all contexts for which it is defined, because among these contexts there are some in which water is a different substance (e.g. XYZ) and which are not yet excluded by the usage of “water” at that time.

not the incomplete knowledge of individuals, but the divergence of metaphysical necessity and apriority even in the case of perfect semantic competence. The difference between communal and individual apriority is not relevant for Kaplan, since it does not show up in indexical expressions like “I”, “here”, “now”, or “this”. Presumably, my communal interpretation of characters is not acceptable to Kaplan, since his characters are designed to capture what’s “in the head”. 
Have matters changed since? Not really. According to my proposal, the informational content of a sentence is represented by the diagonal of its character. Today’s experts believe that water essentially consists of H₂O, and their current methods of identification, like electrolysis, are chosen accordingly. In current English, the diagonal of “water is H₂O” is the set of those contexts in which all those pieces of matter that share the essential properties of the paradigm applications of the current usage of “water” consist of H₂O – where it is part of the current usage that experts believe that the paradigm examples consist of H₂O molecules and thus behave like H₂O, say, in electrolysis. Contexts in which “water” is used to refer to XYZ or to liquids of some other chemical composition no longer belong to the domain of the character of “water”, since they would not conform to our current usage of “water”. Hence it may seem that there are no contexts left in which “water is H₂O” could turn out false, and that the sentence has thus lost its informativity.

However, this would be a false impression. It is still conceivable that we find differences in the microstructure of H₂O molecules which we cannot discover with our present scientific methods. In that case there would be two kinds of H₂O, kind A and kind B. Now suppose that our actual water consists only of H₂O of kind A. Under these circumstances our sentence “water = H₂O” would be false. Hence, the character of “water = H₂O” is defined for such a context, but its diagonal excludes this context. Therefore the sentence is still informative.¹²

4. Implications for meaning change

There are two parameters in the character definition above which determine reference: the usage, providing the fallible epistemic element, and the nature of the paradigms and hence the nature of the context world, providing the realistic element. Both aspects can vary either independently or jointly, and hence we can distinguish three different types of meaning change in terms of character change.

The first type applies to words the usage of which shifts so that different objects become paradigmatic referents and the words thus acquire new in-

¹² This possibility is perhaps exemplified by the discovery of isotopes and of deuterium in particular. However, I do not want to comment on the issue whether or not heavy water is to count as water.
Meaning change as character change

Meaning change as character change in all contexts. “Rabbit”, for instance, was originally applied only to the young of the animal; only in the 17th century the word was extended to include the adult animal. “Dog” originally referred to a specific race and later on extended to the present usage. “Deer”, by contrast, was originally used for animals in general, and later on it was restricted to refer only to antlered ruminants. In this case, usage and intension change together. This type is the familiar case of meaning change as we observe it in everyday language for all sorts of words, not only for natural-kind terms. The criteria of usage are shifted, narrowed, or broadened, and if the word is not a hiddenly indexical natural-kind term, these criteria suffice to fix its extension and intension. In the present context, where we are primarily concerned with scientific language, this is certainly the least interesting case.

Is the definition of character given above still compatible with Putnam’s general views, even though it incorporates an epistemic element? This question will lead us to the second type of meaning change. We have stated above that Putnam’s theory of natural-kind terms was in part motivated by the intuition that the meaning of a word can remain constant despite changes or gains in knowledge and theory making use of that word. Do we still capture this intuition?

At the very least, this constancy is not automatically excluded as it was by description theories of meaning. Consider first our simplest standard case, “water”. According to the notion of usage that I proposed above the present usage of “water” differs from the earlier usage of “water”, say, in 1750, because our theories about water, and hence identification methods, have changed: today, chemists have much more direct access to the molecular structure of a liquid. Yet this gain in knowledge has no impact on the intension of “water”. The intension of “water”, as used today in the actual world \( w_t \), is the function that assigns to each possible world the totality of liquid in that world that consists of \( \text{H}_2\text{O} \) molecules. However, according to our meaning rule the word “water” had exactly the same intension in 1750, because neither the paradigmatic referents nor their nature have changed since. Hence we are still talking about the same natural kind as we did in 1750, although our knowledge about it has tremendously increased.

Still, in another sense the meaning of “water” has changed: its character has changed. In the real world \( w_t \) the past and the present word “water” have the same intension, but this is not true for other context worlds. Considering English “water” at a time \( t \) before the discovery of electrolysis, say

---

13 I have taken these examples from Room (1986).
in 1750, we might imagine possible contexts in which, for some mysterious reason, the English speaking population were confronted with large quantities of XYZ liquid. These quantities of XYZ would then have been integrated into the paradigmatic applications of “water”. Hence, the word “water” in English, at time \( t \), and under such counterfactual circumstances, would denote both substances; its intension would be the function that assigns to each possible world \( w \) the set of pieces of substance consisting of \( \text{H}_2\text{O} \) or of XYZ molecules.

We may place the English language in such contexts also at times \( t' \) after the invention of electrolysis. However, as soon as a substance’s behavior under electrolysis has become one of the criteria for identifying water, all the suddenly appearing quantities of XYZ will no longer be subsumed among the paradigmatic referents of “water”. Hence, the word “water” in English at \( t' \) still refers to \( \text{H}_2\text{O} \), but not to XYZ, and its character yields for that context again the intension that assigns to each possible world \( w \) the set of pieces of substance consisting of \( \text{H}_2\text{O} \) molecules. This implies, however, that the word “water” has different characters, and in this sense different meanings, at \( t \) and \( t' \).

The case is summarized by the figures below which exemplify the character of “water” at 1750 and at 2000. \( w_1 \) is the actual world in which all quantities of water consist of \( \text{H}_2\text{O} \) molecules. \( w_2 \) is to be a world in which there exist both quantities of \( \text{H}_2\text{O} \) and quantities of XYZ, in the vicinity of the language community. Finally, \( w_3 \) is a world containing only XYZ. The horizontal headline lists worlds as points of evaluation, the vertical headline takes worlds as relevant parts of contexts. Thus, a row represents the intension of “water” in the relevant context.

<table>
<thead>
<tr>
<th>1750</th>
<th>( w_1 )</th>
<th>( w_2 )</th>
<th>( w_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_1 )</td>
<td>( \text{H}_2\text{O} )</td>
<td>( \text{H}_2\text{O} )</td>
<td>( \text{H}_2\text{O} )</td>
</tr>
<tr>
<td>( w_2 )</td>
<td>( \text{H}_2\text{O} \cup \text{XYZ} )</td>
<td>( \text{H}_2\text{O} \cup \text{XYZ} )</td>
<td>( \text{H}_2\text{O} \cup \text{XYZ} )</td>
</tr>
<tr>
<td>( w_3 )</td>
<td>( \text{XYZ} )</td>
<td>( \text{XYZ} )</td>
<td>( \text{XYZ} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2000</th>
<th>( w_1 )</th>
<th>( w_2 )</th>
<th>( w_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_1 )</td>
<td>( \text{H}_2\text{O} )</td>
<td>( \text{H}_2\text{O} )</td>
<td>( \text{H}_2\text{O} )</td>
</tr>
<tr>
<td>( w_2 )</td>
<td>( \text{H}_2\text{O} )</td>
<td>( \text{H}_2\text{O} )</td>
<td>( \text{H}_2\text{O} )</td>
</tr>
<tr>
<td>( w_3 )</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
The figures show that the intension of “water” has changed in \(w_2\) and \(w_3\), but not in the actual \(w_1\); hence, the character of “water” has changed as well. (In \(w_3\) the character of “water” as used in 2000 is no longer defined, since that usage does not exist in \(w_3\).)

The example demonstrates that changes in usage need not entail changes in intension. This is so at least if the changes in usage involve at most minor changes in what counts as a paradigmatic referent. Thus we face here a second type of meaning change as character change: change in usage without change in intension.

Putnam’s writings suggest that this type constitutes the normal case of meaning change for scientific terms. It covers not only examples like “water” or “gold”, but also examples like “electron” and “fish”. The latter example, however, shows particularly clearly that our meaning rule above crucially depends on vaguely referring only to most paradigmatic referents: If we assume that the use of “fish” at earlier times included whales and dolphins as paradigmatic referents, as well as herrings and sharks, then the intended result – namely, that the present as well as the past intension of “fish” really covers only fish, and not fish, whales, and dolphins – can only be reached if we can neglect whales and dolphins as paradigmatic referents. However, we are licensed to do so by the tacit assumption (which was made even in earlier times) that all paradigmatic referents are essentially of the same kind. In order to be able to maintain this tacit assumption we may, and need to, exclude whales and dolphins from the paradigm samples.

Admittedly, there is no very clear answer to the question of how to talk when the assumption of a common nature underlying all applications of the usage of a word turns out to be wrong. Asking “how to talk” points to the fact, by the way, that there is also a normative element in fixing meanings, something like a “principle of charity”, as Putnam himself stresses. If we find only a few “black sheep” among the paradigm cases, we are justified in ignoring them. If we find that our paradigms are divided into two or three natural classes equally entitled to the common name, it seems reasonable to allow for a “disjunctive” nature. If, however, it should turn out that the paradigm cases share no essential property whatsoever, there is no choice but to give up the assumption of an underlying common nature and to rely on phenomenal or functional criteria for determining the term’s extension.

Putnam (1975b: 241) offers “jade” as an example for a word that covers two kinds of minerals, namely jadeite and nephrite, and thus has a disjunctive nature. In this case it seems reasonable to assume that the intension of
the word “jade”, today as in the past, is the function that maps every possible world \( w \) onto the lumps of substance that consist either of jadeite or of nephrite.

Other cases, however, are less clear. One of the favorite examples in philosophy of science is the term “mass” and the question of how it changed its meaning when Newtonian mechanics was replaced by Relativity Theory. Newton only knew one mass, described by Newton’s laws. By contrast, Relativity Theory implies that one has to distinguish between the mass of bodies in rest and of bodies in movement, though this distinction will become measurable only for movements at fairly high speeds. At first glance, this looks very much like the “jade” case: Newton really was talking about both kinds of mass all the time, but his experimental and theoretical means were too poor to realize that two different kinds were in play. However, this view is presumably not very plausible. It would imply that two statements “\( a \) has mass \( x \)” and “\( a \) has mass \( y \) (\( \neq x \))” might have been true at the same time. Nowadays, “mass” is an ambiguous word that may mean either “rest mass” or “relativistic mass”\(^{14}\). But do we really want to claim that “mass” was an ambiguous term in Newton’s time as well, although nobody could appreciate the ambiguity?\(^{15}\)

Let us finally turn to the third type of meaning change. It occurs when the usage of a word remains constant, and yet the nature of the paradigm cases changes without the linguistic community realizing it. Such cases are possible in principle, as we saw in the above thought experiment: if speakers of English had encountered large quantities of XYZ before the development of modern chemistry (due to their travels, say, or due to vulcanic activity), then the intension of “water” would have changed unnoticeably and hence without effect on the usage of the word.

\(^{14}\) Cf. Field (1973), fn. 12. \\
\(^{15}\) Field (1973) defends the view that “mass” was referentially indeterminate before the invention of relativity theory, referring partially to relativistic mass and partially to rest mass. Assertions that contain referentially indeterminate expressions are considered true (or false, respectively) iff they are true (false) for all potential denotations, and otherwise undefined. A proposition like “\( a \) has mass \( x \)” was therefore, strictly speaking, undefined in Newton’s time. By contrast, a proposition like “the larger the mass of a body, the more energy is required to accelerate it” would have been true even then. In my opinion, the “mass” example is analogous to the case of Eustacia Evergreen discussed in Kitcher (1978: 526f.).
Meaning change as character change

Eckardt (2001, this vol.) discovered that such a case of character change has actually occurred, namely in the word “jade” in Chinese. The Chinese word *yu* was originally used only for nephrite (plus a number of “black sheep” referents); jadeite was unknown in early China. Starting in the 18th century, jadeite from Burma was imported into China in larger quantities, and these stones were immediately accepted as paradigmatic instances of *yu* (“jade”). When it was discovered, somewhat later, that the paradigm samples of “jade” consisted basically of two minerals, scientific usage at least was standardized so that “jade” (*yu*) covered exactly those two chemical substances.

Here, our meaning rule states that the intension of the Chinese word *yu* before common trade with Burma yielded, for each possible world, only the totality of nephrite in that world (neglecting the black sheep cases). After trade with Burma has developed, the intension of “jade” has changed and turned into that function that yields, for each possible world, the totality of nephrite and jadeite in that world.

Eckardt holds that this is an undesirable consequence of Putnam’s theory: that there may be instances of meaning change that occur entirely unnoticed by speakers (and even experts). I do not think that this objection applies to Putnam directly. On the one hand, Putnam does not want to develop an epistemic notion of meaning and hence needs not respect such intuitions at all. On the other hand, he might point out that the intuition that no change has occurred is already captured by the fact that the usage has not changed in these examples and that in this sense the concept of jade has remained the same.

The objection might be more serious if put forward against my epistemic reinterpretation of Kaplan’s character theory. It has been the aim of my paper to complement intensions as a metaphysical kind of meaning with the diagonals of characters as an epistemic kind of meaning. However, in the case just discussed even the diagonal of “jade” changed. This shows that the diagonal of a character as explained above is still not a purely internal notion of meaning that captures the internal state of a linguistic community.

Indeed, the characters and their diagonals, as I have introduced them above, describe what might be called situated usages. The usage of “water” may be the same on earth and Twin Earth, and the usage of *yu* may be the same in the 17th and in the 20th century. Yet, the mere fact that these usages are situated differently in space and time, e.g., before and after large jadeite imports, entails different characters and diagonals. In my opinion,
this spatio-temporal localization is indeed part of the identity of a language, which would support the picture that I developed here.\textsuperscript{16}

References

Carnap, Rudolf

Chalmers, David

Devitt, Michael and Kim Sterelny

Eckardt, Regine

Feyerabend, Paul


Field, Hartry

Haas-Spohn, Ulrike

\textsuperscript{16} Cf. also Haas-Spohn (1995: 135ff). Kupffer (1999: ch. 5) has shown, however, how we can abstract away from this kind of situatedness and reach a notion of meaning that indeed describes nothing but the internal state of the respective linguistic community. In Kupffer’s account, then, there can be no unobserved meaning changes, and \textit{yu} will not have changed its meaning in Chinese during the crucial period.
Meaning change as character change

Jackson, Frank

Kaplan, David

Kitcher, Philip

Kripke, Saul

Kuhn, Thomas S.

Kupffer, Manfred

Putnam, Hilary

Room, Adrian
Stalnaker, Robert
Meaning change in conceptual Montague semantics

Regine Eckardt

1. Introduction

Truth-value based semantics in the tradition of Montague (1974) has successfully treated a wide range of semantic phenomena. Yet, so far it is a purely synchronic theory. The aim of my paper is to extend this theory by a diachronic dimension in a meaningful way. I will proceed in two steps. First, I will propose a theory of how a new word and its meaning are introduced into language. The proposal will be nominalistic in spirit and replace the currently favored causal theory of reference. Second, I will show how contexts of language use can, in retrospect, be interpreted as hypothetical contexts of meaning introduction, and thus give rise to new meanings of an old word – depending on the historical, cultural, and social circumstances – in a systematic fashion. Finally, I will discuss how the resulting account for word meaning in Montague grammar integrates notions like “stereotype” and “default knowledge”, notions that capture indispensable aspects of our semantic competence and yet often are treated in a way strangely independent of the core semantic theory.

2. The snapshot: Synchrony

2.1. How are new words established in a language?

If we want to investigate the question how speakers can change the meaning of words in a language, it might be a good starting point to ask how speakers can establish the meaning of a new word in the language. Within truth value based semantics, the causal theory of reference (see Kripke 1972 and Putnam 1975 with a wealth of subsequent literature) is at present still the most widely accepted account for this question. The introduction of a new proper name into a language is the paradigm case of the account.
Kripke (1972) argues that the meaning of a name is established in an initial act of ‘baptizing’, in which the name is “attached” to an individual $R$. In terms of intensional semantics, this leads to the so-called rigid designators as in (1). Let $\text{name}$ be a proper name in the language under investigation.

(1) Assume that the proper name $\text{name}$ has been introduced into the language in baptizing an individual $R$. The name $\text{name}$ will then refer to $R$ under all possible circumstances:

$[[ \text{name} ]]$ is that function $f: D_S \rightarrow D_e$ where $f(w) = R$ for all possible worlds $w$.

The account was subsequently extended to natural kind terms (that is, words like “gold” or “water”) which are treated like proper names for natural kinds. While this sounds simple, it implies that one has to say what a natural kind is, and how the kind relates to an isolated sample of that kind. Putnam (1975) proposes that the meaning of a natural kind term is established in an initial act of baptizing with reference to some sample $R$ of the natural kind in question. He furthermore assumes the relation of “being identical as a substance” that states for any other lump of matter (in whatever possible world) whether it is subsidentical to sample $R$. Hence, $R$ and subsidentity will determine the full extension of the natural kind, in all possible worlds. Haas-Spohn (1994; this volume) developed an explication of Putnam’s ideas in a Kaplanian framework. The definition in (2) extends Haas-Spohn’s (1994) theory to the initial act of baptizing a natural kind. Here, $\text{nkt}$ is the prospective natural kind term and $R$ the sample that is pointed at in the baptizing. Subsidentity is meant to be the relation that holds between two lumps of matter just in case if they are of the same substance.

(2) $[[ \text{nkt} ]] = f: D_S \rightarrow D_{(e,t)}$ where $f(w) = \{ b \mid b \text{ is a lump of matter in } w, \text{ and } b \text{ subsidentical to } R \}$

Obviously, the subsidentity relation plays a core role in the account. It rests on the assumption that there is one and only one correct way to categorize.

---

1 Haas-Spohn treats the meaning of a natural kind term at times later than the baptizing. Definition (2) differs from hers in that the initial samples $R$ are usually no longer available to and sometimes even irrelevant for later generations of speakers who instead rely on known paradigm referents of $\text{nkt}$. 
the realm of substances, of animals, of plants. Natural sciences pursue the
task of finding out the correct category boundaries, but scientific theory
might still be fundamentally wrong about categorization – while the correct
boundaries are eternally mirrored by subsidentity. This position is better
known as philosophical realism: Crucially, it leads to the prediction that the
meaning of a word like “gold” does not depend on how much we know
about gold (both as individual speakers and as speaker community). The
position of realism has earned its high reputation in philosophy by the fact
that it can, better than any other position, account for language use in scien-
tific progress (see Devitt and Sterelny 1999 for an introduction).

From a linguist’s perspective, however, the account is less satisfying. I
will list three drawbacks.2

1. The missing “natural kind term” convention: For proper names, there
is a linguistic convention that one name should refer to one per-
son/object, even if speakers might not have operative criteria to iden-
tify the referent or to distinguish it from similar persons/objects. This
convention becomes visible, for instance, when someone realizes
that he has used one name for two people. He will immediately
check which one of these two carries the name, or whether the two
persons happen to carry the same name (i.e., there are two homon-
ymous proper names in the technical sense). No comparable situations
arise for common nouns, which means that there is no comparable
linguistic convention for an alleged class of natural-kind terms in
language. Without such a convention, however, we have no lingui-
stic evidence for a class of natural-kind terms like those posited by re-
alism.

2. The danger of describing the nonexistent: What if nature does not
fall into “natural kinds”? We must already suspect this in the case of
animal and plant species; where modern genetic engineering has
taught us to think in terms of a genetic continuum (at least poten-
tially) rather than a clear-cut taxonomic system. The same is possi-
ble, at least as a thought experiment, for chemical substances. Phi-
losophical realism entails that if there are no natural kinds, then there

---

2 (2.) has already been pointed out in Zemach (1976); (1.) and (3.) to my knowl-
edge have not yet been stated in this clear-cut form. I refer the reader to the article
of Haas-Spohn in this volume where realism is defended against objection (3.)
are can be no natural kind terms either. (Remember that, as shown in (1.), the class of natural-kind terms lacks any linguistic characterization and hence, the only common feature of natural-kind terms is that they refer to natural kinds.) We might describe the non-existent.

3. Spurious meaning changes: I will use the following criterion to diagnose meaning change in an empirical sense: There is a time at which speakers can say: “We would call this thing here XY today, but in former times, they wouldn’t.” Or: “In former times, they would have called this thing here XY, but today, we don’t.” Philosophical realism postulates instances of meaning change where this criterion does not apply. Meaning changes occur without anybody before or later noticing: Whenever speakers intend to use a word as a natural kind term and – rightly, according to their own classification criteria but erroneously, according to the ultimate scientific classification and subsidentity – apply the name to objects of another substance (i.e., something that is not subsidentical to previous samples), realism predicts that the meaning of the word has changed. Of course, this is the core point of Putnam’s twin-earth example (which I will not repeat here, as it is discussed in full length in Haas-Spohn, present volume), but interestingly, real instances of this example have taken place in language history, and it turns out that linguistic intuition in these cases does not exactly match the philosopher’s intuition about the twin-earth case.

Putnam’s twin-earth case is exemplified by the history of the word “jade” (and its Chinese translation yu): The word was originally established on the basis of the substance nephrite. When jadeite was first encountered, speakers extended the word to pieces of jadeite without any notion that two different substances could be involved and they cherished the new pieces as “jade” (yu) of the best kind (for historical details, see Eckardt 2001 and Turner 1996 [entry “China”]). Crucially, the extension of “jade” / yu was not corrected as soon as modern chemistry could identify the “error”. No one felt that the word had been used wrongly at any point, or that earlier and later uses of the word were different. Realism will postulate a meaning change from jade(1) = nephrite to jade(2) = nephrite+jadeite.3

3 Of course, the other way out of the dilemma consists in claiming that “jade” (or, yu) never was a natural kind term. This, however, highlights the prescriptive char-
Besides, the causal theory of reference only applies to a very small fraction of even the nominal lexicon. The vast majority of nouns, adjectives, and verbs refer to classes of objects or events whose boundaries are not drawn by any science, and where boundaries can be drawn tighter or wider with equal justification. Thus, there seem to be several reasons to ask for a more flexible, less externalist account.

2.2. Can we give a more realistic, less “realistic” account of baptizing?

It is certainly an appealing idea that the class of objects denoted by a word \(\Theta\) should somehow “cluster around” those things which were first or previously called “\(\Theta\)”. We will leave it open whether the first uses were deliberate baptizings in the literal sense (like names for babies) or whether the word \(\Theta\) was established in a more implicit manner. Can we retain the appealing aspects of the classical theory while avoiding its drawbacks?

In definition (2), realism is reflected in the relation of subsidentity. Reality itself determines which objects \(a\) and \(b\) are identical in this sense. Subsidentity between \(a\) and \(b\) holds true no matter whether anyone knows this, and no matter whether even a majority of even reputable speakers believe that \(a\) and \(b\) are not of the same kind.

In order to come to a more realistic, intuitively more appealing theory of word meaning, we should integrate the idea that categorization is primarily a human strategy to sort our environment into things that are more or less alike. We can easily do this by replacing subsidentity by one of a family of similarity relations \(\sigma\). Loosely speaking, a word \(\Theta\) applies correctly to the first sample \(R\) which was referred to in introducing \(\Theta\) into the language, plus any other object \(K\) which resembles \(R\) in the appropriate sense (\(=\sigma(K,R)\)). This will require some comments.

Firstly, this account reflects the insight that things can be similar in more than one respect. Different choices of respect-in-which-things-are-compared will also lead to different classes of things-which-are-similar-to-

4 One reviewer offered the simple example of two quantities of liquid, a glass of Kölsch and a glass of Pils, which can be classed as “the same” – both being beer – or as “different” – ask the connoisseur! – with equal right. The same problem was brought up for natural kinds in Sterelny (1983), see below.
A classic discussion of this observation (but by no means the only one) is Goodman (1972). In establishing a new word $\Theta$ by pointing at a sample object $R$, speakers accordingly need to specify the respect in which other referents of $\Theta$ should resemble $R$. This theoretical assumption can be observed at work in experimental studies of word learning like Landau et al. (1998) (discussed below). More relevant to our enterprise, the mechanism is reflected in word history and word polysemy, where we can find many examples in which a word’s readings can be reconstructed as different kinds of generalizations of a common set of sample referents.

Secondly, what is the formal nature of these similarity relations? I propose to adopt a topological notion of similarity like that of Gärdenfors (1998, 2000). He argues that similarity is most naturally determined on the basis of a notion of distance in some conceptual space. An object $b$ is similar to another object $a$ if it is “close enough to $a$” in terms of that distance. If the question is whether $b$ is like $a_1$, or like $a_2$, ... or like $a_n$, where these $a_i$ are prototypical members in different classes which we already know, then we will say that $b$ should go with the prototype $a_j$ which is closest to it in terms of that distance. By comparing the distances of some object $b$ to several other objects, this topological notion can also capture the fact that similarity or non-similarity is often a question of anti-samples as well as samples: Kölsch is certainly “similar to Pils” if the only alternative would be “be similar to milk”. It is not evidently “similar to Pils” if we restrict our attention to alternative types of beer. – Nothing in the following will crucially depend on the formal nature of similarity.

Thirdly, what factors determine the intended similarity relation in question? Sterelny (1983) observed that even assuming realism, the classical causal theory of reference needs to specify the taxonomic level at which “identity” is to be determined. For instance, should a sample dog Fido stand for the class of male dogs, of dogs, of canines, of mammals, ... ? Assuming that biology is on the right track towards the ultimate categorization of animal species, we have to acknowledge that scientifically all these classes are equally well justified. Sterelny argues that this *qua question is determined by the context in which a baptizing takes place. I will adopt his proposal and assume that the utterance context of a baptizing will determine the respect in which the sample is to be generalized. Unlike Sterelny, I will not restrict the choice of classes to those licensed by an elusive reality but will allow for any similarity that suggests itself to the human speaker. Scientific categorization is hence justified by current scientific theory, and we allow for cultural, perceptual, functional categorization as well (mean-
Meaning change in conceptual Montague semantics

ing that Fido might also stand for the class of pets, of watchdogs, of food, etc.).

We will come back to contexts and distinguish several independent contextual factors that conspire in determining similarity, factors like the knowledge of the speakers and their community, the interests they pursue in their conversation, the range of objects under consideration, samples and anti-samples, etc. With this complex notion of “context” in mind, we can state that the *qua question* is answered by context in pointing out some similarity relation $\sigma$. Thus, we come to the following modified theory of baptizing:

(3) If word $\Theta$ has been introduced in some baptizing context $c$ on the basis of sample $R$ then

$$[[ \Theta ]] = \text{a function } f \text{ from worlds to sets of objects such that }$$

$$f(w) = \{ x \mid x \text{ is an object in } w \text{ that is } \sigma\text{-similar to } R \}$$

where the similarity relation $\sigma$ was the most salient one in context $c$.

This proposal can be read as a nominalistic and epistemic interpretation of possible-world semantics. It is no longer reality which cuts out the classes of objects denoted by a word: Similarity is rather a conceptual notion, reflecting how speakers see the world. Moreover, extension and intension depend on speakers’ interests and knowledge at the time of introducing $\Theta$.

In the next sections, we will turn this view of word introduction into a moving picture of meaning development. I want to elaborate the idea that all situations in which speakers utter a word $\Theta$ can in retrospect function as if they had been baptizing situations: The ways in which we used $\Theta$ yesterday will set the limits of today’s readings of $\Theta$.$^5$

---

$^5$ The resulting picture can even be interpreted as nominalistic in a strong sense. Not only are similarities the only universals that we rely on. We will also model the growth of a full spectrum of meanings of a word $\Theta$. Such clusters of meanings all grew out of previous readings at some time in a motivated manner but are, synchronically speaking, indeed only tied together by the same word $\Theta$. 

3. The moving picture: Diachrony

3.1. Words in time: Introducing a temporal index

We need to distinguish the meaning of a word Θ at time \( t \) from the meaning of the same word Θ at time \( t' \). We will make the idealizing assumption that time proceeds in discrete steps. Eventually, we want to derive the potential meaning(s) of word Θ at time \( t+1 \) from its meaning at time \( t \), the exemplars that were considered typical Θ’s at time \( t \), plus the circumstances under which word Θ was predominantly used at time \( t \). (“Predominantly” can mean both “predominating in quantity”, or “by socially dominating speaker groups”.) However, we will first clarify the formal nature of the temporal index.

I will base my considerations on synchronic semantic models \( M \) in the Montagovian tradition, where meanings are objects in models of type theory (Montague and Thomason 1974). These models will have to be augmented by several ingredients which will be introduced in the following sections. As a first extension, all models \( M \) will contain a set \( T \) of points of time which carry a discrete linear order. If \( t \in T \), I will use \( t+1 \) to refer to the point in time which immediately follows \( t \). The trans-temporal meaning of word Θ is a function from \( T \) into appropriate synchronic meanings of Θ.

(4) If \( \Theta \) is an expression in \( L \), then the diachronic meaning of \( \Theta \) relative to some model \( M \) is \( [[\Theta]]^M = \text{a function, mapping } T \to [D_S \to D_{\alpha}] \) where \( \alpha \) is a logical type matching the syntactic category of \( \Theta \).

\( [[\Theta]]^M,t \) the synchronic meaning of \( \Theta \) at \( t \), is \( [[\Theta]]^M \) applied to time \( t \).

Note that \( [[\Theta]]^M,t \) is not the same as “the things we’d call \( \Theta \) and which existed at time \( t' \) in the sense of temporal logic. If we take the noun meat, then \( [[\text{meat}]]^{1000AD} \) is the meaning of meat in the sense used at 1000 AD (roughly “food” in the modern English sense). The extension of modern meat at \( z=1000 \text{ AD} \) in the sense of temporal logic, in contrast, would be “everything in 1000 AD which would qualify as a meat in the modern sense of the word”. Clearly, we need two independent temporal parameters to describe language-in-time in contrast to extension-over-time. Temporal logic is currently treated in an extensional fashion (e.g. Musan 1997 or Kamp and Reyle 1993 who argue against implicit time indices) by assuming an explicit time parameter for at least verb, nouns and adjective mean-
Meaning change in conceptual Montague semantics

ings. It is less clear in what sense more “grammatical” words like conjunctions, particles, or certain adverbials should have a time-dependent extension.

In the next section, I will introduce the notion of “accepted referents of word Θ at time t”. This will require us to link extension-in-time and meaning-in-time (of Θ), because we should reasonably expect that the accepted referents of Θ (in the sense it has at time t) should be in the extension of Θ at times t’ before and up to t. This is not the place to introduce a full version of tense logic into our framework; I will instead make the simplified assumption that all words Θ under consideration are of some type (s, (z, α)) where s = possible worlds, z = times (in M) and α an appropriate type. A common noun, for instance, would be of type (s, (z, (e,t))).

(5) shows the meaning of a common noun over time:

(5) If Θ is a common noun, then the diachronic meaning of Θ relative to some model M, [[ Θ ]]_M, is defined as follows:
[[ Θ ]] := a function, mapping T → [ Ds → [ Dz → D(e,t) ]]
[[ Θ ]]_t will be interpreted as [[ Θ ]] applied to time t, that is, the meaning of word Θ at time t. The extension of Θ (language stage t) at time z and world w is computed as [[ Θ ]]_M(t, z, w)

If we apply this to the above example, we get the following distinctions:

(6) [[ meat ]]_{1000AD} = the intension of the word meat in the sense of 1000 AD.
[[ meat ]]_{1000AD}(w_0)(2001) = all food in our real world in the year 2001
[[ meat ]]_{2001AD} = the intension of the word meat in the current sense
[[ meat ]]_{2001AD}(w_0)(2001) = all meat in our real world in year 2001

6 Note that many details wait to be spelled out in a full version of this idea. Verbs will pose the extra challenge of linking up their event time with the language stage; according to current views they do not carry a time parameter, but tense + aspect is thought to operate on the event parameter.
3.2. The reliable core

If speakers introduce a new word $\Theta$ into their language, they do so on the basis of one or more objects, let us call them “samples”, which they from now on want to call $\Theta$. The full extension of $\Theta$ will cluster around these first samples in the way defined in (3), section two.

We face a different situation if speakers don’t start but continue to use word $\Theta$ at time $t+1$ where $\Theta$ was already part of the language at time $t$. It is reasonable to expect that speakers will use $\Theta$ in a conservative fashion: The things that were traditionally referred to by $\Theta$ at time $t$ should still be called $\Theta$ at time $t+1$. Moreover, the full extension of $\Theta$ at time $t+1$ should cluster around these traditional instances of $\Theta$ at time $t$, much as in (3).

Building on the notation of the previous section, let $\text{SAMPLE}$ be a polymorphous function that maps intensions of type $(s, (z, \alpha))$ at time $t$ onto a subset of the extension up to time $t$:

\[
(7) \quad \text{For word } \Theta \text{ and time point } z, \text{ let }
\text{SAMPLE}( \Theta )(z+1) \subseteq \bigcup_{z' \leq z'} [ \text{M}_z( \text{SAMPLE}( \Theta )(z') )]^z,
\]

be the traditionally accepted, “typical” referents of $\Theta$ at time $z$.

The characterisation of $\text{SAMPLE}$ in (7) ensures basically two things:

(a) The samples of today $z+1$ are among the referents of former times ($z' \leq z$). They will eventually also become referents of today (see below), but it does not make sense to recruit today’s samples from today’s referents.

(b) Only referents in the real world will count for the future development of the meaning of word $\Theta$.

Our theory is thus restricted to words $\Theta$ which take their departure in our actual environment. We do not deal with words that acquire their meaning by pure definition in terms of other words: Such words may have empty extension at the time of introduction, but of course they have a meaning. The theory in its present shape is also unable to capture non-definitional

\[\text{Note that “subset” for relational types means “subset of the relation”, whereas for functional types that do not end in } t, \text{ it means “subset of the function” in the literal sense.}\]
Meaning change in conceptual Montague semantics

words which yet have an ostensive flavor, words like “unicorn”. The word “unicorn” clearly does not have a meaning in terms of a definition, and neither has anyone ever been able to define “unicorn” by ostension in the real world. It seems plausible to capture such words “as if” they were introduced by ostension, with the minor change that the sample in question is given by description. Keep straight that this is not the same as claiming that the word “unicorn” has a descriptive meaning. Describing a sample, one makes the tacit addition “… and all σ-similar animals are also called unicorn”, which invites the speaker to generalize the given description in reasonable ways as suggested by context.

Interestingly, a certain amount of variation is already implicit in this dynamic notion of samples. Let me show why. By introducing the notion of samples, we are freed from the burden of the entire extension of Θ up to time z. It is not required that the samples of Θ of today t need to remain samples at t +1 (= tomorrow). The world may change, and what was yesterday’s typical referent of Θ may have become rare today. To see an example, you only need to consider “good” referents of a word like “car” over time. Formally we refuse restrictions (9) and (10) as universal principles, although something like (11) might usually hold true. It will be an investigation in its own right to ask about developments that lead to more radical re-estimations of sampling and even violate (11).

\[(9) \quad \text{SAMPLE(Θ)(z)} = \text{SAMPLE(Θ)(z+1)} \quad \text{(identity)}\]

\[(10) \quad \text{SAMPLE(Θ)(z)} \subseteq \text{SAMPLE(Θ)(z+1)} \quad \text{(conservativity)}\]

\[(11) \quad \text{SAMPLE(Θ)(z)} \cap \text{SAMPLE(Θ)(z+1)} \neq \emptyset \quad \text{(non saltat)}\]

3.3. Generalizing first samples: Similarity relations

The main claim of section one was that it was worth trying to develop a non-“realistic” version of the causal theory of reference. The core step in that enterprise was suggested in section 3.2, namely replacing subsidentity by similarity-in-a-certain-respect. Section 3.2 also related this suggestion to proposals elsewhere in the literature, but little to no exemplification of the core idea was given. This gap will be filled now.

Shape and function are probably the two most prominent respects in which an initial sample can be generalized. The meaning of words like “tri-
angle” or “circle” (in their everyday sense where also a somewhat shakily drawn triangle counts as a triangle) are clearly based on the idea of “being similar in shape to a prototype triangle”, “~ ... circle”.

Shape even seems to be something like a default criterium of classification (at least for humans). It is the first respect used by children in categorization (see, e.g., Keil 1994). It is the respect both children and adults will use in absence of any other clues (see Landau et al. 1998). And, as you can test for yourself, practically all words which refer to visible things have a reading that can be paraphrased as “looking like a core instance of ...” (The mountain Matterhorn for instance is called so because it looks like an animal horn, not because it actually is one.)

Function is often used as the alternative respect in which things can resemble each other (exemplified once more by Landau et al. 1998). Note, however, that unspecified function will often be too broad for our purposes. Artifacts are usually built to fulfill a specific function, and I take these more specific functions-for to stand for single similarity respects under the general header of “function”. A word like “can opener” exemplifies this kind of similarity. Can openers are created for the purpose of opening cans, and everything which resembles core can openers in this respect will qualify as a can opener. An object which resembles (my) core can openers in the functional respect of “being an instrument with which one could severely hurt enemies” will not qualify as a can opener (wrong function). An object which looks like a core can opener but does not open cans like one will only be a “can opener” in the shape sense of the word.

Yet, things can be even more intricate. Words for traditional tools like, e.g., “hammer” do not depend on function alone. At least I would hesitate to call anything which can draw nails into wood a “hammer”, no matter how it works. The reason is perhaps that a core hammer is also of use for so many other purposes (like opening nuts, flattening dents in my car, ...) which other devices could not fulfill. The fact that the trackball (or, recently, trackpad) of my laptop is not called a “mouse” offers another example. We have to take into account the possibility that similarity can be based on a combination of function and shape.

Worse, we have to face the fact that a certain amount of deliberate definition can play a role in determining whether something x resembles something R sufficiently to call both objects by the same name. In the case of artifacts, the inventor or creator has, to a certain extent, the right to decide in what class the new artifact should fit. The decision will often depend on marketing strategies—the case of the computer “mouse” can once again
serve as an example. Calling your new trackball a “mouse”, you suggest that you have merely improved the old “mouse” somewhat. Calling it a “trackball”, you stress the fact that you have invented a device that may fulfill the old function but is radically different from the old tool in all other respects. (Guess what would be the better marketing strategy.) Such deliberate instances of classification are discussed in Bloom (1996).

3.4. What is a context?

It was proposed in section two that the context (of baptizing, and sometimes of use of a word Θ) will suggest some similarity relation. This similarity will then determine the class of objects that are represented by the samples. Now that I have given a more specific idea of what kinds of similarity will play a role, we can also specify in some more detail the contextual factors that determine this choice.

If someone introduces, or uses, some word Θ in a context c, then c will determine the following four aspects:

(a) When does c take place? – Time of c.

(b) Who were speaker and addressee in c? – Speaker and hearer, and in a broader sense: language community they are part of.

(c) What was the conversation about? – Interests of speaker and hearer in c.

(d) What was the relevant range of objects? – Referents and anti-samples.

Indirectly, the time and speaker community of c will determine another important factor, namely the epistemic state of the speaker community at time t (= “what they know”). Knowledge and interests are two major independent factors that determine the choice of some similarity relation. Let me briefly illustrate this.

Assume that we have two contexts c₁ and c₂ in which speakers with similar interests talk about some lump of matter R. Assume that these interests are such that the speakers are concerned with R as a natural substance. However, assume that the speaker community in c₁ has the epistemic back-
ground of an ancient Greek scholar while the community in \( c_2 \) is in possession of modern chemistry and physics. The similarity relation chosen in \( c_1 \) will therefore be based on features like color, hardness, mass, etc., while the similarity relation in \( c_2 \) might use such sophisticated criteria as mass spectroscopy or DNA sequencing. This can (but need not necessarily) have the results that the extension of a word \( \Theta_1 \) for “\( R \) and what is similar”, introduced in \( c_1 \), is different from the extension of a word \( \Theta_2 \) for “\( R \) and what is similar”, introduced in \( c_2 \). Thus, contexts providing equal interest but different knowledge can lead to different concepts. The experimental studies of categorization behavior of small children described in Keil (1994) confirm the assumption that categorization is dependent on available knowledge.

The opposite case would be a pair of contexts located at the same time, and within the same speaker community, but with different interests. Clearly, different interests will also result in different ways to generalize a sample. This assumption is supported by much of literature in psycholinguistics, recently for instance in Landau et al. (1998).

The range of objects to be classified is another factor in determining whether we want to draw looser or tighter boundaries. As in footnote 4 mentioned above, the question of whether Pils and Kölsch are the same or different things will crucially hinge on the range of liquids that are to be classified. Another simple example can be found in color classification. Carmin red and zinnober are clearly “the same” relative to the full color spectrum but different if attention is restricted to the shades of red.

Why do we have to distinguish between speakers and speaker community? Evidently, we do not want to make word meanings dependent on the knowledge of arbitrary individual speakers. The idea of linguistic “division of labor” was proposed in Putnam (1975), a proposal that has been accepted without much objection. Putnam argued that single speakers are not themselves in full command of the meanings of words because they very often have to rely on experts in their society in order to determine the correct extension and intension of a word. Speakers, so to speak, master “parametrized” versions of meanings that leave the details to be filled in by experts.

In that sense, single speakers in some given context \( c \) may also themselves not be fully aware of the kind of similarity relation in play. They will simply know that word \( \Theta \) covers \( R \) and everything that is “somehow similar to \( R \), and ask the expert what that will mean in detail”. In the case of baptizing, there might still be some sociologically directed correlation be-
3.5. Words in time: Yesterday’s uses determine the meanings of today

Throughout this section I will use the time variable \( t \) to stand for “yesterday” and, consequently, \( t+1 \) for “today”. Furthermore, let \( \Theta \) be the word we want to describe in time (and, to keep matters simpler, we will assume that \( \Theta \) was introduced into the language before time \( t \)). We can now consider the set \( C(\Theta, t) \) of all contexts \( c \) at time \( t \) in which \( \Theta \) was used. At least some of them will determine some similarity relation \( \sigma \). Crucially, these similarity relations \( \sigma \) need not be the same ones as the similarity relation \( \sigma_0 \) on which the meaning of word \( \Theta \) at time \( t \) is based. (In the following, I will consistently use \( \sigma_0 \) in this sense.) For instance, exemplars of a substance might become interesting in a certain function, etc. In order to obtain a similarity relation for each context \( c \), I will assume that in absence of hints as to what similarity is the intended one, the context points out the similarity relation \( \sigma_0 \) leading to the actual meaning of \( \Theta \) at time \( t \).

Speakers at time \( t+1 \) will remember how word \( \Theta \) was used at \( t \). In particular, they will collectively remember if there was a wealth of contexts which all suggested the same similarity relation \( \sigma \). We can safely assume that it is part of our speaking abilities that we master the act of baptizing, or generally, the introduction of a new word into our language. (Otherwise, speakers would not have been able to learn their mother tongue at all. No one could ever become a speaker without this ability.) Thus, speakers can imagine what the meaning of word \( \Theta \) would have been, had all these contexts been contexts of word introduction, not contexts of use. This is reflected in the following definitions. Definition (12a) offers the general case, while (12b) once more applies (12a) to the case of common nouns to show what it comes down to in a concrete case.
(12) a. The set of potential readings of $\Theta$ at time $t+1$ consists of the following functions that map possible worlds and time points into objects of type $\alpha$:

\[
\{ f \mid \text{There is a context } C(\Theta, t) \text{ that renders } \sigma \text{ salient, and for all worlds } w \text{ and times } z, \]
\[
f(w)(z) = \{ x \mid x \in D_{\alpha} \land \exists R ( R \in \text{SAMPLE}(\Theta)(t+1) \land \sigma(x,R) \text{ in } w \text{ at time } z) \}\}
\]

b. The set of potential readings of some noun $\Theta$ at time $t+1$ consists of the following functions that map possible worlds and time points into sets of objects:

\[
\{ f \mid \text{There is a context } C(\Theta, t) \text{ that renders } \sigma \text{ salient, and for all worlds } w \text{ and times } z, \]
\[
f(w)(z) = \{ x \mid x \in D_{e} \land \exists R ( R \in \text{SAMPLE}(\Theta)(t+1) \land \sigma(x,R) \text{ in } w \text{ at time } z) \}\}
\]

This definition sets limits within which the word $\Theta$ may develop new readings at time $t+1$. Not all potential readings need to become actual readings. If a certain reading was only suggested by one single context at time $t$, chances are good that this potential reading passes away unnoticed. However, if a wide range of contexts at time $t$ suggests a new reading, speakers may decide to establish this reading as a new sense in which the word $\Theta$ can be used. Note, moreover, that the original meaning of $\Theta$ will be transported automatically to time $t+1$, due to the fact that all neutral contexts of use formally render $\sigma_0$ salient.

4. Summary and further questions

4.1. Summary

I have presented the outlines of a theory that augments formal semantics by a diachronic dimension and links up subsequent language stages in a meaningful way. Generalizing the mechanisms of naming by ostension, I proposed a formalism that allows us to derive word meanings at stage $t+1$ from language use at time $t$. The core idea is the following:

It is part of our language ability that we can establish the meaning of a word on the basis of some (core) samples, and a respect in which subse-
quent objects in the extension of the word must be similar to these first samples.

This ability is exercised in word introduction (baptizing). Yet, it also influences the subsequent lives of words. Our knowledge of the meaning of a word $\Theta$ today is based on the fact that we know what was typically called $\Theta$ yesterday, and the further fact that we know in what respect other things that we can call $\Theta$ must resemble these typical $\Theta$’s of yesterday.

Assume that some speakers are faced with samples $R$ for $\Theta$ – how do they know the appropriate similarity relation $\sigma$ needed to form a class of objects around $R$? I suggested that contextual factors will determine the choice of $\sigma$ (with the provision that perhaps not all contexts are rich enough to suggest some choice; however at least for contexts of baptizing it is vital that some $\sigma$ is salient.) The contextual factors which play a role in this choice comprise at least the knowledge of the speaker community at the time, and the interests of the speakers interacting in the specific context.

Speakers of today will realize whether for some word $\Theta$, in the contexts in which $\Theta$ was used yesterday, a similarity respect $\sigma$ became salient that was different from that respect $\sigma_0$ which leads to the "old" meaning of word $\Theta$. Consider some examples: If the substance name “chalk” is predominantly used in contexts where that substance serves as a writing tool, this suggests a generalization of the referents in the functional dimension. The Latin *pipa* (= reed) was encountered in varying functions leading both to *pipe* (like flute) and *pipe* (like tube). Examples can also be found in the verbal domain. Eventualities of *watching* (“being awake”) gain most attention when the state of being awake is maintained against one’s inclination to sleep: *watch* in the sense of “remaining awake”. One step further, “remaining awake” is often for a specific purpose, reflected in the nautical term *watch* (like guard; a similar development has taken place for German *wachen* = be awake to *wachen* = remain awake on purpose, for instance as a religious exercise). We may only speculate as to what contexts of “remaining awake” and “guarding”, presumably in the presence of enemies or captives, invited the transitive reading of *watch*. All these word histories exemplify how salient similarity respects in a context of use can open up a new potential reading of word $\Theta$. It is, to a certain extent, a question of historical accident which ones of those are actually established.

---

8 The interesting history of the English verb *watch* is traced in detail in Verdauger and Poch (1998).
A further source of potential meaning shifts lies in the assumption that the “best” referents of $\Theta$ of yesterday are the samples of today. The nature of these best referents $\Theta$ can change, and due to such changes at the core of a category, the boundaries may change as well. Moreover, core exemplars of a different nature also can lead to new salient similarity relations.

This account offers an explicit link between non-linguistic external factors, and linguistic factors interacting in meaning change. This should not be surprising for an account of diachronic semantics, as it is almost an armchair insight that historical accident may drive word meanings in quite arbitrary directions. Let me briefly label the parameters of the suggested account, according to the academic discipline that would host them.

The nature of the samples (= the most common referents of $\Theta$ of yesterday) will contingently depend on historical and factual accidents. If you live in Europe, it is a matter of fact that the typical “swan” is a white, elegant waterbird. This is not a biological necessity, neither is it a linguistic constant – it is simply an accident.

The nature of the contexts of use of $\Theta$, as well as the frequency of contexts with equal interests and knowledge, will depend on historical accident. It was a historical accident that people at a certain time talked about the natural product “horn” (=part of an animal) with an interest in drilling holes into it to produce a musical instruments. By this accident, the similarity relation $\sigma = \text{“similar in the respect of producing musical sounds”}$ (roughly) became salient enough to lead to a second reading of “horn” (= brass instrument). Only a change in the nature of the typical horn-instrument can have led to the classification as a brass instrument, because the original musical horn evidently wasn’t made of brass. Interestingly, another reading of “horn” arose in contexts where (musical) horns were blown mainly in the function of “giving alarm”: Car horns are not musical instruments but share the function of alarming.9

The ability to recognize similarities in various respects, and the ability to use them in categorization, is part of the bio-psychological makeup of human beings (and not only theirs). Without this ability there would be little point in talking about individuals, or objects, or relations, or categories. In this respect, the position developed here is in sharp contrast to the leading metaphysical doctrine of realism. Due to limitations of space, the reader is referred to other work where this kind of “non-realism” is defended in

---

9 I want to thank the reviewer who brought up this further step in the development of “horn”.
more detail, notably the introductory part of Jackendoff (1983), and Gärdenfors (1998, 2000).

The ability to turn these ingredients into words with meanings, finally, is clearly a linguistic ability of humans, and of humans alone. Note that this ability goes beyond merely attaching some sound label to some idea/concept. It is part of our knowledge of the meaning of a word that we also know how to combine it with other words to construct larger meaningful objects. Following this initial idea, my account was carefully tailored as a conservative extension of truth-value based semantics.

As a result, my notion of “meaning” combines aspects of conceptual and logical semantics. On one hand, words denote objects in the realm of type logic. In this sense, I follow the tradition of analytical philosophy according to which a word’s extension and intension are the cornerstones of its meaning, the safest and clearest aspect of the meaning of a word such that we would say: “Two speakers agree on the meaning of a word \( \Theta \) if and only if they agree on its actual extension (here) and its possible extensions (counterfactually), that is, its intension.” On the other hand, these denotations come with an internal structure. Extensions derive from sets of objects plus strategies of generalization, and importantly these strategies of generalization reflect human cognitive abilities and are not the result of the one, eternal, independent and true way of cutting the world into slices. In this respect, the approach takes up leading ideas of cognitive semantics, relates them to new empirical data and stands in opposition to the picture drawn by philosophical realism.

4.2. Discussion

Previous approaches to word meaning, especially those that are couched in logical semantics, usually restrict their attention to the case where words are assigned meaning by definition. It is known from mathematics that cross-refering sets of axioms can successfully code the meaning of such simple relations as “smaller than” or operations like “plus”. Yet, the claim that a full natural language can get all the semantics it needs from a huge collection of meaning postulates remains to be demonstrated. It is more likely that the lexicon is based on a set of words with meanings-by-ostension and extended recursively by definition and more ostension. My account focusses the basic ostensive level of the lexicon.
Still, one might ask how the descriptive word meaning as produced by single speakers, or the knowledge offered by traditional encyclopedia, relates to my account of word meaning. It is one advantage of the approach that we can easily locate what has elsewhere been called “stereotype knowledge” or “default knowledge”: This is simply knowledge about the samples for a certain word. It is usually sufficient to describe samples of a class and leave it to the speakers to do the necessary generalization. Different kinds of vagueness can arise at this point. Let me give two examples.

Take the word “fox”. I know enough about foxes to be able to describe, and maybe even find, a typical exemplar of a fox. I also know enough about biology and English to know that the similarity relation in question should be “similar in the sense of being of the same species”. However, I must leave it to experts to explicate the criteria that apply in the case of foxes. In other words: I can not, for an arbitrary animal, determine whether it is a fox or not. Nevertheless, this kind of knowledge seems sufficient for me to know the meaning of the English word “fox”.

However, I must confess to use words for which I know typical exemplars but have no idea as to how these should be generalized. To reveal my ignorance of English: I know what a typical instance of “cattle” looks like, but I am uncertain whether these should be generalized to all exemplars of the species, or whether the boundaries are drawn on the basis of functional aspects (like being owned and exploited by a farmer). In this case, it is questionable whether the speaker (= I) knows enough to know the meaning of the word (= cattle). Yet, that speaker (once again, me) can fight her way through a good deal of conversation without this lack of knowledge becoming apparent.

This kind of ignorance has, of course, previously been acknowledged as a lamentable but inevitable communicative practice. Putnam’s theory of stereotypes, among others, attempts to describe the semantic competence of all those speakers who are not in possession of the real meanings of words. Yet, Putnam’s stereotypes are semantic objects that live in complete isolation from what he considers to be the true meanings of words (as characterized in (2)). It is not clear in his account whether and how true meanings can be approximated by or computed from stereotype knowledge. This situation is particularly unsatisfying for the working linguist because it

---

10 And remember that there are quite a lot of them, according to Putnam. It is questionable whether the true meaning of any word can ever be mastered by a single speaker, in terms of Putnam’s theory.
suggests that she has the choice between a neat-and-clean semantic theory for a kind of language that nobody uses or an “ugly”, poorly developed semantic theory that, however, will at least apply to some real speakers.

The present picture offers reason for hope: Stereotype knowledge is one ingredient for computing word meaning, the missing or implicit ingredient being generalization by appropriate similarity. This explains why it is often sufficient to provide stereotype knowledge in order to convey a word’s meaning to everyone’s satisfaction (what an encyclopaedia does) and why ignorance is often harmless in communication.

References

Bloom, Paul

Devitt, Michael and Kim Sterelny

Dik, Simon C.

Eckardt, Regine

Gärdenfors, Peter

Geeraerts, Dirk

Goodman, Nelson
Kamp, Hans and Uwe Reyle

Keil, Frank
1994  Knowledge and Category Formation. In Lexical acquisition. Lingua 92, special issue, L. Gleitman and B. Landau (eds.).

Kripke, Saul

Landau, B., L. Smith, and S. Jones

Lambalgen, Michiel and Jaap van der Does.
1998  Logic and Cognition. Reader, provided at the European Summer School in Logic, Language and Information ESSLLI ’98, University of Saarbruecken.

Montague, Richard

Musan, Renate

Pelletier, Francis J. and Nicholas Asher

Pessin, Andrew and Sanford Goldberg (eds.)

Putnam, Hilary

Sterelny, Kim

Stern, Gustaf
Meaning change in conceptual Montague semantics

Turner, Jane
1996

Verdauger, Isabel and Anna Poch
1998

Zemach, Ernst
1976
Tense in time: The Greek perfect

Eva-Carin Gerö and Arnim von Stechow

1. Survey

The paper will deal with the diachronic development of the meaning and form of the Greek Perfect. The reason for focusing on this language is two-fold: first, it has often been neglected in the modern linguistic literature about tense; secondly, in Greek, it is possible to observe (even without taking into account the meaning and form of the Perfect in Modern Greek) a very interesting diachronic behavior of the Perfect which may have a great deal to tell about this “tense” more generally.¹

In early literary texts (e.g. Homer) the Perfect, morphologically characterized by reduplication and (often but not always) by a special Perfect-suffix -κ-, has two distinct semantic values: an “intensive” (iterative or simply intensified) meaning and a resultative one. The Perfect form

¹ For the reader not so familiar with the chronology of the Greek language we describe here, in very broad outline, the various periods of its history.
Archaic Greek: 700–500 BC, e.g. Homer, Sappho, and Alcaeus. This stage of the language is highly synthetic and in many respects not as grammaticalized as Classical Greek; the Perfect has a different meaning than that of later stages (especially salient in the so-called intensive use).
Classical Greek: 500–300 BC, eg. Plato, Xenophon, and Sophocles. The language of this period is still very synthetic but more grammaticalized than AG; the Perfect, ex hypothesi, acquires a new meaning.
Hellenistic and Roman-Imperial Greek: 300 BC–600 AD, e.g. Polybius, New Testament, and Lucian. Shows analytic tendencies and, partly, exhibits another pattern of use of e.g. the moods and the two negatives than CG; the core meaning of the Perfect is basically the same as in CG.
Byzantine Greek: 600–1450 AD, e.g. Johannes Malalas and Photius. The literary language is a mixture of “Volkssprache” and a classically influenced idiom; the synthetic Perfect is often replaced by analytic constructions.
Modern Greek: from 1450 AD. Much more analytic than Ancient Greek; the Perfect, as well, is always analytic.
dédorke (from dérkomai ‘see’) thus means ‘gazes’ (= sees in a prolonged or intense way), the Perfect Participle dedorkós ‘flashing (fire from ones eyes)’. The resultative meaning we find e.g. in the Perfect form réthnêka (from thnêiskô ‘die’), ‘has died’ (= ‘is dead’). In Classical Greek the Perfect often yields a resultative meaning (although it can also, as we will see, convey other typical Perfect-shades-of-meaning). During this period we also find Perfect forms with what seems to be a Simple Past/Aorist meaning (e.g. in Pl. Cri.44 a tekmatromai dê ēk tinos enupníou, ho heôraka oligôn próteron taútês tês nuktôs ‘I conclude this from a dream which I had (lit. have seen) somewhat earlier tonight’). In Postclassical Greek the “Simple Past”-use of the Perfect gets to be quite widespread (e.g. in Pol. 3.1.2 ...en têi trítêi bûblôi dedêlôkamen; homoîôs dê kata tás aitías en autêi ekeînêi diesaphêsamen ‘In my third book I explained (lit. have explained)...; I likewise set forth in the same place the reasons...’). Gradually, (in nonartificial Byzantine Greek), the synthetic Perfect disappears and is replaced by analytic forms.

As in the analysis of many other languages, the correct interpretation of the Greek Perfect is a controversial matter. As far as Classical Greek is concerned (the language period most focused on in traditional Greek grammar), two important types of interpretation can be distinguished, viz. (i) the theories which focus on resultativity, and (ii) the Reichenbachian framework (following Reichenbach 1947) the meaning of the various tenses (e.g., in English) is described by means of the parameters E, R, and S. E is the event time (the point or interval at which an event takes place); S is the utterance time, and R the reference time. The possible relations between these various “temporal entities” are the following: S<R (“speech time before reference time”): Future; R<S (“reference time before speech time”): Past; S,R (“speech time and reference time identical”): Present; E<R (“event time before reference time”): Perfect; R<E (“reference time before event time”): Prospective. The various tenses, then, are the result of a combination of specific relations. E.g., the representation of the Present is the result of the combination S,R with E,R to yield S,R,E, and the representation of the Present Perfect is the result of the combination S,R with E<R, to yield E<R,S.
Tense in time: The Greek perfect

253

theories. Representatives of the resultative “school” are Chantraine (1927), Humbert (1960) and McKay (1980, 1981). The Reichenbachian approach we find in modern linguistic commentaries on the Greek Perfect (e.g., Mourelatos 1978), as well as in newer school and university grammars of the Greek language, e.g. in Blomqvist and Jastrup (1996).

Depending on what views the theoreticians on Greek grammar have concerning the core meaning of the (Classical) Perfect, their theories about the diachronic development of this “tense” vary a great deal. Focusing on resultativity, quite a number of scholars of traditional Greek grammar (e.g., Chantraine 1927) have reconstructed the following history of the Perfect: in Homeric Greek the Perfect, which had an “intensive” or (usually) resultative meaning, was almost always intransitive4; in Classical Greek the resultative Perfect was distinctly extended to transitivity and therefore came to be more past-oriented (an object as a result of an event leads the hearers’ thought to the event in the past which caused the result (“I have born a child”) in a different way than the old result-on-the-subject (“I have born = I am a mother”); the synthetic Perfect then, in Postclassical Greek, gradually acquires a Past meaning (but probably not quite the same meaning as the Simple Past/Aorist), gets confused with the Aorist (also morphologically), and comes to be extinct. The Reichenbachian approach to (middle and later stages of) the history of the Greek Perfect is the following: since the Perfect of Classical Greek can be seen as involving two reference-points, Present and Past (like S,R & E< R in Reichenbach 1947), the Perfect is present-oriented and past-oriented. Because of the past-orientation the Perfect can also be used as a retrospective tense without present-orientation, which is increasingly true of later Greek, where it is used with a Simple Past-meaning. (Note that the first type of theory seems to suggest that there was no real change of meaning – from Perfect into Simple Past – between the Classical and the Postclassical period, whereas the Reichenbachian type implies that there was such a change.)

After having looked at a great deal of Greek data, as well as having considered the applicability on Greek of traditional as well as various modern linguistic approaches to the Perfect in other languages, we tend to think that an Extended-Now-analysis, suggested, e.g., for the English Perfect by McCoard (1978) and Stechow (2001a) is to be preferred to a resultative

4 Exceptions to this rule are Hom. Il. 2. 272 ἐ δὲ μάρι’ Ὀδυσσεὺς ἐσθλὰ ἐόργη (‘For sure, Ulysses has accomplished innumerable good deeds’) and Od. 17. 284 κακὰ πολλὰ πέπονθα (‘I have suffered many evils’).
mode of explanation (whether traditional, as in Chantraine or McKay, or modern like Kamp and Reyle 1993). The Reichenbachian account of the Perfect sometimes adopted in modern treatments of the Greek tenses is, at most, adequate as a means of analysis of the usage at certain stages of development in this language, and it fails, in any real sense, to give insight into the mechanisms behind the development per se. The Extended-Now-approach, as we shall see, covers much more ground. To our knowledge, no other thorough attempt has been made to interpret the Ancient Greek Perfect in this way, either synchronically or diachronically. The “intensive” use of the Perfect we will almost entirely exclude from our investigation.

In the account of the diachrony of the Greek Perfect something will also be said about the early emergence of analytic Perfect constructions in this language, and of the fate of the synthetic vs. analytic Perfect in MG.

---

5 According to the Extended-Now-account the Present Perfect denotes an interval which includes the event time and extends up to the speech time. The Extended Now is then (e.g. by Dowty 1979) treated as a predicate XN, with the following truth conditions: XN(t) is true at a time t’ if and only if t’ is a final subinterval denoted by t. For a similar approach, cf. Iatridou et al. 2001: “...the Perfect sets up a time span. The right boundary of that time span is oriented towards a contextually relevant point in time (not necessarily by identity, but also by approximation, as with lately)...The left boundary of the Perfect time span depends on the lexical meaning of the adverbial. When the Perfect adverbial is since, the left boundary is the overt argument of since (e.g. since 1991)”. For a more explicit analysis, cf. Stechow (1999b).

6 It deserves, however, to be mentioned that the famous American pioneer of Greek and Latin grammar, Basil Gildersleeve, who always had very sound intuitions about the languages he interpreted, to all appearances looked upon the Perfect in some such way, cf. Gildersleeve (1900–1911: 99): “The perfect looks at both ends of an action. The time between these ends is considered as a present. When one end is considered, the present is used; when the other, the aorist”. In his treatment of verbal aspect in New Testament Greek, Fanning (1990) briefly discusses the Extended-Now-theory, as found in McCoard (1978). He rejects it for NT Greek, however (with implied claims of nonvalidity also for Ancient Greek generally) in favor of a combined Tense/Aspect/Aktionsart-approach. As for the synchronic interpretation of the MG Perfect, an Extended-Now-approach has been proposed recently by Iatridou et al. (2001).
2. On the tense/aspect/Aktionsart- architecture

The aim of our investigation is an analysis of the system in syntactic and semantic terms. We will use an approach in the style of recent syntax/semantics (cf. Chomsky 1995; Stechow 1993, 1996; Giorgi and Pianesi 1998, Stechow 1999a). The main characteristic of such systems is the assumption of functional projections for Agreement, Tense, and Aspect. One of the great advantages of this approach is its ability to provide a transparent and simple analysis of the meaning of Tense and Aspect and to combine this with the semantics of adverbs and negation. We will see that this framework enables us to describe linguistic variation through time and languages within a stable underlying “universal grammar”.

The syntactic analysis proposed in the following is rather abstract. In particular, we do not commit ourselves to details of Greek surface syntax. Our syntactic structures abstract from linear order and scrambling. To give an idea of the difficulties and the required abstractness of the analysis, we ask the reader to consider the following example:

(1) \( tòn d’ aútê prosēipe períphôn Eurúkleia \)
    Him then answer\textsubscript{AOR. wise} Eurycleia
    \( (\text{Hom. Od. 20.134}) \)

To our mind, this sentence expresses the statement that the PAST spoken about contains an event which is an answering done by Eurycleia and directed towards Ulysses. In somewhat more formal terms, this reading can be expressed as:

(2) \( \exists e [e \subseteq \text{PAST} & e \text{ is the giving of an answer to him by Eurycleia}] \)

PAST is the semantic past, i.e., a time before the speech time specified by the context of use. The information \( e \subseteq t \) is the meaning of the perfective aspect (“PERFECTIVE”). The AG Aorist morphology contains the information \( \text{PAST} + \text{PERFECTIVE} \), which is spread to different functional nodes in the syntax. A syntactic tree for the sentence could be this:
The semantic tense PAST and the semantic aspect PERFECTIVE are located in the specifier of the tense phrase (TP) and in that of the aspect phrase (AspP) respectively. This information is abstract, i.e., phonetically invisible. Semantically, the finite Aorist is a combination of PAST + PERFECTIVE. Morphologically, the PERFECTIVE is associated with the stem, for we have an infinitive Aorist in AG, which has no PAST-meaning. The information PAST comes with the augment, which is traditionally regarded as its carrier (but which may be zero in Homeric Greek). The situation is complicated somewhat by the fact that the finite Aorist form has special personal endings different those of other nonpast indicative tenses. In any case, it is safe to say that the finite Aorist form has the morphological features [past] and [perfective], which are located in the heads of TP and AspP, respectively. The finite verb may check these by successive movement. It ascends gradually to the past-position. The pronoun tôn, ‘him’, is scrambled to the “Wackernagel position”, i.e. to the position in front of TP. This gives us the observed surface word order.
We will see that the meaning (2) can be calculated from this tree by functional application of PFV to the VP-meaning and by functional application of PAST to PFV VP. We will now explain our system in more detail.

The essential idea of the analysis is that we carefully distinguish between morphological feature and semantics. A verb has certain morphological features, which check the presence of the functional projections containing the respective features as heads. The specifier of a functional projection contains semantic information, which matches the morphology in the unmarked case, but by no means always does so. For instance, the German perfect morphology is ambiguous in many ways: it may express a semantic Past tense, an Extended-Now-Perfect, a Future Perfect, and a Perfect of result:

(4) a. *Gestern sind wir in Rom gewesen.
   yesterday are we in Rome been

b. Ich habe seit zwei Stunden auf dich gewartet.
   I have since two hours for you waited

   tomorrow have I the paper delivered

d. Eva hat Tübingen seit einem Jahr verlassen.
   Eva has Tübingen left since one year

The Italian passato prossimo, which is morphologically equivalent to the German Perfect, has a much more restricted meaning, namely that of the AG aorist. For instance, (5a) is unacceptable. We have to use the imperfetto instead of the passato remoto, if we want to get an idiomatic text. The examples are taken from (Giorgi and Pianesi 2001).

(5) a. *(Alle tre) Mario ha mangiato una mela e la sta mangiando tutt’ora. [22]
   (At three) Mario has eaten an apple and he is still eating it

b. (Alle tre) Mario mangiava una mela e la sta mangiando
tutt’ora.
   (At three) Mario ate (ipfv) an apple and he is still eating it

\[\text{Cf. Stechow (2001b).}\]
Obviously, we could produce exactly the same contrast for the Aorist vs. Imperfect for Ancient Greek. Below we give an exposition of the tense/aspect/Aktionsart-architecture which we have in mind. As has been said above, we will strictly distinguish between morphological and semantic tense and aspect. The semantic notions will be represented by capital letters, the morphological ones by small letters. This is the architecture of the finite clause:

(6) The tense/aspect/Aktionsart-architecture

```
TP
  SEMANTIC TENSE
    morphological tense
    T'
      AspP
        SEMANTIC ASPECT
          morphological aspect
          Asp'
            VP
              AKTIONSART
```

The standard semantic tenses are PRESENT, PAST, and FUTURE, but there may be some additional ones as well. The morphological realization of the tenses can vary from language to language. A complication may arise from the fact that TENSE and ASPECT are not always realized by separate morphemes. A stable factor in Greek morphology, on the other hand, is that the Aorist morphology of finite forms (the Aorist marking proper together with the so-called augment) expresses PAST + PERFECTIVE, whereas the imperfect morphology expresses PAST + IMPERFECTIVE.

The notion of aspect is two-faced, which easily leads to confusion. The semantic side of the aspect morphology is complicated by the fact that we have to distinguish between two different notional categories, viz., asp·ctual class and asp·ctual relation. In the semantic literature asp·ctual class usually is a synonym for Aktionsart in the sense of Vendler (1957). In other...
words, an Aktionsart is an accomplishment/achievement, activity, or state.\(^8\) The Vendlerian Aktionsart is expressed by the tense- and aspectless VP. We will assume that an accomplishment/achievement is a class of events, whereas Vendlerian states are sets of times or states. Following Herweg (1990) and Katz (1995), we will not distinguish between times and states.

An event or state is located in time by means of a relation that connects the reference time with the event time or the event state. With Klein (1994), we will call these relations ASPECTS, more accurately, SEMANTIC ASPECTS. The following ASPECTS are used in many languages: (i) the reference time may INCLUDE the event time or the state time, (ii) the reference time may be INCLUDED in the event time, or (iii) the reference time may follow the event time or state time, being POST. One and the same morphological aspect may license different ASPECTS. For instance, Russian perfective morphology selects Vendlerian accomplishments/achievements and licenses either INCLUDES or POST.

Let us make this more precise. Recently, a version of Partee’s (1973) deictic theory of TENSE has become increasingly popular. We follow the proposal given in Heim (1994), according to which tenses restrict the interpretation of temporal variables:

\[7\text{ Semantic tenses}\]

are symbols of type \(i\) which bear time variables as indices. Let \(c\) be the context of the utterance with \(t_c\) as the speech time.

a. \(\|\text{NOW}\|^{c}\) is the speech time conceived as a point.

b. \(\|\text{PAST}\_j\|^{c}\) is defined only if \(g(j)\) precedes the speech time \(t_c\). If defined, \(\|\text{PAST}\_j\|^{c} = g(j)\).

c. \(\|\text{FUTR}\_j\|^{c}\) is defined only if \(g(j)\) follows the speech time \(t_c\). If defined, \(\|\text{FUTR}\_j\|^{c} = g(j)\).

\(^8\) In Slavistic philology, the term Aktionsart is used a bit differently. In this tradition, Aktionsarten are properties of actions such as inchoativity, terminativity, frequentativity, and what other characteristics can be expressed by affixes modifying the verbal root. The Vendlerian Aktionsarten are classes of events defined by their temporal properties. The term does not touch upon the problem of how the VP expressing such an Aktionsart is built up by means of internal Aktionsarten in the sense used by Slavistic philologists.
We will use the terminology introduced by Reichenbach (1947) and call the
time denoted by a semantic tense the \textit{reference time}.\footnote{It is what Bäuerle (1979) and Fabricius-Hansen (1986) call \textit{Betrachtzeit}, what Klein (1994) calls \textit{topic time or time of the claim}, and what Musan (2000) calls \textit{tense time}. I see little point in contributing to the general confusion by introducing a new terminology; we accept any term, provided it has a clear interpretation.} Below is a list of the \textbf{SEMANTIC ASPECTS} mentioned.

(8) \textbf{Semantic aspects}

a. INCLUDES $= \lambda P \lambda t \exists e. \tau(e) \subseteq t \ & P(e), \ P \text{ of type } <v,t>$ \footnote{$i$ is the type of times, $v$ that of events.} \text{ ("PERFECTIVE")}

b. POST $= \lambda P \lambda t \exists e. \tau(e) < t \ & P(e)$ \text{ ("PERFECT")}

c. INCLUDED $= \lambda P \lambda t \exists e. \tau(e) \subseteq \tau(e) \ & P(e)$ \text{ ("IMPERFECTIVE")}

e can be an event or a time. \(\tau(e)\) is the time of \(e\). If \(e\) is a time, then \(\tau(e) = e\).

In order to keep \textbf{SEMANTIC ASPECT} distinct from morphological aspect, it would be safe to use the technical terms INCLUDES, POST and INCLUDED throughout the paper. But after Klein (1994) the practice has spread of using the names PERFECTIVE, PERFECT and IMPERFECTIVE for these relations (cf., e.g., Kratzer 1998). We will adhere to this usage, but, at the same time, we want to formulate a serious warning: One has to understand these terms in their technical senses, which are given by their definitions. One should avoid confusing \textbf{SEMANTIC ASPECT} with morphological aspect. Henceforth, we will use the abbreviations IPFV for IMPERFECTIVE, PFV for PERFECTIVE, while PERF will stand for PERFECT.

This said, let us evaluate in detail the LF conveyed by tree (3). We will not analyze the VP in detail, but we will assume the following meaning:

\begin{equation}
\| [\text{VP periphrôn Eurúcleia tôn} \ \text{proséeípe}] \|_{\mathcal{F}} = \\
\lambda e. e \text{ is the giving of an answer to } g(j) \text{ by wise Eurycleia}
\end{equation}
The only important thing to keep in mind is that the VP is semantically tense- and aspectless, i.e., it is interpreted exactly as if it were a non-finite construction. The semantic tense + aspect, i.e., PAST + PERFECTIVE, is located in distant functional nodes. The Vendlerian Aktionsart is not altogether clear. The VP is presumably best interpreted as an accomplishment; it could be an activity as well, but not a state.

The tree is evaluated by ignoring the morphological nodes and by performing type driven functional application (FA), the way taught by standard textbooks, e.g., Heim and Kratzer (1998).

\[
\lambda \exists e \in \tau(e) \subseteq t \land \lambda e[\text{e is the giving of an answer to } g(j) \text{ by wise Eurycleia}] \]

meaning of the VP

\[
\lambda - \text{conversion}
\]

\[
\exists e[\tau(e) \subseteq g(i) \land \lambda e[\text{e is the giving of an answer to } g(j) \text{ by wise Eurycleia}]]
\]

meaning of the VP

\[
\lambda - \text{conversion}
\]

3. Chronology of tense/aspect systems

In what follows we give a survey of some important historical stages of the Greek tense/aspect systems. The representations will associate different verbal forms with the functional nodes checked by the form. The functional nodes are associated with their morphological features and their semantic interpretations. As before, the morphological features are written in small letters, whereas capital letters are used for the semantics. It should be obvi-
ous that the survey is highly idealized and incomplete, as any attempt to introduce order into the actual historical development is bound to be.

3.1. Archaic Greek (700 – 500 BC)

The following charts will not distinguish between heads and specifiers of the functional categories. Furthermore, no account will be given of the person features.

<table>
<thead>
<tr>
<th>Form</th>
<th>Tense</th>
<th>Aspect</th>
<th>root-meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>márptô</td>
<td>Present</td>
<td>neutral</td>
<td>marpt-&quot;seize&quot;</td>
</tr>
<tr>
<td>márpsô</td>
<td>Future</td>
<td>neutral</td>
<td>&quot;</td>
</tr>
<tr>
<td>émarpton</td>
<td>Past</td>
<td>ipfv</td>
<td>&quot;</td>
</tr>
<tr>
<td>émarpsa</td>
<td>Past</td>
<td>pfv</td>
<td>&quot;</td>
</tr>
<tr>
<td>mémarpa</td>
<td>Perfect</td>
<td>ipfv</td>
<td>INTENS or RESULT + &quot;seize&quot;</td>
</tr>
<tr>
<td>(e)memárpea</td>
<td>Pluperfect</td>
<td>ipfv</td>
<td>INTENS or RESULT + &quot;seize&quot;</td>
</tr>
</tbody>
</table>

A look at this chart reveals an asymmetry of the system. While we find perfective/imperfective opposition for the past tenses, we do not find it for the future. Modern Greek has introduced an “Ausgleich” here. The present form not being aspectually marked (i.e., its being “neutral”) does not come as a surprise. The semantic PERFECTIVE cannot be combined with PRES, which is conceived as a point of time and must therefore be combined with IMPERFECTIVE.

To all appearances, the Imperfect and the Aorist Indicative are genuine Past tenses that are distinguished by aspect\(^{11}\). The Aorist selects the PERFECTIVE, while the Imperfect selects the IMPERFECTIVE.

---

\(^{11}\) This is indeed the communis opinio found in the traditional descriptions of the Greek tense and aspect system. (E.g., Bizos 1981, Humbert 1960, Kühner and Gerth 1898). It should also be mentioned, however that some recent theoreticians favor another view, viz., that of the Aorist (also the Indicative) not being a Past tense, but basically a tenseless form, which only expresses perfectivity (cf. McKay 1988, Rijksbaron 1988).
Tense in time: The Greek perfect

PERFECTIVE seems to be the same notional category as in the Slavic languages, i.e., it expresses completion/telicity of the event type denoted by the VP. This follows from the semantics of PERFECTIVE, which requires that the VP-event be included in the reference time. This is only possible for completed events.

The Perfect and the Pluperfect are the most complex forms. The Perfect (unlike the Pluperfect) does not contain a semantic PAST. It is a presentic form expressing either intensity/iteration or resultativity. The traditional cover term for this usage is “Pure Perfect”. We have retained this terminology, as a convenient accommodation to the conventions of traditional Greek grammar. Semantically, the term covers two completely different operations, namely INTENSivitY and RESULTativity. There have been attempts in traditional Greek grammar to reduce Intensity to Resultativity (as reflected, e.g., in Smyth 1956: 435), but we do not believe in the possibility of such a reduction. We will not speculate on the question why the two meanings are encoded in the same form. As said already, we will neglect the intensive meaning.

Before we consider the Perfect, let us make some remarks about the interpretation of imperfective forms. Imperfective morphology requires IPFV. As a consequence, the embedded Vendlerian Aktionsart must have the subinterval property. If the VP expresses a state or an activity, this raises no problems. But if it is an accomplishment or achievement, we must stativize it by means of semantic operations such as the Progressive, Habituality, Iterativity, or Modality (e.g., Possibility). These operations may be morphologically visible or not. Below is a first approximation to the meaning of a habituality and an iterativity operator:

(10) **Two temporal stativizers**

a. Habilituality.

   \[ \text{HAB} := \lambda P \lambda t \exists \tau [t \subseteq \tau \& \text{there are many } e: e \subseteq t & P(e)] \]

b. Iterativity.

   \[ \text{ITER} := \lambda P \lambda t [\text{There are many } e: e \subseteq t & P(e)] \]

Here are some examples illustrating the application of these stativizers.

(11) \textit{autârko doûre dúô kekoruthména chalkoi pallôn Argeîôn prokalîzeto}

and spears two tipped bronze\textsubscript{Dat} brandishing of Argives challenge\textsubscript{ipfv}

\[ \text{pántas aristous antibion máchesthai en ainêi dêiotêti} \]
all the best face-to-face to fight in dread combat
“and brandishing two spears tipped with bronze he challenged
(‘went about and challenged’) all the best of Argives to fight with
him face to face in dread combat”
(Hom. Il. 3. 18)

“He challenged the best” is an accomplishment. We have to convert this
Aktionsart into a state in order to combine it with IPFV. The operation
stativizing the VP is presumably iterativity. This leads to the following
analysis.

(12) PAST past IMPFV ipfv [vp ITER [vp pro prokalízeto arístous]]

We are using “pro” as a representation for the missing subject pronoun. The reader may check for him-/herself that this analysis provides a reason-
able interpretation: PAST is contained in an interval which itself contains
an iteration of challenges.

The accomplishment in the following example is presumably best stativ-
ized by means of the HAB-operator.

(13) pàr dè zôstêr keíto panaíolos, hôi r’ ho geraiòs zônnut’h,
beside and girdle lieIMPF flashing whichDat. the old man girdIPFV.med/pass.
Hôt’ es pólemon phthisénona thôrēssóito
whenever to war baneful arraypres.Opt.med/pass.
“and by his side lay the flashing girdle, wherewith the old man
used to gird himself, whenever he arrayed himself for battle, the
bane of men”
(ibid. 10.78)

The logical form of the sentence is therefore something like this:

(14) [TP PAST [Ag IPFV HAB [vp ho geraiòs zônnto ] ] ]

The sentence means that the reference time is a subinterval of the period in
which the old man had the habit to gird himself with the flashing girdle.

The introduction of the modal stativizers Progressivity (PROG) and
Possibility (CAN) requires a modal framework, which we cannot seriously
develop in this paper. The semantics for the Progressive will follow Dowty
(1979) and Landman (1992), which means that PROG is treated as a modal
operator conveying the meaning that the event in question is completed in the normal course of events, in the “inertia worlds”.\(^{12}\) CAN means that the proposition modified is compatible with a modal background, which is a proposition specified by the context (Kratzer 1977).

(15) **Two modal stativisers**

a. \(\text{PROG} := \lambda P \lambda t \lambda w \forall w'[\text{Inert}(t)(w') \rightarrow \text{There is a superinterval } t' \text{ of } t \text{ such that } t \text{ is not a final part of } t' \& \exists e \subseteq t' \& P(e)(w')]\)

b. \(\text{CAN} := \lambda P \lambda t \lambda w[H_{w_t} \text{ is compatible with } P], \) where H is a modal background

Inert\((t)(w')(w')\) means that the world history \(w'\) is the same as \(w\) until the end of \(t\), and that \((w', t')\) makes every proposition true which should be true in \(w\) at \(t'\) in the normal course of events, where \(t'\) is the said superinterval of \(t\). A modal background, used in the definition of modalities like CAN, is a function which assigns a proposition to a world and a time.\(^{13}\)

The following data present a case where the embedded accomplishment may be stativized by means of the PROG-operator.

(16) *kai sphìn áchos katà thumòn egígneto* derkoménoi Si Tròas and them\(\textit{Dat.}\) sorrow in heart become\(\textit{ipfv}\) beholding\(\textit{Dat.}\) Trojans\(\textit{Acc}\) “and sorrow grew (‘was growing’) in their hearts, as they beheld the Trojans”

(Hom. *Il.* 13.86)

Consequently, the LF of the sentence may be something like the following:

(17) *egígneto áchos sphìn katà thumón*

\([TP \text{PAST} \text{[Ans]} \text{IPFV} \text{[VP \text{PROG} \text{[VP \text{BECOME} sorrow them in the heart]]]}]}\)

\(^{12}\) For this term, see Dowty (1979: 148ff.)

\(^{13}\) This semantics is simplified because it neglects inconsistent backgrounds. We assume that they are treated along the lines indicated in Kratzer (1978) or Kratzer (1981).
Note that we do not classify PROG as a semantic aspect. It is a stativizer, which belongs to the VP, not to the AspP. As to the semantics of the BECOME-operator, the reader should consult Dowty (1979: chapter 3). This predicate is an “eventizer” which converts a state into an accomplishment or achievement. The predicate tells us that the state does not obtain at the beginning of the event, but only at the end of the interval (and is not defined in between).

Finally, the reader is asked to consider modalization by the CAN-operator, which requires the presence of a negation in the syntax (Cf. Chantraine 1953). In the following example it is the negated modality/possibility which converts the embedded accomplishment into a state:

(18) áll' oúd' hős toû thumôn enî stêthessin ēpeithon, prîn g’ hôtê dê
but also not so his heart persuadeipfv.3plur. before
thálamos pûk’ ebálleto
chamber much batteripfv.3.Sing.Pass
“yet not even so could they persuade the heart in his breast, until
at last his chamber was being hotly battered”
(II. 9.587)

The position of the negation is not entirely clear to us. Let us assume that it is located under the semantic aspect. In this case, the LF of the example could be as follows:

(19) [TP PAST [Asp IPFV oûd’ [VP CAN ēpeithon toû thumôn ] ] ]
PAST [IPFV NEG [VP CAN they persuade his heart]]

This means that the PAST is contained in a stretch of time during which they (i.e., the people around Meleager) where not able to persuade him (i.e., Meleager himself).

Finally, let us consider the Perfect. During this period, the Present Perfect always expresses a presentic statement. It follows that the semantic aspect IPFV is unmarked in the morphology. There are two meanings con-

---

14 Interestingly, exactly this configuration is exhibited by Russian and other Slavic languages (Schoorlemmer 1995: chapter 4).
Tense in time: The Greek perfect

267

connected with this form, viz. the Intensive and Resultativity. No PAST meaning is witnessed to our best knowledge.15

As has already been mentioned above, the Greek Perfect may express Intensity. The following example illustrates this kind of Perfect.

(20) máchê pôlemós te dédêe
‘Fight and war are raging (intensively)’

Here the active Perfect form dédêe (from daíô ‘light up’, ‘kindle’) has the intransitive meaning of “burn”, “rage”, otherwise connected with the passive forms of this verb.

In Homeric Greek this use of the Perfect is very frequent with verbs whose meaning involve a) sounds, e.g. bébruche ‘he bellows’, bebruchós ‘bellowing’ (Pres. brucháomai), kékraga ‘I scream’ (Pres. krázô), b) activities of the senses (or feelings), e.g. ópôpa ‘I see’ (Pres. horáô), ódôde ‘smells’ (Pres. ózô), géérêthe ‘is delighted’ (Pres. géthô), d) gestures/facial expressions, etc., e.g. pephríktôs ‘shuddering’ (Pres. phríssô)16.

Intensifiers are something like adverbs of degree or manner adverbs. They measure distance from normality. The INTENS-operator means approximately “very much”. When the activity denoted by the verb has a component that can be escalated, the meaning of the Intensive will be “to a high extent”; when there is no such component (or sometimes even when there is one), the meaning of the Intensive will, so to speak, enter another dimension, yielding readings like “rapidly” or “repeatedly”. As is to be expected, these meanings are stativizers.

We will make no attempt to analyze this use any further. Let us only remark that there is no plausible connection between the intensive use of the perfect and the simultaneously existent resultative use to which we turn next.

15 Examples like Il. 2.272 and Od. 17.284 (cf. above, page 253, note 4) are sometimes believed to represent an intermediary state between the Present Resultative and the more Past-oriented Experiential use. Cf. for instance Hedin (1999).
16 The meaning of the Present forms is usually that of the Perfect form minus the intensive “grade”. A good example is also bébêka ‘be in the act or attitude of going’ (Pres. baínô ‘walk’), as in Od. 1. 360 hê mên thambêsasa pálin oikónde bébêkei (“but she (sc. Penelope) was astounded and took her way back to her rooms”).
The prevailing use of the Perfect is that of a RESULT-operator. This use has to be assumed in the following example:

\[(21) \quad \text{êôs dé moí estin êdê duodekâtê hôt’ es Ilion eilêloutha} \]
\[\text{morning and I.Dat. be.PRES. now twelfth when to Troy come.PERF.1.Sing.} \]
\[\text{“it is now the twelfth morning of my being in (lit. having come to) Troy”} \]
\[\text{(Hom. II 21.81)} \]

RESULT is a stativizer which gives us the resultant state of an accomplishment or achievement, exactly in the same way as in the case of the English adjectival passive or the German “Zustandspassiv”. In German, we can translate the example into a presentic statement:

\[(22) \quad \text{Es ist der zwölfte Morgen, seit ich nach Troja gekommen bin.} \]
\[\text{it is the twelfth morning since I to Troy come am} \]

In Archaic Greek the resultative use is usually connected to intransitivity. A closer inspection would reveal that these verbs are typically those that have been called “unaccusatives” in the generative literature. There are quite a number of theories about the Perfect of result. We will simply adopt the proposal made in Kratzer (2000). Kratzer assumes that the perfect of result can be formed by verbs which have an accessible resultant state in the lexical semantics. For instance, the verb “x comes to Troy” may be analyzed as:

\[(23) \quad \text{A transformative verbal} \]
\[\text{come to Troy} := \lambda x \lambda e \lambda s [\text{cause(s)(e) & in-Troy(y)(s)}] \]

Applied to a subject x, this verb x is true of an event and a resultant state s, if e causes s and s is the state that x is in Troy. The RESULT-operator takes such a verb; furthermore, it existentially quantifies the generating event and gives us the resultant state. Consequently, its semantics is this:

\[(24) \quad \text{The RESULT-operator} \quad (\text{Kratzer 2000}) \]
\[\text{RESULT} = \lambda R \lambda s \exists e [R(e)(s)] \]
The RESULT-operator is not a stativizer in the sense that it converts an accomplishment/achievement into a state. It rather chooses the state as an option. This makes it an “Aktionsart chooser”. In consequence, the logical representation of the tenseless perfect form “I have come to Troy” is:

\[
(25) \quad [VP \ \text{RESULT} \ I \ \text{come to Troy}]
\]
\[
= \lambda s. \text{cause}(s)(e) \ & \ \text{in-Troy}(I)(s)
\]

This form may duly be combined with PRES + IPFV. So the LF for our example is this:

\[
(26) \quad [TP \ \text{PRES} \ [\text{AspP} \ IPFV \ \text{RESULT} \ [\text{pro} \ \text{Ilion} \ \epsilonι\ellουθα]]]
\]

This structure means that the PRESENT is contained in a state s such that I am in Troy at s and s is caused by a previous event. The “pluperfect” has exactly the same interpretation with PAST instead of PRES.

We have to add that the logical type of this sort of unaccusative requires an extra logical operation for ordinary active statement. We need an “eventizer”, which is the mirror image of the RESULT-operator. For instance, the sentence “I came to Troy” must have the following LF:

\[
(27) \quad \text{PAST} \ \text{PFV} \ \text{EVENT} \ I \ \text{come to Troy}
\]

The semantics of the “eventizer” is given by the following rule:

\[
(28) \quad \text{An eventizer}
\]
\[
\text{EVENT} = \lambda R \lambda e \exists s[R(e)(s)]
\]

The evaluation of the precise relation of this theory of the resultative Perfect to traditional analyses (e.g., Chantraine 1927) we leave to future research.

One wonders by what historical coincidence two very different meanings, Intensity and Resultativity, were associated with the same morphology. We will see that the double meaning of the form does not survive in history. Traditionally, it has been assumed that intensity had to give way to resultativity. As we will see, however, not all Posthomerics uses of the Perfect are resultative: in Classical and Postclassical Greek we find all the readings of the Perfect commonly associated with, e.g., the English Perfect (U-reading, the various E-readings, see explications p. 272).
3.2. Classical (500 – 300 BC)

In traditional descriptions the TA-system of Classical Greek is said to be just like that of Homeric/Archaic Greek, except for the Perfect, which in Classical Greek is supposed to have resultative meaning or occasionally a Past meaning, while the intensive use is disappearing\(^\text{17}\). The grammarians also list several periphrastic constructions, which they describe as having a particularly strong resultative meaning (\textit{eimí} ‘be’ with Perfect Participle, \textit{échô} ‘have’ with Aorist or Perfect Participle\(^\text{18}\)). The semantics of these constructions in Classical and Postclassical Greek will not be analyzed in this paper.

As for the diachronics of the synthetic Perfect, the following functional explanation has suggested itself to the traditional grammarians: the Resultative has a “posteriority” component, which the Intensive has not. The Intensive is a sort of Present Progressive and points toward the near future, the end of the action. Classical Greek would consequently specialize towards the “posteriority” perspective.

In the new analysis which we are going to suggest, we will start our completion of the traditional picture by the observation that other uses of the Perfect appear already in the classical period, where the Perfect shows all the readings that are typical for e.g. the English Perfect.

In archaic Greek, the resultative use is almost exclusively combined with intransitivity; in the classical period the meaning is generalized to

\(^{17}\) The Intensive Perfect during the Classical period is indeed scarce and has turned into a kind of “formula”. We do not agree with some commentators of Greek tragedies who seem to think that the use of the Intensivum was still really productive at this time (cf., e.g., Kannicht’s commentary to Euripides’ \textit{Helen} [p. 155]).

\(^{18}\) An example of the periphrasis with \textit{eimí} is Ar. \textit{Pl.} 867 \textit{ho dè polù màllon eniuous estin exolôlekôs} (‘Some of us truly he has brought to ruin rather than to wealth’); the construction \textit{échô} with Aorist Participle we find e.g. in Hdt. 6. 12 \textit{...andri Phôkaei alazônì...epitrëpsantes hêméas autoìs échomen} (‘...we have committed ourselves into the hands of this Phocaean braggart...’); an example of \textit{échô} with the Perfect Participle is Pl. \textit{Theaet.} 200 a \textit{epistêmên ára oiêsetai tethëreukôs échein, all’ ouk anepistêmosûnên} (‘Thus he will think that he has caught (and has) knowledge, not ignorance’). Note, however, that the periphrasis with \textit{eimí} is often used also without special emphasis on resultativity, so regularly in the third person plural, as well as in the optative and subjunctive moods of the Perfect.
transitivity\textsuperscript{19}. In the examples below (29) illustrates the frequent resultative meaning of one transitive Perfect form (\textit{apoleloípasin} ‘have left (us)’ of two intransitive ones (\textit{apodedrákasin} ‘have gotten away’ and \textit{apopepheúgasin} ‘have escaped’)), (30) the almost extinct “intensive” use, and (31) the new “Past” meaning (cf. above, p. 252).

(29) \textbf{apoleloípasin hêmâs Xenías kai Pasíôn.} all’ eû ge méntoi epistásthôn,
    leave\textsubscript{perfect.3.pl} us Xenias and Pasion but well however know\textsubscript{imper.3.plur.}
\textit{hóti oûte apodedrákasin - oûda gár, hópei oîchontai - oûte}
that neither get away\textsubscript{perfect.3.pl}, know\textsubscript{1sing.}; for where be\textsubscript{3.pl.}, neither
\textit{apopepheúgasin -échô gár triêreis hôste heleîn tò ekeînôn ploion}
e scape\textsubscript{perfect.3.pl}; have\textsubscript{pres.1.Sing.} for triremes so that take\textsubscript{inf.} the their ship
“X. and P. have deserted us. But they ought to know that neither
\textit{have they gotten away} because I know in what direction they have
gone - nor \textit{have they escaped} - because I have triremes\textsuperscript{20} so that I can
overtake their ship”
(Xen. \textit{An.} 1. 4. 8)

(30) \textbf{oûd’ éti pûr epibômion en Troíai theoîsin lêlampen}
and no longer fire on-the-altars in Troy to the gods shine\textsubscript{perfect.3.sing.}
\textit{kapnôi thuôdei smoke\textsubscript{Dat.3Sing} fragrant}
“\textbf{No more} on the altars of the Immortals \textbf{is the old fire ablaze}, fragrant with incense”
(Eur. \textit{Andr.} 1025)

\textsuperscript{19} Some theoreticians writing on the Greek Perfect have focused heavily on the
distinction intransitive/transitive as being of crucial importance for the (temporal/aspectual) semantic development of this “tense”.
\textsuperscript{20} I.e., bigger and faster ships that of Xenias and Pasion.
The innovation is illustrated by the last example (31). The employment of the adverbial τότε (‘then’) suggests that this is a PAST-meaning of the Perfect. In Archaic Greek, when the Perfect usually had a presentic value (intensive or resultative) the combination with past-oriented adverbs was not possible (more precisely: is not attested).

We have already mentioned that the proponents of traditional Greek grammar look upon the resultative use of the Perfect as more or less the only meaning of this tense in Posthomeric, classical Greek (cf. e.g., Rijksbaron 1984, Smyth and Messing 1956). To be sure, the Greek Perfect during the classical period often yields a resultative reading. After closer inspection of the classical texts, however, we find practically all the different readings that are commonly associated with e.g. the English Perfect.

In the linguistic literature four major uses of the Present perfect have been identified (cf. for instance Dowty 1979, McCoard 1978). The U(niversal) Perfect conveys the meaning that the predicate holds throughout some interval stretching from a certain point in the past up to the present (“John has been sick since 1990” on one reading of the sentence); the Experiential Perfect asserts that the subject of the sentence has a certain experience (“I have been to Athens five times”); the Resultative Perfect expresses that the effect of an underlying event still holds (“I have lost my glasses”); the Perfect of Recent Past is employed, when one wants to report an event which just happened (“John has just graduated from college”).

21 “The perfect denotes a completed action the effects of which still continues in the present” (Smyth and Messing 1956); “the primary perfect indicative (commonly: perfect) signifies that the state resulting from the completion of the verbal action exists at the moment of utterance (the ‘present’)” (Rijksbaron 1984). In most grammars there is also mention of the “generic” Perfect (cf. below, p. 274, n. 22), which is sometimes recognized as a kind of Experiential Perfect (Smyth uses the term “empiric Perfect”).

(31) ἱππρίσμαι mēn egō kai propeπέλακιστai tò sòma
offend\PERF.Past.1.Sing. for-sure I and drag-in-the-dirt\PERF.Past.3.Sing. the
body\Nom.
toumòn tòte...
my then

“It was for sure I who was then offended and whose person was dragged in the dirt”

(Dem. 21.7)
E(xistential) Perfect is sometimes used as a cover term for the last three uses, but the Experiential Perfect is usually looked upon as the prototypical Existential Perfect.

In Greek as well we find cases of the Perfect that allow existential readings which are not resultative, as well as cases of the Universal Perfect. We start with an experiential reading.

(32) τούτῳ ἡμῶν δέομαι: εὰν διὰ τὸν αὐτὸν λόγον ακοῦὲτέ μου
this of-you begpres.1.Sing:
if through the same words hearpres.Subj.2.Plur IGen.Sing
apologouménon, δι᾽ ἑνὸπερ εἶόθα
defend-myselfpres.Past.Gen.Sing through which be-accustomed-toI.Sing
légein kai ἐν αγοραὶ ἐπὶ τὸν τραπέζων, hína ἡμῶν πολλοὶ
speakInf. both in market at of-the-bankers'-tables,
where of-you many
akēkóasi, καὶ ἄλλῳ μῆτε θαυμάζειν μῆτε θορυβᾶτιν
hearperfect.3.Pl., and elsewhere neither be-surprisedInf.
nor make-disturbanceInf.
touτου hêneka
of-this because
“And, men of Athens, I urgently beg you if you hear me making my defense with the same words with which I have been accustomed to speak both in the market place at the bankers’ tables, where many of you have heard me, and elsewhere, not to be surprised or to make a disturbance on this account”
(Pl. Apol. 17 c f.)

In the following example the use of the Perfect also has a purely existential, but not resultative meaning (on any natural reading of the sentence):

(33) σκέψασθε...εἴ τις χειρὶ Καὶ Ἐρατοθένει ἐχθρὰ πόποτε
considerimper.2.pl. if I.Dat and Eratosthenes.Dat. enmity ever
géγενεται πλὴν ταύτης
ariseperfect.3.Sing. beside this
“ask yourselves, if any enmity ever has arisen between me and Eratosthenes beside this one”
(Lys. 1.43)
For examples of the U-Perfect, cf. below (a propos participles). With adverbs like duó/tría/... étê (‘for two/three/... years’) the Greek Perfect is ambiguous between the U- and the E-reading. With adverbs like duóin/... etôn or pollôn etôn (‘in two/.../many years’) the Perfect can only have the E-reading (which in case of negation equals a U-reading). The following sentence illustrates the latter claim.

(34) ouk oísth' hóti pollôn etôn Agáthôn entháde ouk epi
dedêken...

not know2.Sing. that many-yearsGen. AgathonNom. here not be-
homeperfect.3.Sing.?
“Don’t you know that Agathon has not been home in many
years?”
(Pl. Symp. 172 c)

It should be obvious from these examples that the Perfect of Classical Greek looks a lot more like the Perfect of various modern languages than the traditional grammarians have thought22.

For certain languages, e.g. English, it is appropriate to analyse the Perfect in terms of an Extended Now (cf. McCoard 1978, Pickbourn 1798, Stechow 1999a). Since for Greek a focus on resultativity, if carried out in a traditional way, fails to account for uses of the Experiental Perfect, like in the three examples just quoted, or within a modern linguistic framework (like Kamp and Reyle 1993) leads to an overcomplicated analysis (cf. Stechow 2001b), it seems attractive to try the Extended-Now-approach also here. In order to check whether this leads to a feasible account of the data, we first need to look at some further examples.

When the Present Perfect is used in Greek, it frequently seems to be the case that the event denoted by the VP either continues after the speech time or that it at least continues up to the speech time (in an inclusive way). In terms of an XN-analysis, the speech time can be seen as a final subinterval of an interval which reaches into a contextually or lexically determined past (cf. Stechow 1999a, 2001b).

The following example is a Greek “garden-variety” use of the Perfect, which yields an XN-reading:

---

22 Naturally, also the “generic” Perfect, much discussed in Greek grammar, e.g. Xen. Anab. 3.1.389 hê ataxía polloûs édê apolôleken (‘lack of discipline has already been the ruin of many’), should be analyzed as an Experiental use of the Perfect (not unknown also in other languages).
Tense in time: The Greek perfect

(35) \textit{oligon gár chrónon allēlois dieilégmetha}
    short\textsubscript{Acc} for time\textsubscript{Acc} each-other\textsubscript{Dat} converse\textsubscript{perfect.1.Pl.}
    “for we have conversed with each other only a little while”
    (Pl. Apol. 37 a)

The larger context shows that the speaker, Socrates, is here referring to the
court situation in which he participates\textsuperscript{23}. This is depicted as “conversation”,
which has some duration in the past and includes the speech time.
Let us use the abbreviation \textit{XN(t,n)} for “t is a time interval that extends up
to n (and possibly includes n)”. The meaning of the last example can then
be paraphrased as “there is a time t: \textit{XN(t,n)} & we converse with each other
at t and the beginning of t is only a little while before n.”

In participle constructions, which are very frequent in Greek, the dis-
tinction between “conclusion before the speech time” and “continuation at
the speech time” can be quite accurately observed\textsuperscript{24}. A case where it is
clear that the activity denoted by the VP no longer continues at the speech
time is the following:

(36) \textit{humîn dè memarturêkasin hoi próteron ergazómenoi kai pollà}
    you\textsubscript{Dat.} and attest\textsubscript{PERF.3.Pl.} the earlier cultivate\textsubscript{PRES.Part.Pl.Nom} and many
    \textit{été par’ emû memishómenoi mé eînai sêkôn en tôi chôriói}
    years from me tenant\textsubscript{perfect.Part.Pl.Nom} not be\textsubscript{inf.} olive-stump on the piece-
of-land
    “and the previous cultivators, who rented it from me for many
    years, have testified to you that there was no stump on the plot”
    (Lys. 7.11)

\textsuperscript{23}Cf. the continuation 37 b “I believe that if you had a law, as some people have,
that capital cases should not be decided in one day, but only after several days, you
would be convinced; but now it is not easy to rid you of great prejudices in a short
time” and 39 e “But with those who voted for my acquittal I should like to con-
verse (hêdêos aû dialechtheîen) about this which has happened...Wait with me so
long, my friends; for nothing prevents our chatting with each other while there is
time”.

\textsuperscript{24}When looking for good examples in the Greek texts (e.g. through the TLG (The-
saurus Linguae Graecae)), one is soon reminded of the tendency of the Greek lan-
guage to put temporal information into participle constructions. This should con-
sequently be seen as another “excuse” for focusing so much on this type of material
in this context.
The speaker has just made a list of the earlier tenants, who no longer hold the piece of land (7.10 “In the fourth year I let it to Alcias, who is dead. After that Proteas, too, hired it in the same state for three years”). Once more, the Perfect expresses the relation XN, since an existential reading is available here: \( \exists t: XN(t,n) \land \exists t' \subseteq t(t' \text{ is many years } \& \exists t'' \subseteq t': x \text{ rents the land from me at } t'') \), where \( x \) are the people in question.

On the other hand, when we have a conclusion before the speech time (or before some other reference time) the Aorist Participle is preferred. This can be observed in almost any Greek text. (37) provides an example, where the Aorist participle yields an anteriority reading with respect to a time of reference in the past.

(37) *hoútô mèn tà tôn Hellênôn prágmata ephthárê hèx étê*

thus the of-the-Hellenes affairs\textsubscript{Nom.Pl.(Neutr.)} ruin\textsubscript{aor.Pass.3.Sing} six years\textsubscript{Acc.} polemésanta
make-war\textsubscript{aor.Part.Nom.Pl.(Neutr.)}”

“Our thus undertaking of the Hellenes came to naught after a war of six years”
(Thuc. 1. 110. 1)

With durational temporal adverbs like *polùn chrónon (édê)* (‘(already) for a long time’) we find the Perfect or the Present Participle. In (38) we find the Perfect, in (39) the Present Participle.

(38) *épeitá eisin hoútoi hoi katêgoroi polloi kai polùn chrónon édê katê-gorékôtes.*
further be\textsubscript{3.Pl.} these the accusers many and long time already accuse\textsubscript{perf.Part.} ,
étique dé kai en taútêi têi hêlikái légontes pròs humâs en héi ãn mâlista moreover and in that the age speak-pres-Part to you in which most episteúsate, would-believe
païdes óntes énioi d’ humôn kai meirákia
children be\textsubscript{pres-Part} some of you and youths

“These accusers are many and have been making their accusations already for a long time, and, moreover, they spoke to you at an age at which you would believe them most readily (some of you in
youth, most of you in childhood)"
(Pl. Apol. 18 c)

(39) *emoû gár polloi katêgoroi gegónasi pròs humâs kai pálai pollà êdê étê kai*

of-me for many accusers arise before you and already many years and

*oudên aléthês légontes, hoús egô mâllon phoboûmai ê toûs amphi Ánuton*

nothing true say, whom I more fear then the around Anytos

“For many accusers have risen up against me before you, **who have been speaking for a long time, many years already, and saying nothing true**: and I fear them more than Anytos and the rest”

(Pl. Apol. 18 b)

With adverbs semantically equivalent to the English *since* (or German *seit*) the Present is preferred, but here also the Perfect occurs. Here is an example with a *since*-adverbial and the Present Participle.

(40) *légein keleiôn tòn Tissaphérnê pròs autoûs hôs Athênaîoi ek*

say urge the Tissaphernes to them that Athenians since

*pleîonos chrónou epistêmones óntes toû nautikoû triôbolon longeur time experienced of-the-naval three-obols*

*toîs heautôn didóasin*

the of-them selves give

“He urged Tissaphernes to tell them that the Athenians, who **had had experience in naval matters for a longer time, gave only three obols to their men**”

(Thuc. 8.45.1)

---

25 The complication with present participles is that they express simultaneity with respect to the matrix tense. In (39), the main verb introduces an Extended Now and says that many accusers have arisen in that period; the participles specify what they did in that period and what they presumably still are doing. (The example is still interesting, however, especially as a pendant to (38), belonging to the same context and referring to the same category of people.)
In the following example we have a “since” adverbial combined with a Perfect participle.

(41) nûn tí deî makrêgoreîn, hôns toûs mèn
now what be-need-for3.Sing speak-at-great-lengthInf, of-which
someMasc.Pl.Acc.
dedoulôménous horâte, toûs d’ epibouleúontas
autoûs...kai ek polloû propareskeuasménous, ei pote
theyAcc. and since long time make-preparationsPerfect.Part.Masc.Pl.Acc., if
sometime polemêsontai;
make-warfuture.3.Pl.
“What need is there of a long harangue, when you see that they have
enslaved some of us and are plotting against others...and that they
have long (lit. since long time) been making their preparations
with a view to the contingency of war?”
(Thuc. 1.68.3 f.)26

The examples show that measure adverbials such as ek pléonos chrónou
“*since much time’ or ek polloû ‘*since long’ have a wider range of application
than their English counterpart and are more similar to German or
Swedish. To begin with, English since does not combine with a measure
term whereas German seit and Swedish sedan do: seit langer Zeit, seit
langem/sedan lange. But English since-adverbials are Perfect level adver-
bials, i.e., they only combine with the Present Perfect, Past Perfect, or
Future Perfect. In these combinations, the adverbs characterize an XN-interval
introduced by the tense form (cf. McCoard 1978, Dowty 1979). No such
restriction holds for German seit, which combines with simple tenses as

26 In Greek, also the Perfect and Present Infinitive behave in a similar way in their
way of combining with adverbs. Sometimes, with durational adverbs, we even find
them side by side, like in Pl. Apol. 31 b ou gár anthròpônìs òiıkê to emê tôn mèn
emautòù hapántôn ìmelêkêtaî kai ònécheîtaî tôn oikeîôn anelouvémôn tosaûta
édê êtê (‘it does not seem human that I have neglected all my own affairs (neg-
lect.PERF.Inf.) and have been enduring (endure.PRES.Inf.) the neglect of my
concerns all these years’).
well and introduces an XN-reading in such contexts (Stechow 2001b) (In Swedish the Present with *sedan...tillbaka* yields an XN-reading, cf. below).

The other, more important point is the following: in spite of the quite frequent use of the Present Participle (and generally the Present) with adverbs like *pollà étê/połun chrónon* and *ek polloû chrónou*, it could now be argued that the classical Greek Perfect introduces an Extended Now. In Greek, as in certain other languages, e.g., Swedish, an adverb may sometimes introduce an Extended Now. In such cases the Present is chosen. Cf. Swedish:

(42)  *Han bor här sedan fem år tillbaka.*
‘he lives here since five years back’

In other cases, usually with other adverbs, it is the Perfect which introduces the Extended Now. Cf. Swedish:

(43)  *Han har bott här i fem år.*
‘He has lived here for five years (and still lives here)’

Adopting such an analysis in (40), we claim that the adverb *ek pléonos chrónou* (*since a longer time*) introduces an Extended Now, whereas in (41) the Perfect is the introductory “item”, the adverb qualifying only the Extended-Now-interval\(^{27}\). Another important fact in this context is the ob-

\(^{27}\) German behaves differently in this respect. *Seit* combines with the Perfect, but only in its Preterite use; with the Present, on the other hand, *seit* introduces an Extended Now:

Er wohnt hier seit 5 Jahren
‘he lives here since 5 years’

Sometimes the Perfect introduces the Extended Now, usually with other adverbs. Schiporeit (1971) and Rathert (1999) give arguments that XN readings are triggered by the presence of the adverb *schon*:

| Schiller hat immer gute Freunde gehabt | No XN |
| Schiller has always good friends had | |
| ‘Schiller always had good friends’ | |

*Schiller hat schon immer gute Freunde gehabt*  XN

Schiller has already always good friends had
‘*Schiller always has had good friends’
servation that the combination of the Greek Perfect with *since*-adverbials indeed *necessitates*, not only *allows* an XN-reading.

The introduction of the XN-perfect conveys an important change of the system: while the Homeric Perfect was an “Aktionsart chooser”, the XN-perfect has a temporal meaning, but it still does not quite fit into the conventional picture drawn by Reichenbach (1947), Klein (1994), and others who look upon tense as a relation between the utterance time and the reference time. An XN rather stretches the reference time into an indefinite past. In consequence, it is a tense *sui generis*, say an “auxiliary tense”, because it requires the presence of another tense. (44) proposes a formal analysis which is sufficient for the purposes of the present discussion:

(44) **The XN-Perfect**

\[ \text{XN-PERF} = \lambda P \lambda t \exists t'[\text{XN}(t', t) \land P(t')] \]

Recall that XN(t', t) means that t is final subinterval or point of t'. Here is an analysis of the perfect statement (35):

(45) **olígon gàr chrónon allêlois dieilégmetha**

“For a short time we have been conversing with each other”

\[ \left[ \text{T}_{\text{PRES}} \left[ \text{T}_{\text{XN-Perfect}} \text{for-short-time} \left[ A_{\text{we talk to each other}} \right] \right] \right] \]

\[ \exists t' [\text{XN}(t', \text{PRES}) \land t' \text{ is short} \land \exists e \subseteq t'[e: \text{we converse with each other}]] \]

Returning to example (34), we now get the following analysis: with an Extended-Now-analysis the Perfect denotes a time interval extending from the reference time to an indefinite past. The temporal adverb “for n years” is predicated of that interval and tells us that it is of a specific length. The temporal adverb “in n years” picks out a subinterval of that interval. The aspect phrase is located in the restricted interval. Below is a picture of the scenario:
Tense in time: The Greek perfect

pollôn etôn Agáthôn ouk epidedémêken.
{--------[ ///////////// ]---------PRES}
{------}  Perfect interval
[-------]  “in many years”-interval, “frame interval”
/////////  Negation of the (perfective) VP “There is no being home”

The frame interval may extend up to the speech time PRES, but it needs not. Note that the impression that the embedded VP is stative comes from the fact that the occurrence of the event is negated. Here, finally, is the picture for a for-interval in a construction expressing a Universal Perfect.

polla étê Agathôn ouk epidedémêken.
{ ///////////// /////////////PRES}
{------}  Perfect interval = frame interval, many years long
/////////  Negation of the (perfective) VP “There is no being home”

3.3. Hellenistic and Greek-Roman (300 BC – 600 AD)

In the Greek of this period the intensive meaning of the Perfect is completely lost. The core meaning of the Perfect is ex hypothesi still an Extended-Now-meaning. All the various E-Perf ects found in classical Greek still occur, e.g., in the New Testament. The non-active intransitive Perfect is gaining importance, especially the Perfect Participle. The PAST meaning, i.e., a variety of the XN-Perfect, is becoming more and more frequent.

The examples below illustrate (i) the resultative use (46), (ii) the exper iential use (47), and iii) the Past use (37).

(46)  teuthnêkasin gár hoi zêthoûntes tên psuchên tou paidiou
die,perfect.3.pl for the search Pres.Part.Nom.Pl the life Acc. the child Gen.
“those who sought the child’s life have died ( = are dead)”
(Matth. 2: 20)

The present participle zêthoûntes could not correctly be integrated into our account if it contained a semantic PRESENT. This is a separate issue, which cannot be treated here. The intuitively most attractive interpretation
of example (46) seems to be in terms of the resultative aspect, though one may hold the view that the sentence illustrates the "experiential" use.  

(47)  *Theon oudeis heôraken pòpote*  

\[\text{god Acc. nobody Nom. see} \text{perfect.3.sg ever}\]  

"Nobody has ever seen God"  

(Joh. 1: 18)  

(48)  *Toûton ho theòs [kai] árchonta kai lutrötên apéstalken sin*  

\[\text{he Acc. the god Nom. (both) leader Acc. and deliverer Acc. send} \text{perfect.3.Sing. with}\]  

\[\text{cheirì aggéloü tou òphthéntos autôi en téi bátôi}\]  

\[\text{hand Dat. angel Gen. see} \text{aor. Part. Marc. Sing. Gen. in the bush Dat.}\]  

"God sent him as [both] leader and deliverer by the hand of the angel that appeared to him in the bush"  

(Acts 7: 35)  

To our mind, these examples should be interpreted in terms of an Extended Now. (They can all be seen as instances of the E-Perfect). As for cases like (48), the interpretations of the traditional grammarians are at variance. Some look upon this type of Perfect as semantically equivalent to a Simple Past (Aorist) (e.g. Mandilaras 1973, Wackernagel 1920–24), some as, in some sense, a Perfect of Result (e.g. Chantraine 1927 and, with some reservation, Fanning 1990). Within an XN-analysis such cases yield an E-Perfect-interpretation: there is an event and there is a time such that XN, etc. A special problem is posed by examples where the (morphological) Perfect and the (ditto) Aorist are used side by side with apparently very little difference in meaning. It is, however, quite possible, that an Aorist-

---

29 The issue is hard to decide, and the context does not help us any further. The passage reports the words of an angel who tells about the death of Herodes. Surely, Herodes is dead at moment of the speech, but he died shortly before. If the sentence contained the adverb \(nûn\) ‘now’ it clearly would express a perfect of result.  

30 Such an example is the following:  

\[\text{heurôn de hêna polûtîmon margarítên apelthôn pépraken pânta hòsa eîchen kai} \]  

\[\text{égorasen autôn pânta hòsa eîchen kai} \text{égorasen autôn}\]  

\[\text{find aor. Part. Masc. Sing. Nom. and one precious pearl Acc. go-}\]  

\[\text{away aor. Part. Masc. Sing. Nom. sell} \text{perfect.3.Sing.}\]  

\[\text{all Acc. that Acc. have} \text{impf.3.Sing. and buy aor. 3.Sing. it Acc.}\]  

"...who on finding a pearl of great value went and sold (lit. has sold) all that he had and bought it"  

(Matt. 13: 46)
like (i.e., Past-oriented) E-Perfect could be used together with a real Aorist with almost, but not quite, the same meaning. The present account clarifies this often invoked difference: the Aorist reports an event which is over; it can never have a reading which expresses continuation of a state up to the speech time. The Perfect, by contrast, can also be used in such cases, as we know from examples expressing the U-Perfect. Furthermore, the Aorist and the XN-Perfect are, of course, radically different from a conceptual point of view, even if they figure in sentences that are truth-conditionally equivalent.

3.4. Byzantine period (600 – 1450 AD)

During this era the morphology of the Perfect and the Aorist become more and more similar: reduplication is used instead of augmentation in Aorist forms and the suffix -k-, typical of the Perfect (but also earlier found in some Aorists, the so-called kappa-Aorists), spreads more generally within the Aorist paradigm; additionally, the Perfect and Aorist personal endings become more alike (cf., e.g., Hedin 1999, Ruge 1991).

Typical for nonartificial language of this period is the replacement of the synthetic Perfect by analytic constructions, formulae which were used in earlier Greek to express a special resultative meaning. The periphrases in use were (i) eimí (‘be’) with Aorist or Perfect Participle, and (ii) échô (‘have’) with Present or Aorist Participle, e.g. (i) (replacing the Pluperfect) Theoph. 260. 1 έν κελεύσας (‘had ordered’), Cedr. II 331, 21 τοὺς μετ’ αὐτοῦ ἐκσκαθομένους ὄντας (‘those of his men who had been scattered’), (ii) Leo Neap. 76. 5 εγώ ἐχό ἡξέκοντα έτε ἡμίπων διὰ τὸν ἐργὸν μου τὸν Χριστόν (‘Through my deeds I have affronted Christ for sixty years’) and Duk. 210. 21 (again for a Pluperfect) εἶχε περάςας τὸν Ἰστρόν (‘he had crossed the Istrus’) (cf. Psaltes 1974).

In traditional treatments of the Perfect of Postclassical Greek such instances of the Perfect are often highlighted as proofs of an Aorist meaning of this tense (e.g. in Mandilaras 1973).

Already in Chantraine (1927) we find the suggestion that the Postclassical “aoristic” Perfect has its own (Perfect) meaning, not identical to that of the genuine Aorist.

Cf. above, p. 270.
Towards the end of the Byzantine era we also find a periphrasis involving the auxiliary échein and the supine of the verb, the aparemfeito. It is not our intent to give a precise account of the history of this form, but we remind the reader that the appearance of the analytical perfect is not a geographically isolated phenomenon; analogous constructions develop in the great majority of European languages during the medieval era.

The emerging system is quite mixed, but already very close to what we find in Modern Greek. Here is a survey.

<table>
<thead>
<tr>
<th>Form</th>
<th>T</th>
<th>Aux</th>
<th>Aspect</th>
<th>root-meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>gráfo</td>
<td>PRES</td>
<td></td>
<td>IPFV</td>
<td>graf-“write”</td>
</tr>
<tr>
<td>grápsa</td>
<td>FUTR</td>
<td>neutral</td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>thélo gráfin/ na gráfo</td>
<td>FUTR</td>
<td>thélo</td>
<td>IPFV</td>
<td>“</td>
</tr>
<tr>
<td>thélo grápsa/ na grápsa</td>
<td>FUTR</td>
<td>thélo</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>égrafon/égrafo</td>
<td>PAST</td>
<td></td>
<td>IPFV</td>
<td>“</td>
</tr>
<tr>
<td>égrapsa</td>
<td>PAST</td>
<td></td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>gégrafa</td>
<td>PRES?</td>
<td></td>
<td>IPFV?</td>
<td>RESULT+“write”?</td>
</tr>
<tr>
<td>écho gráfon</td>
<td>PRES</td>
<td>echo XN-</td>
<td>IPFV</td>
<td>“</td>
</tr>
<tr>
<td>écho grápsas</td>
<td>PRES</td>
<td>echo XN-</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>eimí grápsas/ écho grápsi</td>
<td>PRES</td>
<td>echo XN-</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>ícha gráfon</td>
<td>PAST</td>
<td>ícha XN-</td>
<td>IPFV</td>
<td>“</td>
</tr>
<tr>
<td>ícha grápsas</td>
<td>PAST</td>
<td>ícha XN-</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>ên grapsas/ ícha grápsi</td>
<td>PAST</td>
<td>ên/ícha XN-</td>
<td>PFV</td>
<td>“</td>
</tr>
</tbody>
</table>
Synthetic Perfect forms are not mentioned in the chart. They are still found throughout the Byzantine period, but the meaning is often that of a PAST (or possibly an PRES + Extended Now). Here is an example:

(49) kontaréas dedôkamen
wounds-with-lance give.Perf.3.plur.
“we wounded (lit. have wounded) them with our lances”
(Digenis Akritas 6.727)

It is not clear to us whether the resultative interpretation of transitive forms persists in this period. Note further that we have given not at all a semantic analysis of transitive Perfects of result in this paper. See Stechow (2001b) for an analysis of transitive resultatives.

The following example exhibits the combination of échô (‘have’) with an aparémfato. The meaning is, however, perhaps that of an Existential Perfect. It is very well possible that the periphrastic perfect is analyzed as a combination of PAST + ASPECT. The Byzantine system shows a tendency towards symmetry: we find aspectual oppositions in all the tenses including the analytical Perfect. In comparison, the modern Greek system seems somewhat impoverished, since the supine only exhibits perfective morphology. The following example illustrates the aparémfato construction.

(50) o kápoios Frágkos eugenês, ánthrôpos paideuménos,
the some Frankish nobleman, person educated,
apó tén pólín éxei élthei apó ton basiléan
from the city,Acc. have.pres.3.sing. come.Supin. from the emperor,Acc.
“Some noble Frank, an educated man, has come from the city, from the emperor”
(Morea-Chronicle 4900–1)

3.5. Modern Greek (From 1450 AD)

In MG, the synthetic Perfect and the synthetic Future disappear completely. On the other hand, the morphological opposition Perfective/Imperfective is rather systematically implemented. Past (Aorist/Paratatikos) and Future exhibit an aspectual opposition, but the perfect participle does not. There is
no semantic reason for this gap; e.g., in Bulgarian, the Perfect Participle realizes the said opposition Iatridou et al. (2001). As in English (or in earlier stages of the Greek language), we do not find a perfective Present form.33

As is well known, the Balkan languages have lost the infinitive. As a consequence the future auxiliary thélo embeds a finite form or a supine.

<table>
<thead>
<tr>
<th>Form</th>
<th>T</th>
<th>Aux (+C)</th>
<th>Aspect</th>
<th>root-meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>gráfo</td>
<td>PRES</td>
<td></td>
<td></td>
<td>graf‘“write”</td>
</tr>
<tr>
<td>thélo gráfi(n)</td>
<td>FUTR</td>
<td>thélo</td>
<td>IPFV</td>
<td>“</td>
</tr>
<tr>
<td>thélo na/théna gráfo</td>
<td></td>
<td>thélo na</td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>thélo grápsi(n)</td>
<td>FUTR</td>
<td>thélo</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>thélo na/</td>
<td></td>
<td>thélo na</td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>théna grápsi</td>
<td></td>
<td></td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>tha grápsi</td>
<td></td>
<td></td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>égrafa</td>
<td>PAST</td>
<td></td>
<td>IPFV</td>
<td>“</td>
</tr>
<tr>
<td>égrapsa</td>
<td>PAST</td>
<td></td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>écho grápsi</td>
<td>PRES</td>
<td>écho</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>XN-Perfect?</td>
<td></td>
<td></td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>ícha grápsi</td>
<td>PAST</td>
<td>ícha</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>XN-Perfect?</td>
<td></td>
<td></td>
<td></td>
<td>“</td>
</tr>
<tr>
<td>tha écho grápsi</td>
<td>FUTR</td>
<td>tha écho</td>
<td>PFV</td>
<td>“</td>
</tr>
<tr>
<td>XN-Perfect?</td>
<td></td>
<td></td>
<td></td>
<td>“</td>
</tr>
</tbody>
</table>

Like the Perfect of Classical and Postclassical Greek, the MG equivalent of this tense may yield the resultative or the experiential reading:

(51) Écho grápsi to grámma. “Result”

   have.1.Sing. write.supine the letter

   ‘I have written the letter’

(52) Échis pái sti Germanía? “Experience”

   have.2.Sing. go.supine to-the Germany

   ‘Have you been to Germany?’

33 The most obvious and perhaps correct explanation is that the speech time is conceived as point of time and is therefore too short to embrace an entire event. We assume that this explanation is originally due to Dowty (1979).
A difference from older Greek is that the MG Perfect never yields a Universal reading. E.g. the combination of the Perfect with the adverbial pάntα ('always') is ungrammatical.

(53) *Écho pάntα zίsi stin Athīna
    have.1.Sing. always live supín in-the Athens
    'I have always lived in Athens'

Iatridou et al. (2001) explain the ungrammaticality of the example as follows. According to them, the analytic Perfect in MG denotes an Extended Now34. Given that the aspect of the supine is perfective (both morphologically and semantically), we cannot have a U-Perfect, which requires the IPFV by definition.

We accept this theory for the time being, but with reservations. Typologically, it is highly unlikely to find an XN-Perfect without a U-reading, for it is precisely the existence of this interpretation which motivates the theory that there is such a fancy tense. We also ask why the PFV/IPFV opposition should be missing exactly in this construction. Recall that we still found it in the Byzantine period. Consequently, there is some evidence that the XN-Perfect still survived at this stage of history, but more data would be required to establish the point.

As for MG, one rather gets the impression that the analytic Perfect is a competing form for the Aorist. If this turns out to be true, the analysis given must be revised in an obvious way: the Present Perfect (écho grάpsι) must be analysed as PAST + PFV, the Past Perfect (ícha grάpsι) must be analysed as PAST + PERFECT (= POST) and the Future Perfect, i.e., the form tha écho grάpsι, must rather be FUTR + PERFECT (= POST). We think that this alternative is rather promising, but leave it to future research to work out the details.

4. Conclusion

Re-examining the diachronic development of the Greek Perfect in the light of older as well as more recent theories of the Perfect “tense” we have

34 A similar analysis of the MG Perfect we find earlier in Hedin (1987), only in a different framework (Hedin in her dissertation draws upon Bull’s concept of an “extended present”).
found that the Archaic (Homeric) Perfect displays properties most typical of an intensifier or an aspect (result). The two meanings connected with the Perfect of this time are, however, too different to allow serious assumptions as to their (historical) common denominator.

In Classical Greek the Perfect seems to have developed into something between a stativiser and a tense. Its core meaning is now an Extended Now, but the resultative reading is still available. As the above discussion of the data has shown, the XN-approach accounts for much of the material that causes problems for a theory focusing on resultativity, be it in a traditional or a modern linguistic framework.

In Postclassical Greek the Perfect still denotes an Extended Now. There is no reason to assume that during this time the Perfect assumes a meaning of a Simple Past (Aorist). The criteria formulated by grammarians like Mandilaras (taken over partly also by, e.g., Fanning 1990) for judging when to regard a Perfect form as having a Past meaning may seem attractive at first glance, but they do not really provide much help in narrowing down the semantic value of a Perfect form in concrete cases. (One argument used by Mandilaras is the side-by-side use of Perfect and Aorist forms found in Postclassical texts; another one is the combination of the Perfect with past-oriented, punctual adverbs [cf. Mandilaras 1973: 221]).

As for the classical “Present Perfect-puzzle”35, it is true that the Greek Perfect during this period quite frequently combines with past-oriented adverbs. This, however, does not really constitute evidence against a preserved RESULT-meaning of the Perfect morphology in Postclassical Greek, since the use of the Perfect may still be more now-oriented in such cases36. Definite, punctual adverbials, then, when combined with the Perfect, do not refer to the Extended-Now-interval, but to the event time (already in Classical Greek, cf. examples above, e.g. (21)).

The reason, then, for the demise of the synthetic Perfect is to be found on the semantic and the morphological level. On the one hand, the synthetic

35 The semantic problem constituted by the extremely limited occurrence of the Present Perfect with punctual adverbs such as yesterday in languages like English and Swedish (cf. e.g. Klein 1992).
36 Cf. for instance Hedin (1999) apropos this kind of data in Classical Greek: "The fact that the perfect is combined with a definite time adverbial does not necessarily implicate, for instance, that it is used narratively...but (it could mean) that the event is connected to the present in another way. Perhaps as a kind of background to what is happening now" (our translation).
Perfect had not acquired the same meaning as the Aorist, as is often assumed, but nevertheless – mainly through its Experiential use – a similar meaning. On the other hand, the morphological confusion led to a loss of understanding of the distinction between the two tenses. Finally, competing analytical constructions in later Greek were there to take over the old XN-meaning of a form no longer completely understood.

The “career” of analytical constructions has only briefly been touched upon in this paper, since our focus has been more on the semantic development of the synthetic Perfect. However, it is important to note that the early periphrastic constructions (Perfect or Aorist Participle with εἶναι ‘be’ or ἔχειν ‘have’ in Classical Greek) often have a very pronounced resultative meaning and therefore require a different semantic analysis than the synthetic Perfect. (Even in its resultative use the synthetic Perfect seems to have a different meaning from the analytic resultative constructions. For discussion of such constructions in other languages, e.g., Swedish, cf. Dahl 1985: 133f.). One of these analytic constructions has survived in MG, retaining a resultative meaning (the ἔχο grammēno-periphrasis). Note, however, that the regular analytic Perfect in MG has developed from a completely different construction, ἰκχα (‘had’) with the Aorist infinitive (as a kind of futurum preteriti; cf. Browning 1969; Hedin 1987; Ruge 1991).

The major claim made in this paper is that the Greek Perfect up to the Byzantine period, with the exception of Archaic Greek, should be understood in terms of an Extended Now. According to Iatridou et al. (2001), this claim holds even for MG. Thus the Greek Perfect, as we have pointed out, shows much similarity with the Perfect, e.g., in English and Swedish. In other languages, however, the Extended-Now-approach to the Perfect is under debate.

References

Chantraine, Pierre

Chomsky, Noam

Dahl, Östen

Dowty, David

Fabricius-Hansen, Cathrine

Fanning, Buist M.

Gildersleeve, Basil Lanneau

Giorgi, Alessandra and Fabio Pianesi

Hedin, Eva

Heim, Irene
1994  Comments on Abusch’s Theory of Tense. Manuscript. MIT.

Heim, Irene and Angelika Kratzer

Herweg, Michael

Humbert, Jean
Iatridou, Sabine, Elena Anagnostopoulou, and Roumyana Izvorski

Kamp, Hans and Uwe Reyle

Katz, Graham

Klein, Wolfgang

Kratzer, Angelika

Kühner, Raphael and Bernhard Gerth

Landman, Fred

Mandilaras, Basileios G.

McCoad, Robert W.

McKay, Kenneth L.

Mourelatos, Alexander Peter D.

Musan, Renate

Partee, Barbara

Pickbourn, James
1798 A Dissertation of the English Verb: Principally Intended to Ascertaın the Meaning of its Tenses. London: J. Davis

Psaltes, Stamiatos B.

Rathert, Monika

Reichenbach, Hans

Rijksbaron, Albert

Ruge, H.

Schiporeit, Luise

Schoorlemmer, Maaike
Smyth, Herbert W. and Goldon M. Messing

von Stechow, Arnim


Vendler, Zeno

Wackernagel, Jacob
Light verbs in Urdu and grammaticalization

Miriam Butt and Wilhelm Geuder

1. Introduction

This paper focuses on complex predicates of the V-V type, as illustrated in (1) for the South Asian language Urdu. These types of complex predicates consist of a full lexical verb accompanied by a so-called light verb (other names include vector or explicator verb). The examples in (1) show the light verbs ‘give’, ‘take’, ‘sit’, ‘fall’, and ‘go’, for instance. Their designation as “light verb” is due to the fact that their semantic contribution to the predication is weak and tantalizingly elusive:

(1) a. $\text{yaasiin=nee keek k\textp{\textdegree}aa lii-yaa}$
   \hspace{1cm} Yassin.M=Erg cake.M.Nom eat
   take-Perf.M.Sg

   ‘Yassin ate the cake (completely, for the benefit of himself).’

* An earlier version of the historical material contained in this paper was presented at the Workshop on New Reflections on Grammaticalization in Potsdam, June 1999. A preliminary version of the synchronic analysis of light verbs was first presented at the Workshop on Semi-lexical Heads in Tilburg, May 1999. The paper has profited from discussions with Ashwini Deo, Regine Eckardt and Shin-Sook Kim as well as from the comments of two anonymous reviewers. The diachronic discussion is part of on-going joint work with Aditi Lahiri and can be found in more detail in Butt and Lahiri (2002). Many thanks go to Karin Schunk for deeply appreciated help with the Sanskrit examples. The work was written with support from the SFB 471 “Variation und Entwicklung im Lexikon,” funded by the Deutsche Forschungsgemeinschaft (DFG).

1 The South Asian languages Urdu and Hindi are closely related. Both are among the 16 official languages of India and are spoken primarily in the north of India. Urdu is the national language of Pakistan. The data presented in this paper are drawn primarily from the dialect of Urdu spoken in Lahore, Pakistan, as well as from examples cited in the literature on both Urdu and Hindi.
Miriam Butt and Wilhelm Geuder

b. naadyaa=nee xat lik5h dii-yaa
   Nadya.F=Erg letter.M.Nom write
give-Perf.M.Sg
   ‘Nadya wrote a letter (completely, forcefully, for somebody else).’

c. naadyaa aa bEt5h-ii
   Nadya.F.Nom come sit-
Perf.F.Sg
   ‘Nadya has arrived.’

d. naadyaa gaa par5-ii
   Nadya.F.Nom sing fall-
Perf.F.Sg
   ‘Nadya burst into singing (fell to singing).’

e. naadyaa gir ga-yii
   Nadya.F.Nom fall go-Perf.F.Sg
   ‘Nadya fell (down).’

It is perhaps this elusive quality of the light verbs that has prompted a number of synchronic studies on V-V constructions not only for Urdu/Hindi (e.g., Hook 1974, Abbi and Gopalakrishnan 1991, Singh 1994, Butt 1995), but also for other South Asian languages (e.g., Vale 1948, Zbavitel 1970, Dasgupta 1977, Ramchand 1990, Gopalakrishnan and Abbi 1992). In addition, these light verbs have been discussed from a historically oriented point of view in terms of the idea of progressive grammaticalization and concomitant semantic bleaching (Hook 1991, 1993, Hopper and Traugott 1993, Abbi 1999). The idea is that the formation of these types of V-V complex predicate constructions can be explained through a progressive semantic bleaching from the full verb meaning of ‘give’ or ‘go’ to a “lighter” meaning, which takes on a more grammatical, as opposed to lexical, function in the language through the loss of some of its semantic content.

The present paper takes issue with this view of light verbs as grammaticalized elements, and presents a synopsis of the historical development and the synchronic status of light verbs in Indo-Aryan, with special reference to Urdu. While the explanation in terms of grammaticalization is intuitively appealing at first glance, several observable properties of the light verbs in
do not quite submit to the proposed scenario. We therefore put forth an alternative account that stresses the lexical, as opposed to grammatical, nature of light verbs. Diagnosing light verbs as “bleached” variants of full verbs may by itself not be objectionable, depending on the exact sense ascribed to the term “bleaching”. What is to be denied, however, are the historical and syntactic implications that derive from identifying semantic bleaching with grammaticalization. Rather, our main point is that the pairing of light and full verb readings of the same item constitutes a case of lexical polysemy, but not of grammaticalization.

We begin by setting out the grammaticalization hypothesis for light verbs and discuss some further syntactic and semantic aspects of the problem. This is followed by a synchronic analysis of light verbs in Urdu and by a diachronic investigation into the origins of the light verb construction. From both of these perspectives, the grammaticalization thesis proves to be very doubtful. In the final section, we sketch a lexical-semantic account of light verbs that reconciles the semantic intuitions and the observed data with the lack of grammaticalization effects.

2. A grammaticalization account for light verbs

Given that light verbs exhibit a fairly weak and elusive semantics, and given moreover that for each light verb a corresponding full verb variant can be found with a stronger meaning, it is tempting to assume that light verbs are diachronically derived via semantic bleaching of a full verb, and hence are the outcome of a grammaticalization process. Such an account has been proposed by Hook (1991), who offers a comparative study of light verbs in a number of Indo-Aryan languages. Hook proposes that the different ranges of light-verb phenomena that can be found in these languages can be ordered along a grammaticalization cline. The endpoint of the development, on his account, is a light verb in the function of a marker for perfective aspect. This hypothesis is bolstered by a diachronic study of the use of light verbs in texts dating from 1800 to the present for Hindi/Urdu, from 1600 for Marwari, and from 1400 for Bengali (Hook 1993).

The central point of Hook (1991, 1993) is that he observes a correlation between the distribution of light verbs and their semantics: the more widespread light verbs are in a language, the more their meaning proves to be weak and detached from the lexical meaning of the corresponding full verb. This state of affairs shows that the range of uses is a direct result of the weakness of the lexical meaning, i.e., of the degree of semantic bleaching that they have undergone.

Hook’s (1991) first observation concerns the frequency with which light verbs occur in texts: in this respect, a scale can be established that reaches from Kashmiri and Marathi, the languages with the lowest text frequencies, up to Hindi/Urdu, the language with the highest. Second, conditions on the possible contexts are identified, which are seen as originating from the semantic contents retained by light verbs. Hindi/Urdu is again found at the end of the scale in that it allows uses of verbs like ‘give’ that do not appear to conform at all to the literal meaning of ‘give’. This is evidenced by combinations of light ‘give’ with full verbs as in the following examples (cited from Hook 1991: 67):

(2)  
\[\text{aaxirkaar } us=nee \text{ muf}^\text{b}=koo\]  
in the end Pron.Obl=Erg I.Obl=Dat  
paise money.M.Pl.Nom  
give give-  
\text{Pf.M.Pl}  
‘In the end he gave me the money.’

(3)  
\[\text{is}=me$s$e$ \text{us}=nee \text{teer}-aa\]  
this=in Pron.Obl=Erg  
your-M.Sg  
naam name.M.Sg.Nom  
lee take give-  
\text{Pf.M.Sg}  
‘In this he took your name (he implicated you in this, he named you).’
In example (2), the light verb is the same as the full verb, and so the specific contribution to the predication seems quite mysterious. In (3), on the other hand, full and light verbs are antonyms, so this sentence should somehow feel contradictory (which it does not) if the light verb had its normal lexical meaning. Likewise, (4) describes a situation in which nothing whatsoever is “given”, because the main verb ‘feel’ describes a situation that is more similar to an instance of receiving something. All these cases point to a use of ‘give’ in which its normal lexical-semantic features don’t play any role. The reason seems to lie with specific developments of the light verb ‘give’ in Urdu, for the corresponding constructions in Marathi are unacceptable even though this language does have complex predicate constructions (but in the cases below, a simple main verb would be used): 2

(5) a. *SevaTI tyaa na ma-laa paise de-un dile finally he ERG me-to money give-CP gave
   ‘Finally, he gave me the money.’
   (Marathi, from Hook 1991: 67; cf. (3) above)

b. *hyaacaat tyaa na tudzha naaw ghe-un dila this-LOC he ERG your name take-CP gave
   ‘He implicated you in this.’
   (Marathi, from Hook 1991: 67; cf. (4) above)

2 The examples are shown in the transliteration used in Hook (1991). “T” is a retroflex t, “S” a palatal fricative.
Given that the semantic contribution of the light verb in the Urdu examples (2) to (4) cannot be identical to the normal lexical meaning and, moreover, is so elusive, it is tempting to think that the light verb has been reduced to some kind of grammatical marker. Hook shows furthermore that there is a correlation between the semantic weakening of verbs in Urdu complex predicates and their greater textual frequency, and concludes that light verbs in Hindi/Urdu are grammaticalized to a higher degree than in other closely related languages (like Marathi).

The data shown so far, however, can strictly speaking only be used to show that the lexical meanings in question are generalized. It still has to be shown that we are dealing with a grammaticalization phenomenon. Concerning the claim of a historical development in Hook (1993), it is not clear whether the reported increase in light verb usage is statistically significant. This is a particular worry for Hindi/Urdu, where the usage only varies between 12% and 18% in the texts under investigation, and does not show a consistent increase in usage over time. Furthermore, Meißner’s (1964) examination of a version of the Ramayana dating from 1575–1600 indicates a very well-established and wide-spread use of the light verbs ‘go’, ‘come’ and ‘put’. These data have not been factored into Hook’s account.3

To further establish the grammaticalization account, Hook (1991) claims that the light verbs in Urdu/Hindi have gained the new grammatical function of aspect marking, i.e., are seen as markers for perfectivity. Slavic scholars of the Indic languages have generally reported intuitions of a correspondence between light verbs and Slavic perfective aspect. Apart from such intuitive judgments, Hook presents two criteria to show that the function of the light verbs is really the marking of perfectivity. He identifies two contexts which in Slavic languages always require perfective morphology, and shows that Urdu requires the light verbs in the same contexts (more specifically, the light verb ‘give’ appears). The relevant contexts are: sentences describing possibilities which occur as the object of attitude verbs like ‘fear’ (see (6) and (8)), and sentences that are introduced by the

---

3 These diachronic data have also not been taken into account in Abbi’s (1999) claim that the light verb ‘go’ is a product of grammaticalization in South Asian languages in general. Abbi (1999) comes to much of the same conclusions as Hook. However, as Abbi’s claim is based on a synchronic typological study, but no diachronic data, we here focus on Hook’s work.
conjunction “until”, highlighting the inception of an event (see the Russian examples (7) and (9)):

(6)  
\[
\begin{array}{llllll}
\text{mat’} & \text{bojalas’} & \text{kak-by} & \text{eē} & \text{syn} & \\
\text{mother} & \text{feared} & \text{lest} & \text{her} & \text{son} & \\
\text{ne} & \text{zabolel} & & & & \\
\text{NEG} & \text{took-sick-PFV} & & & & \\
\end{array}
\]

‘The mother was afraid lest her son might get sick.’

(7)  
\[
\begin{array}{llllll}
\text{ona} & \text{budet} & \text{ubajukivat’} & \text{rebēnka} & \\
\text{poka} & \text{she} & \text{will} & \text{rock} & \text{child} & \\
\text{while} & & & & & \\
\text{on} & \text{ne} & \text{zasnēt} & & & \\
\text{he} & \text{NEG} & \text{sleeps-PFV} & & & \\
\end{array}
\]

‘She will rock the baby till it goes to sleep.’

(8)  
\[
\begin{array}{llllll}
\text{muj’e} & \text{d5ar} & \text{t’aa} & \text{ki} & \text{kahi$i$i} & \\
\text{tum} & \text{I.Obl.Dat} & \text{fear.Nom} & \text{be.Past-M.Sg} & \text{that} & \\
\text{lest} & \text{you} & & & & \\
\text{dee} & \text{use} & \text{cit5$^b$i} & \text{naa} & & \\
\text{doob} & \text{Pron.Obl.Dat} & & & \text{letter.F.Sg.Nomnot} & \\
\text{give} & \text{give.Imp} & & & & \\
\end{array}
\]

‘I was afraid that you might give him the letter.’
(9) tum yaha$a$ t5$h$aharoo jab-tak voo
you.Nom here stay.Imp until
Pron.Nom
tumhe$ cit5$t$h$ii naa dee dee
give.Dat letter.F.Sg.Nom not give
give.Subj

‘Wait here until he gives you the letter.’

To sum up so far, Hook (1991) establishes a semantic gradient of generalization in light verb meanings; the most general variants of light verbs (those occurring in Hindi/Urdu) are analyzed as aspectual markers. Consequently, he posits a process of grammaticalization that leads from fully lexical verbs to grammatical aspect. While we believe that Hook’s observation of a correlation between semantic generality and breadth of distribution is valid, his characterization of the semantic gradient of generalization will be found to be problematic; we return to this topic in section 5.

2.2. The grammaticalization account in broader perspective

In this section, we discuss the current ideas on historical change as they could be applied to the formation of light verbs.

2.2.1. The grammaticalization cline

The theory of grammaticalization describes a type of historical change which follows a developmental cline (Hopper and Traugott 1993). Along this gradient of grammaticalization, the loss of semantic content goes hand-in-hand with a progressive loss of the material integrity of a lexical item and a loss of its independent status in the syntax. Applied to our case, the loss of lexicality would be seen in the transition from full verb to auxiliary:
an auxiliary is generally deemed to be more of a functional, rather than a lexical element. 4

(10) Grammaticalization cline
full verb > (light verb) > auxiliary > clitic > affix

While there have been new developments in the theory by which grammaticalization is also sometimes assumed to be linked to the appearance of more semantic or pragmatic content, rather than less (see Traugott 1999 for some discussion), for our purposes the theory as summarized above is relevant. Note that according to this grammaticalization cline the light verb stage is optional, for clearly main verbs may develop directly into auxiliaries and from there into clitics and affixes.

We argue in this paper that light verbs can be clearly distinguished from main verbs in Urdu, yet this distinction is not to be taken as an indication that they have begun to undergo grammaticalization. While the grammaticalization cline appears to hold true for the development of auxiliaries from full verbs and their further slide into more and more of a functional rather than lexical status, light verbs will be found to differ markedly from this picture. In particular, the prediction that light verbs develop further into auxiliaries cannot be confirmed.

2.2.2. Semantic bleaching and metaphor

On the semantic side, grammaticalization is often characterized as going hand-in-hand with a process of “bleaching”, i.e., a loss of meaning that might be conceived as a loss of semantic features of a word. Loss of content alone is not the whole story, however. It has been pointed out by many authors that a process of metaphorization plays an important role, as well. An approach that is of special interest to us is the proposal made by Sweetser (1988). The overall structure of her account will be found to share some similarities with ours, though we do not agree with some of her basic assumptions.

4 This view of progressive grammaticalization of verbs also fits in with Ramat’s (1987: 8–13) four-stage universal process of “auxiliarization” by which one begins with full verbs, moves on to periphrastic auxiliary constructions via a predicative stage, and from there on to the reduction to a morphological affix.
Sweetser’s (1988) account is embedded in the cognitive-semantic view that there is a realm of so-called meaning schemata, which serve to structure lexical meanings and also occur, in a pure form, as “grammatical meanings”, i.e., the interpretation of grammatical categories. For example, the grammatical category of tense structures time via a schema of one-dimensional succession, and the same schema is shared with the representation of space-time that underlies lexical verbs of movement. Metaphor, on Sweetser’s account, is defined as the transfer of one and the same schema across different conceptual domains (e.g., from space to time). This is why verbs of movement (like ‘go’) can develop into markers for (future) tense. Metaphor in the lexical domain is viewed as a process of extracting the schema that structures some lexical concept, and transferring it to another domain of lexical meanings where it serves to structure new and different conceptual contents. Grammaticalization is said to consist in the use of a schema, extracted from a lexical content, as a grammatical schema. The difference between lexical metaphor and grammaticalization therefore lies only in the fact that, in lexical metaphor, a schema that is transferred to another concept is “fleshed out” anew with conceptual contents of the new domain, while in the case of grammaticalization the schema remains what it is, a purely schematic representation.

We need not take detailed issue with the theory of metaphor in Sweetser’s account, but it must at least be made clear that metaphor does not generally involve semantic bleaching: in productive, creative lexical metaphor, it is clearly felt that the metaphor adds a secondary structuring onto a domain that is already structured (cf. Indurkhya 1994). If we say, for instance, *Our linguistics department is a kindergarten*, then the original meaning of ‘kindergarten’ is actually accessible as a whole and can be used to elaborate the metaphor in creative ways, e.g., by paralleling not only the linguists to children, but also the computer equipment to toys, or academic exhibitionism to childish playing, etc. Furthermore, it is felt that this is a metaphor, i.e., a secondary and contrived structuring of a concept, unlike the (persisting) original one.

The semantic developments in grammaticalization could therefore indeed be loosely identified as “metaphorical”, in the sense that transfer of a schema to a new cognitive domain is involved. Nevertheless, we would prefer to introduce a finer distinction between two subtypes of the transfer of schemata, keeping the notion of metaphor distinct from “pure schema transfer”. On the one hand, there is lexical metaphor, i.e., the identification of a schematic similarity shared by two fully lexical words. Lexical meta-
phon does not “destroy” the original lexical sense of a word – it just posits a similarity relation between two lexical meanings, which consists in their sharing a certain implicit schema. Another process can now be more narrowly identified as a pure schema transfer: it provides a first-hand structuring for a new domain and thus requires the creation of a new linguistic item with an entirely schematic meaning.

2.2.3. Syntactic fusion

Grammaticalization is a process that unites changes in content (of the kind sketched above) with changes on the formal side. Typical results of grammaticalization are the prosodic and phonological reduction of the items in question and, ultimately, their morphological attachment to some grammatical host. With respect to light verbs one would expect that these would also be undergoing a partial loss of their syntactic independence. The claim that light verbs arise via grammaticalization thus calls for an account of their diachronic syntax. The obvious historical scenario for our case would be a development by which one verb is semantically weakened and thus loses its ability to project its own domain of predication; hence the syntactic development might involve the fusion of two domains of predication to a single one. This is the approach usually taken in historical work on the development of auxiliaries. (Recall that in the grammaticalization cline of Hopper and Traugott 1993 cited above, light verbs are indeed assumed to be a possible source for developing auxiliaries).

Syntactic accounts of relevant cases have been put forward, for example, by Roberts (1993), Roberts and Roussou (1999), or Harris and Campbell (1995). Roberts (1993) and Roberts and Roussou (1999) propose a view of syntactic change which assumes a basic mechanism whereby syntactic structures tend towards “simplification”. A part of what triggers the simplification of complex structures is the reclassification of a formerly lexical item as being a functional element. In sixteenth century English, for example, the modal mote is analyzed as being a fully lexical item that needed to move into the functional position T for various language internal
reasons, as shown in the constructed example in (11) (Roberts and Roussou 1999: 1023).5

(11) a. The kynge mote speken.

b. [TP The kynge [T mote] [VP t_mote [TP np speken]]

c. [TP The king [T must] [vp speak]]

In modern-day English, essentially the same surface string of words has been reanalyzed as in (11c), whereby a formerly biclausal structure has become a monoclausal structure in which the modal is simply generated in T, rather than being moved there every time. Hence, the essential historical mechanism which triggers a “clause union” type of effect is viewed as the process of reanalysis of a syntactic structure; the nature of the semantic change involved in the development of a formerly full lexical item into a functional element is not the subject of investigation.

An interesting prediction ensues from the work on historical change of Harris and Campbell (1995). These authors formulate a number of principles for a wide range of syntactic developments. Among them is the phenomenon of “clause fusion”, which in principle could be taken to apply to the development of complex predicates, where two or more predicates jointly contribute to the predicational structure of a clause. Harris and Campbell, too, see clause fusion as something which serves to simplify biclausal structures and posit that the right way to think about the formation of monoclausal sentences from formerly biclausal ones is in terms of a historical process of reanalysis, by which the surface appearance of the sentence is not changed, but the underlying structure is taken to be a different one than before.6 Their theory makes certain predictions about the grammatical properties of the reanalyzed construction. In particular, they posit the following principle, called the “Heir-Apparent Principle”:

5 In this example, np indicates some type of dependent silent element (like PRO), whose precise nature Roberts and Roussou (1999) do not elaborate on.
6 Roberts and Roussou (1999) follow Lightfoot (1979) in assuming that syntactic change takes place cataclysmically, rather than gradually, or via a period of variation and concurrent structural ambiguity. Since we have no direct evidence that can weigh in directly on either side of this long-standing debate, we do not address this question within the scope of the paper.
When the two clauses are made one by diachronic processes, the main verb governs the syntax of the reflex clause.

A verb is taken to govern the syntax of the clause if “(a) it determines the number of arguments, the argument roles they fill, and the marking they bear, (b) it determines whether any lexically-conditioned obligatory synchronic rules are conditioned (for example, Inversion), (c) it determines whether the clause can undergo optional synchronic rules (for example, Antipassive), and (d) it determines any exceptional behavior (for example, Quirky Case, suppletion).” (Harris and Campbell 1995: 191). In a newly created auxiliary construction, the new “main verb” is the former subordinate verb. The point of the principle is thus that it predicts semantic role selection and case assignment to be controlled by the former subordinate verb. In contrast, the old matrix verb is predicted to lose these abilities in the process of reanalysis to an auxiliary.

As stated so far, this point of view could indeed apply to the formation of complex predicates, and, hence, light verbs. However, all the examples cited by Harris and Campbell involve cases in which full verbs develop into auxiliaries. So the question that is of central interest now is whether light verbs can be aligned with auxiliaries with respect to their diachronic and syntactic properties. In the following, we attempt to provide both historical and modern data in order to make a case for our own perspective, namely: a) light verbs must be clearly distinguished from auxiliaries both synchronically and diachronically; b) light verbs in comparison with auxiliaries show relative historical stability; in so far as developments can be found, they do not fit into the picture sketched above for auxiliaries; c) these differences can be attributed to the rather special semantic function of light verbs.

3. The grammar of light verbs and auxiliaries in Urdu

As mentioned at the beginning of this paper, it has been a popular move to regard light verbs as a subtype of auxiliary. In this section, we show that auxiliaries and light verbs in Hindi/Urdu synchronically display clear behavioral differences. In a further section, these findings are tied to differ-
ences in the historical development of light verb and auxiliary constructions.

3.1. An introduction to Urdu/Hindi auxiliaries

In order to be able to make the comparison between auxiliaries and light verbs in Urdu more immediately accessible to the reader not familiar with this language, the table in (13) shows the forms that are clearly auxiliaries in Urdu.\(^7\)

(13) Urdu auxiliaries

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
<th>Inflection</th>
<th>Defective cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>hoo-naa</td>
<td>‘be’</td>
<td>Pres/Fut/Impf/Perf</td>
<td>Past</td>
</tr>
<tr>
<td>(t)-aali/ee</td>
<td>‘be’ (Orig. ‘stand’)</td>
<td>Past</td>
<td>All Others</td>
</tr>
<tr>
<td>rah-naa</td>
<td>Progressive (Orig. ‘stay’)</td>
<td>Perf *</td>
<td>Pres/Past/Prog *</td>
</tr>
</tbody>
</table>

* (Fut/Impf only with differing morphology on the main verb)

Some explanation of these forms is in order. Synchronically, all the ‘be’-auxiliaries are regarded as belonging to the same paradigm, whereby \(t\)-functions as the past form. However, as is evident from differences in agreement properties (the hoo-auxiliaries inflect according to tense, person, and number, while \(t\)-inflects according to gender and number), these auxiliaries really represent a case of a suppletive paradigm. The fact that the auxiliaries inflect differently can be traced directly to their historical origins: the hoo auxiliary is the descendant of the verb ‘be’, while the \(t\)-is derived from the past participle form of the Sanskrit verb \(sth\text{\`a} \) ‘stand’. Both of these ‘be’-forms are used for the formation of periphrastic tenses (e.g., the past perfect) and as a copula.

The progressive rah is form-identical with the full verb ‘stay’. It is not a light verb because its syntactic distribution differs from that of light verbs and it has a defective paradigm, unlike light verbs. The morphological inflection on the auxiliary rah is identical to perfective and imperfective morphology. These bits of morphology have been reinterpreted to denote

\(^7\) Modals such as ‘can’, ‘want’, etc., behave much like main verbs in Urdu from a syntactic point of view and are thus kept out of the discussion here.
progressive action in the case of the perfective morphology and repeated action in the case of the imperfective morphology. Unlike auxiliaries, light verbs always span the entire verbal paradigm, as shown in table (14) for the third person masculine singular form of the complex predicate *gir jaa* - ‘fall go’, which we render as ‘fall down’. Here *gir* ‘fall’ is the main verb and *jaa* ‘go’ is the light verb (note: the only verb that expresses the present tense morphologically in Urdu/Hindi is *hoo* ‘be’). Compare this table with the table for the simple verb *gir* ‘fall’ in (15).

(14)

<table>
<thead>
<tr>
<th>Pres</th>
<th>Past</th>
<th>Future</th>
<th>Imperfect</th>
<th>Perfect</th>
<th>Prog</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>gir gayaa</em></td>
<td><em>gir jaegaan</em></td>
<td><em>gir jaataa</em></td>
<td><em>gir gayaa</em></td>
<td><em>gir jaa rahaan</em></td>
</tr>
<tr>
<td></td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td></td>
</tr>
</tbody>
</table>

(15)

<table>
<thead>
<tr>
<th>Pres</th>
<th>Past</th>
<th>Future</th>
<th>Imperfect</th>
<th>Perfect</th>
<th>Prog</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>giraan</em></td>
<td><em>giregaan</em></td>
<td><em>girtaa</em></td>
<td><em>giraan</em></td>
<td><em>gir rahaan</em></td>
</tr>
<tr>
<td></td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td></td>
</tr>
</tbody>
</table>

Thus, light verbs are not restricted to appear with just one tense or aspectual form and never display a defective paradigm, which is again unlike

---

8 The tense/aspect system of Urdu/Hindi is rather complex as it also includes, for example, the possibility of reduplication for repeated action. It also includes the constructions in (i) and (ii), pointed out by a reviewer as a potential counterexample to our claim that light verbs do not display a defective paradigm.

(i)    *tum*      *likh-3-ee*    *jaa-oo-g-ee*
      you    writing-Impf.M.Pl   write-2.Sg-Fut-M.Pl
      ‘You will go on writing.’

(ii)  *naadyaa*  *aa-yaa*  *kar-ee-g-ii*
      Nadya  come-Perf.M.Sg   do-3.Sg-Fut-F.Sg
      ‘Nadya will keep dropping by.’

These constructions certainly do not fall under the type of V-V complex predication we have discussed. They are probably not instances of complex predicate formation at all. However, a complete understanding of this and other complex ways of expressing tense/aspect distinctions in Urdu/Hindi falls outside of the scope of this paper; it is far from being a simple task.
what one might expect of an auxiliary. Furthermore, as shown in the next subsection, auxiliaries and light verbs show distinct syntactic behavior with regard to case marking, word order, reduplication, and topicalization. The following list shows most of the verbs that commonly occur in the function of a light verb:

(16) Common light verbs

<table>
<thead>
<tr>
<th>verb</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>lee</td>
<td>‘take’</td>
</tr>
<tr>
<td>d5aal</td>
<td>‘put’</td>
</tr>
<tr>
<td>de</td>
<td>‘give’</td>
</tr>
<tr>
<td>aa</td>
<td>‘come’</td>
</tr>
<tr>
<td>par5</td>
<td>‘fall’</td>
</tr>
<tr>
<td>jaa</td>
<td>‘go’</td>
</tr>
<tr>
<td>bEt5</td>
<td>‘sit’</td>
</tr>
<tr>
<td>nikal</td>
<td>‘emerge’</td>
</tr>
<tr>
<td>ut5</td>
<td>‘rise’</td>
</tr>
<tr>
<td>maar</td>
<td>‘hit’</td>
</tr>
<tr>
<td>mar</td>
<td>‘die’</td>
</tr>
</tbody>
</table>

3.2. Grammatical properties of Urdu light verbs

This section illustrates several differences between auxiliaries and light verbs and establishes light verbs as a distinct class of their own.

3.2.1. Case

One difference between light verbs and auxiliaries is that light verbs in Urdu/Hindi have an influence on the case marking of the subject that goes beyond the usual type of effect observed with split-ergativity (see Butt 1995). As can be seen in (17), the verb ‘write’ usually takes an ergative subject when formed with the perfect tense, regardless of the auxiliary used.9 Light verbs like par5 ‘fall’ in (18), on the other hand, have the ability to require that the subject be nominative, even in the perfect.

---

9 The progressive rah cannot be used in this test as per definition the sentence then ceases to be in the perfect.
(17) a. \textit{us=nee} \quad /*\textit{voo} \quad \textit{xat} \\
letter.M.Nom \quad lik^{h}-aa \\
\textit{write-Perf.M.Sg} \\
\text{‘He wrote a letter.’}

b. \textit{us=nee} \quad /*\textit{voo} \quad \textit{xat} \\
letter.M.Nom \quad lik^{h}-aa \quad (\textit{hE} \quad it^{h}-aa) \\
\textit{write-Perf.M.Sg} \quad \textit{be.Pres.3.Sg} \\
\text{‘He wrote (has/had written) a letter.’}

(18) a. \textit{*us=nee} \quad /\textit{voo} \quad \textit{xat} \\
letter.M.Nom \quad lik^{h} \quad par5-aa \\
\textit{write} \quad \textit{fall-Perf.M.Sg} \\
\text{‘He fell to writing a letter.’}

b. \textit{us=nee} \quad /*\textit{voo} \quad \textit{xat} \\
letter.M.Nom \quad lik^{h} \quad lii-yaa \\
\textit{write} \quad \textit{take-Perf.M.Sg} \\
\text{‘He wrote a letter (completely).’}
3.2.2. Reduplication

In complex predicates the light verb, the main verb, or both may be reduplicated (see Fitzpatrick-Cole 1994, 1996 for a detailed study of the sister language Bengali). Example (19b) illustrates reduplication of the light verb.

(19) a. *voo soo jaa-tii t^h^-ii
   Pron.3.Sg.Nom sleep go-Impf.F.Sg be.Past-F.Sg
   ‘She used to go to sleep.’

   b. voo soo jaa-tii
      Pron.3.Sg.Nom sleep Impf.F.Sg
      vaa-tii t^h^-ii
      go.Redup-Impf.F.Sg be.Past-F.Sg
      ‘She used to keep going to sleep (at inopportune moments).’

An auxiliary may not be similarly reduplicated, as (20) and (21) show.

(20) a. voo soo rah-ii t^h^-ii
   Pron.3.Sg.Nom sleep Prog-F.Sg
   be.Past-F.Sg
   ‘She was sleeping.’

   b. *voo soo rah-ii vah-ii t^h^-ii
      Pron.3.Sg.Nom sleep Prog-F.Sg
      Prog-Redup be.Past-F.Sg
      ‘She was sleeping.’

(21) a. voo soo-tii t^h^-ii
   Pron.3.Sg.Nom sleep-Impf.F.Sg
   be.Past-F.Sg
   ‘She used to sleep.’

   b. *voo soo-tii t^h^-ii s&-ii
      Pron.3.Sg.Nom sleep-Impf.F.Sg be.Past-
      F.Sg be.Redup-F.Sg
      ‘She used to sleep.’
3.2.3. Word order

The order of constituents within a clause is generally quite free in Urdu/Hindi and is determined on the basis of discourse notions (e.g., see Kidwai 1997, Butt and King 1996). Word order within the verbal complex, in contrast, is very rigid. The main verb comes first, optionally followed by a light verb, a passive auxiliary, the progressive rah, and a ‘be’ auxiliary, in exactly that order: Main verb (light verb) (passive) (progressive) (be auxiliary).

Even further evidence for a clear difference between auxiliaries and light verbs comes from topicalisation. As shown in (22a–b), a main verb may be topicalised away from the light verb:

(22) a. *baccaa soo ga-yaag PerMg
    child.M.Nom sleep go-
    ‘The child has gone to sleep.’

b. soo too baccaa ga-yaag go-PerMg
    sleep Top child.M.Nom
    ‘The child has gone to sleep.’

However, a main verb may not be topicalized away from a cluster of auxiliaries, as (23) illustrates:

(23) a. soo too baccaa ga-yaag go-PerMg
    sleep Top child.M.Nom
    ‘The child has gone to sleep.’

b. *soo too baccaa rah-aa hE be.Pres.Mg
    sleep Top child.M.Nom Prog-
    ‘The child is sleeping.’
3.2.4. Conclusion

While it is not immediately possible to derive a particular structure for the light-verb construction from the above observations, the data at least point towards the conclusion that light verbs form a distinct class in the synchronic syntax of Urdu and other Indo-Iranian languages. In particular, they do not fall into a class with auxiliaries. Light verbs are not exactly like main verbs either: they do not have the same meaning as the full verb version, and they occupy a special syntactic slot in Urdu/Hindi. Still, they pattern with fully lexical verbs in a number of important grammatical properties such as their ability to undergo reduplication and to govern the assignment of subject case.

The finding that light verbs impinge on case assignment is very important because it runs counter to the “Heir-Apparent-Principle” stated in section 2.2.3. above: If light verbs were derived via grammaticalization from a biclaual embedding construction in exactly the way in which auxiliaries are derived in Harris and Campbell’s examples, this principle would predict that in the end only the main verb would be responsible for the government of the subject case. However, we need not necessarily conclude that Harris and Campell’s principle has been falsified, rather we can point out that the distinction between light verbs and auxiliaries may be of critical importance here. The Heir-Apparent-Principle is presumably tailored to the diachrony of auxiliaries and therefore does not apply with respect to our data.

4. The diachronic issues

In addition to the synchronic differences between light verbs and auxiliaries observed above, historical processes also seem to be able to distinguish between the two categories quite clearly: auxiliaries cheerfully enter the grammaticalization cline and undergo a comparatively rapid loss of independence. Light verbs, on the other hand, are found to resist grammaticalization. As an illustrative example for the grammaticalization of auxiliaries that works as expected theoretically, we present the case of the Urdu future inflectional morphology first.
4.1. Grammaticalization: The diachrony of the Urdu future

Sanskrit expressed tense and aspect via a very complex inflectional system. This inflectional system has been lost and replaced by a mixed periphrastic and inflectional system in modern-day Urdu. In particular, Urdu appears to contain no real tense inflections, with the exception of the future.

An overview of the basic tenses/aspects in modern Urdu is given in table (24) for the 3rd singular masculine form of the verb maar ‘hit’.

(24)

<table>
<thead>
<tr>
<th>Pres</th>
<th>Past</th>
<th>Future</th>
<th>Imperfect</th>
<th>Perfect</th>
<th>Prog</th>
</tr>
</thead>
<tbody>
<tr>
<td>maaraa</td>
<td>maarega</td>
<td>maar</td>
<td>maaraa</td>
<td>maar rahaa</td>
<td></td>
</tr>
<tr>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td>+ Aux (be)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The past, imperfect, perfect and progressive forms by themselves only inflect for number and gender. The auxiliaries either inflect for person and number, or for number and gender, depending on which of the ‘be’ auxiliaries is used. The future is not used in combination with auxiliaries and by itself inflects for all three features: number, gender, and person. In addition, the -g- form appears to signal the future: this piece of morphology remains invariant throughout the paradigm, as shown in the table below.

(25)

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>maar-u$u$-g-i/aa</td>
<td>maar-e$e$-g-iSi$/ee</td>
</tr>
<tr>
<td>2nd (tum)</td>
<td>maar-o-o-g-i/aa</td>
<td></td>
</tr>
<tr>
<td>3rd</td>
<td>maar-e-e-g-i/aa</td>
<td>maar-e$e$-g-iSi$/ee</td>
</tr>
<tr>
<td>aap (polite you)</td>
<td>maar-e-e-g-i/aa</td>
<td></td>
</tr>
<tr>
<td>tuu (rude you)</td>
<td>maar-‘hit’</td>
<td></td>
</tr>
</tbody>
</table>

The consensus in the available literature is that the future -g- morpheme is derived from a Sanskrit participle of the verb gā ‘go’ (Kellogg 1893: 231, Beg 1988: 191, McGregor 1972).

(26) Development of the Urdu future marker
Sanskrit gata5’ > Prakrit *gao > ... > Modern Urdu/Hindi -g-
The gender number agreement (ii/aa/i$i$/ee) of the future can thus be associated with the gender and number agreement generally found with old participles (such as the perfect and the imperfect). The vowel following the stem and preceding the -g- morpheme indicates person agreement. According to McGregor (1968: 161) this morpheme represents the leftovers of the hoo 'be' auxiliary that was trapped between the stem form and the past participle of 'go' when the periphrastic construction was formed. Support for the analysis which posits a periphrastic construction as the origin for the Urdu inflectional future comes from the fact that the emphatic particle hii could intrude between the suffix and the rest of the verb in the middle ages (Kellogg 1893: 231).

(27) maan-ee-hii-gii
    heed-3.Sg-Emph-Fut.F
    '(She) will see reason.'

The Urdu future thus appears to have conformed to the grammaticalization cline posited by Hopper and Traugott (1993) in that a participle form of the Sanskrit verb 'go’ first formed a periphrastic construction with another verb and then slowly grammaticalized down to an affix. However, also note that the same Sanskrit verb gā/gam ‘go’ is the ancestor of the Urdu/Hindi main verb jaa (perfect form: ga-) ‘go’. As we have seen in previous sections of this paper, the main verb jaa ‘go’ also has a light verb use, as in (28), repeated from above.

(28) naadyaa   gir   ga-yii
    Nadya.F.Sg.Nom    fall    go-Perf.F.Sg
    ‘Nadya fell (down).’

And at this point, a stark contrast between auxiliaries and light verbs emerges: the futurate ‘go’ split off from the main verb ‘go’ and entered a very individual line of development that resulted in an affix – but the light verb use of the same verb ‘go’ continues to be form-identical to the full lexical verb and does not appear to have undergone any separate historical development. Further and more detailed evidence along these lines can be found in Lahiri (2001) and Butt and Lahiri (2002).
Light verbs in Urdu and grammaticalization

Light-verb uses in general appear to have existed side-by-side with their main verbs for quite a number of verbs derived from Sanskrit (as will be shown in the next few subsections). Thus, only the future auxiliary could be taken to represent a well-behaved case of grammaticalization. Our initial inspection of the light verb data in section 2.1 might have left room for the assumption that light verbs behave the way they do simply because they are still in an incipient stage of grammaticalization. However, the difference in the pace of development between futurate uses and light uses of one and the same Sanskrit verb ‘go’ strongly discredits such an assumption. Instead, the diachrony of light verbs indicates a qualitatively different behavior.

4.2. The ancestry of the Urdu light-verb construction

This section compares the usages of the modern V-V light verb construction with the usage of its ancestral construction in Sanskrit and Prakrit and its development and use on down the ages, as far as it can be documented. For readers who are unfamiliar with the linguistic time line in South Asia, the major linguistic eras are usually divided up into three main times. The oldest attested stage of Sanskrit is Vedic. Sanskrit and Vedic are often grouped together under the label Old Indo-Aryan (1200 BCE – 600 BCE). Middle Indo-Aryan (600 BCE – 1000 CE) is attested via the inscriptions of the Emperor Aśoka, early Buddhist writings in Pāli, various different Prākrits (basically “vernacular” forms of Sanskrit), and the Apabhramśa. New Indo-Aryan is then dated from about 1000 CE on and includes modern (north) Indian languages such as Bengali, Hindi, Urdu, Marathi, Punjabi, or Gujarati.

The modern main-verb form found in complex predicates originates in an indeclinable Sanskrit participial form generally referred to as the “gerund” (Chatterji 1926, Hook 1991). This gerund was formed via the suffixation of tvā or (t)yā/(t)ya on a verb stem (see Whitney 1889: 345–360, Tikkanen 1987 for details on the morphological formation). As one main use of the gerund was a type of clause chaining, exemplified in (29), this suffix is also sometimes referred to as a conjunctive participle (e.g., Hook 1991).
The use of *tvā* was not confined to this type of clause chaining, but was manifold and varied. Indeed the discussions surrounding its usage, which goes hand in hand with some syntactic peculiarities, and the attempts to pin down its elusive and varied temporal and aspectual meaning are strongly reminiscent of the discussions surrounding V-V complex predicates in many of the grammars of the modern Indo-Aryan languages. Tikkanen (1987) uses the constructed example in (30) to illustrate the wide variety of interpretations one and the same V-*tvā* V sequence could have in Sanskrit.10

(30) *indram árabhya cara*

Indra.Acc grasp.Gd go.Imp.2.Sg

‘Having taken hold of Indra, move!’ […]
‘Take hold of Indra and move!’
(‘Move by taking hold of Indra!’)
(‘Go to take hold of Indra!’)
‘Keep yourself to Indra!’
(Sanskrit, Tikkanen 1987: 7)

One of the interpretations, namely the last one, shows a “light” usage of the verb ‘go’ and cohesion of meaning which is immediately reminiscent of the modern V-V complex predicates.

10 A reviewer points out that the first four translations are actually truth-conditionally identical. This does not detract from the fact that the same sequence could be interpreted in a number of ways.
Like in Sanskrit, the V-V string can be interpreted in various ways in Urdu (this is taken up in some more detail in section 4.3). Unlike in Sanskrit, the first verb in the sequence contains no special morphological marking. Compare, however, the corresponding examples in (31) from a modern sister language, namely Bengali.

\[(31)\]
\begin{align*}
a. & \quad \text{ram} & \text{bag}^{h\cdot t5a-ke} & \text{mer} & \text{p}^{h}\text{el-lo} \\
& \quad \text{Ram.Nom} & \text{tiger-Class-Acc} & \text{hit-Perf} \\
& \quad \text{throw-Past.3} \\
& \quad \text{‘Ram killed the tiger.’} \\
& \quad \text{(Bengali)} \\
\end{align*}

\begin{align*}
b. & \quad \text{ram} & \text{eb}^{e}\text{-e} & \text{por-l-o} \\
& \quad \text{Ram.Nom} & \text{come-Perf} & \text{fall-Past.3} \\
& \quad \text{‘Ram arrived.’} \\
& \quad \text{‘Having come, Ram fell.’} \\
& \quad \text{(Bengali)} \\
\end{align*}

Here the first verb in the sequence is marked with an -e, which is commonly glossed as perfective. Thus, what we find in the modern languages is either a null form, as in Urdu/Hindi, or a form that has been reduced to a vowel. The path of morphological development can be sketched as follows (Chatterji 1926, Beg 1988: 185, Kellogg 1893: 341). The Sanskrit participial suffix in \(nv\) had several allomorphs, among them \(ya\) and \(ya\). The glide in these allomorphs was often reduced to \(ia\) in the Prakrits. In Modern Colloquial Bengali the surviving vowel is realized as \(e\), while the marking was lost in Urdu/Hindi.

Besides the morphological changes that have taken place, apparent changes along the syntactic and semantic dimensions must also be acknowledged. In terms of the syntax, the adjacency requirement between the two verbs has become stricter. Whereas in Sanskrit word order was fairly free and discontinuous constituents were common, modern Urdu/Hindi allows discontinuous constituents in only some parts of the grammar. V-V sequences are separable (this shows that they are not items formed in the lexicon, but are formed syntactically, see Butt 1995), but they may only be separated by emphatic clitics or by negation. The main verb may be topicalized away from the light verb, but not vice versa (see section 3.2.3). Thus the V-V sequences form a more cohesive unit in the modern language. However, it is not clear whether this greater cohesiveness stems
from progressive grammaticalization, as argued by Hook (1991), for instance, or whether they are symptomatic of a general loss of freedom in word order.\footnote{Note that Whitney (1879: 340–343) reports the number of gerunds (tvā constructions) to be three times greater in the Nala and Bhagavadvētā (Classical Sanskrit) than in the Úgveda (Vedic Sanskrit), which has ten times more verb forms. Thus the older form of Sanskrit, namely Vedic appears to rely less on the use of the gerund (Tikkanen 1987: 50). This is in line with a view in which the basic construction type was not “created” from Old to Middle Indo-Aryan, it was already in place, but gained more currency in usage over the ages.}

Having briefly given some background as to the known morphological evidence for the historical development of the ancient construction to the modern V-V complex predicate, we now turn to an examination of the syntactic (and semantic) peculiarities that characterized the construction synchronically in Sanskrit and compare these to the modern day situation. This discussion with respect to Sanskrit is based almost exclusively on the findings in Tikkanen (1987). More detailed discussion of the syntax and morphophonological development of these V-V sequences can be found in Butt and Lahiri (2002).

4.2.1. Clause chaining

As mentioned previously, the main use of the tvā construction appears to have been to effect a kind of clause chaining, whereby several differing events which stand in close connection to one another can be described as parts of a greater event. The Sanskrit example in (32) is repeated from above.
Light verbs in Urdu and grammaticalization

(32)  
...[uptvā keβaβmaβrÊ‘i]  

[naκhâni nik<yay>a]  
nail.Nom/Acc.Pl in.do.Gd  

[âjyâ{-}]  

bhyajya]  

...  

clarified-butter.Nom./Acc.Pl  

smear.Gd  

“... having shaved his hair and beard, cut his nails, smeared his eyes and anointed his body, ...”  

(Jaimin<yabarhama‘a, from Tikkanen 1987: 187)

Note that this type of usage is still current in modern-day Indo-Aryan languages. So much so that Chatterji (1926: 1011) marvels at the seemingly infinite number of clauses that can be strung together in Bengali. He cites an example with 15 subclauses, of which we only show 3 here. Note that the last two verbs are an instance of a complex predicate in which the verb ‘go’ is functioning as a light verb. The morphological form of the main verb ‘walk’ does not differ at all from the embedded verbs, however the syntactic function is very different: ‘walk’ is the main verb of the sentence

(33)  
b<ore  tu<e-e,  βtUNUSED βeOne  po$uc<e,  
early rise-Perf station reach-Perf  

†iki†  kor-e,  co<e  jeo...  
ticket buy-Perf walk-Perf go-Imp  

‘get up early, get to the station, buy a ticket and leave’  
[‘having gotten up early, having gotten to the station, having bought a ticket , leave...’]

An Urdu equivalent is given in (34).12

(34)  
sub<aa-sub<aa u<e kar,  βtUNUSED βeβan  pao$  

12 Note that a form of kar ‘do’ has been grammaticalized as a complementizer in order to replace the tvā in some instances. This kar is generally optional when the subordinated verb is adjacent to the main verb example.
early-early rise having, station.Nom reach

kar, Êïka† xariid kar, nikal jaa-oo having, ticket.Nom buy having,
emerge go-Imp.

‘Having gotten up early, having reached the station, having bought a ticket, leave.’

As was seen above and is repeated in (35) for Urdu and (36) for Bengali, all V-V sequences of the type under discussion here are in principle ambiguous between a complex predicate and a clause-chaining reading, however, for some V-V sequences the complex-predicate reading is the dominant one. This is true for the examples in (35) and (36). The native Urdu/Hindi speaker would therefore prefer to insert the optional kar ‘having’ between the two verbs in (35) and the native Bengali speaker would prefer to augment the second verb in the sequence with a light verb for the second ‘having’ reading in order to make clear that there are two main verbs involved.

(35) a. naadyaa=nee xat likÆ dii-yaay
    Nadya.F=Erg letter.M.Nom write
give-Perf.M.Sg
    ‘Nadya wrote a letter (completely).’
    ‘Having written the letter, Nadya gave it (to somebody).’

b. naadyaa aa bEtsÆ-ii
    Nadya.F.Nom come sit-Perf.F.Sg
    ‘Nadya has arrived.’
    ‘Having come, Nadya sat (down).’

(36) a. ram bagÆ-t5a-ke mer-e pÆel-lo
    Ram.Nom tiger-Class.Acc hit-Perf
throw-Past.3
    ‘Ram killed the tiger.’
In the absence of such preferred clues, the two readings – clause chaining vs. complex predicate – can nevertheless be disambiguated on the basis of prosodic and contextual information. In addition, the syntactic properties of complex predicates differ from those of a matrix verb that is modified by an adverbial clause (the clause chaining). In the clause-chaining reading, the embedded verb (‘write’ and ‘come’ in the Urdu examples) forms a prosodic and syntactic constituent with its object. As such, the collocation “letter-write” can be moved around as a constituent. In the complex-predicate interpretation, the two verbs form a prosodic and syntactic constituent (Lahiri and Fitzpatrick-Cole 1999, Butt 1995, Butt and King 1998).

Thus, the examples in (35) and (36) above can be disambiguated on the basis of prosodic and contextual information, and by applying synchronic syntactic tests. This type of disambiguation is difficult to perform with respect to Sanskrit. That it would be illuminating to be able to perform such disambiguations in the ancient language is evident from the vast amount of discussion the Sanskrit tvā gerund has generated (see Tikkanen [1987] for a very nice summary of the main issues and authors). For one, the interpretation of the gerund cannot always be rendered as simple chaining of clauses as in (32) above, where one action can simply be seen as preceding the next, and the next, and so on, until the sequence of actions is done. Tikkanen (1987: 104) identifies three general temporal interpretations of the gerund in relation to the main predication: simultaneous or overlapping, relative past (as in the clause-chaining effect seen above), or relative future. The precise interpretation of the gerund clause is tricky and differs from example to example as it is contextually dependent.\textsuperscript{13}

From our perspective, the possibility of a simultaneous or overlap reading would appear to be one of the crucial preconditions for complex predications as in ‘grasp go’ which can be interpreted as “keep oneself to somebody” (see example (30) above). In fact, the Sanskrit gerund shows some syntactic properties which would seem to indicate that a kind of “clause

\textsuperscript{13} Tikkanen tries to make the argument that the default interpretation should be taken to be the relative past, as in example (35) above.
union” or complex predicate would be an adequate analysis for the phenomena observed at least with respect to some of the available examples.

4.2.2. Complex predicates vs. adverbial clauses

The Sanskrit tvā gerund is argued to come from a frozen form of an original oblique nominal inflection (Tikkanen 1987, Whitney 1889), namely an instrumental singular of a stem in -tu- (Hendriksen 1944: 142). As a nominal construction, it did not allow for the overt expression of an agent (or whichever the highest argument is). As an adverbial clause, this requirement is reflected in the fact that the tvā clause is generally dependent on the matrix clause for the expression of the subject argument. This is reflected by word order distribution in Sanskrit, whereby the (internal) arguments and adjuncts of the tvā gerund were generally grouped so that they preceded the tvā gerund and formed a constituent with it, whereas the external argument was supplied by the matrix clause. As a nominal construction, it did not allow for the overt expression of an agent (or whichever the highest argument is). As an adverbial clause, this requirement is reflected in the fact that the tvā clause is generally dependent on the matrix clause for the expression of the subject argument. This is reflected by word order distribution in Sanskrit, whereby the (internal) arguments and adjuncts of the tvā gerund were generally grouped so that they preceded the tvā gerund and formed a constituent with it, whereas the external argument was supplied by the matrix clause. In a continuation of this ancient pattern, in the modern languages (e.g., Urdu/Hindi or Bengali), the adverbial clause clearly requires subject control of a PRO argument (cf. Mohanan 1994, Butt 1995).

However, there are some examples which indicate that the tvā gerund and the “matrix” verb in fact formed a single clause. These examples involve cases in which negation or wh-operators which are “embedded” in the tvā gerund are able to take scope over the entire clause (Tikkanen 1987: 16). An example with the wh-word kas is shown in (37), where the wh-word kas appears to be situated within the adverbial clause. This is unusual in that the subject of the adverbial clause is generally a PRO controlled by the matrix clause. In (37), however, the wh-word appears to be an argument of the gerund as well as of the matrix clause. One analysis that would get the clausal scope of kas right would be to assume a type of complex predication or clause union of the two predicates.

\[\text{14\ Note that the scenario presented here is a slight abstraction over some of the more murky details discussed at length in Tikkanen (1987).}\]
Thus, in both the ancient Sanskrit and the modern languages, the same surface string may prove to be either a case of complex predication, or a case of adverbial modification. Note that this is reminiscent of the situation in Romance languages where one and the same sequence of two verbs could either be a complex predicate formation or could represent a case of a matrix-verb embedding a VP (e.g., Rosen 1989, Alsina 1996, Andrews and Manning 1999). In Italian, for instance, clitic climbing provides one reliable test for complex predication.

Note that meaning alone is not necessarily a reliable indicator of the presence or absence of complex predication. In the Romance languages, the string *fare riparare* will always mean “make somebody repair something”, but depending on the precise syntactic environment (i.e., clitic climbing or not), the string of verbs will either form a complex predicate or a straightforward syntactic-embedding relationship. However, even from a purely semantic perspective, behavioral similarities can be identified between the ancient and the modern construction.

Tikkanen (1987: 92–95), for instance, observes that the gerund cannot be formed with atelic verbs such as ‘be worthy’ or ‘be jealous’, ‘be pleased’. This is again paralleled by the modern situation in Urdu as statives (atelic) generally do not felicitously combine with light verbs (see Butt 1995: 107–118 for some detailed discussion).
Recall Hook’s (1991) argument that the use of light verbs in modern Indo-Aryan languages should be viewed as an instance of aspectogenesis in the sense that the light verbs appear to denote perfectivity in the modern languages (we argue against this point of view explicitly in section 5). Now compare that claim with the following observation from Tikkanen (1987: 132): “The propensity of the past gerund to express perfective (bounded) situations in narrative discourse is therefore a consequence of its basically relative past tense rather than vice versa”. Again, it seems that the modern and the ancient constructions share a semantic characteristic in that both seem to be doing some kind of aspectual work that is difficult to pin down.

A similarity in use is also apparent from a comparison of the Sanskrit example in (39) with the modern Urdu in (40).

(39) \[ tato mak\text{\textregistered}i\text{ko	ext{"o}\text{"i}ya} gat\text{\textregistered} \]
then the fly-fly.Gd go.PastParticiple
‘then the fly flew away’

(Pan$\text{\textregistered}$cathantra 122 (ed. Kielhorn 1902: 91, 1.14), taken from Tikkanen 1987: 176)

(40) \[ kabutre \text{\textregistered}u\text{\textregistered} ga-ye \]
pigeon.M.Pl.Nom fly go-

Perf.M.Pl
‘The pigeons flew away.’

Finally, Tikkanen (1987: 174–175) observes a fluidity between concrete spatial meanings of the inflected verb and more abstract, or metaphorical
uses in Sanskrit that go hand in hand with a more restricted syntax. This observation again corresponds to the modern situation quite accurately.\textsuperscript{15}

Since these verbs also appear as regular main verbs, the only difference between a peripheral and a nuclear juncture may then be the level of abstraction. Thus in (402) the meaning of \textit{ās} ‘sit’ is conceived of as referring to a concrete rather than abstract state, whereas in (403) the demoted state is abstract rather than concrete. This difference in meaning is also reflected in tighter word order constraints for the periphrastic reading.

(Tikkanen 1987: 174–175)

\begin{center}
\textit{\textsuperscript{402} \textit{ās} \textit{villain} \textit{loosened} ‘May the villain sit, having loosened his belly, …’}
\end{center}

\begin{verbatim}
sit.3.Sg.Imp villain.Nom.Sg belly.Acc.Sg loosened.Gd
\end{verbatim}

\begin{center}
\textit{\textsuperscript{403} \textit{ās} \textit{woman} \textit{not} \textit{talk} \textit{food} ‘Therefore one should not talk with a woman wearing (blood-)stained clothes, one should not sit in her company and one should not eat her food, for she is dressed in (lit. sits having put on) the color of brahmin murder.’}
\end{center}

[Delbrück (1888: 408) “sie trägt die Farbe des Brahmanenmordes (eigentlich: sitzt da, nachdem sie angezogen hat)’]

To summarize, while many of the morphosyntactic properties of Sanskrit have generally not survived the journey through the millennia down to modern day Urdu, the similarities in syntactic and semantic distribution between the Sanskrit \textit{tvā} gerund and the modern day V-V sequences are indicative of a large degree of historical stability with respect to the fundamental ability to form complex predicates. For the sake of completeness,

\textsuperscript{15} Note that the close gloss in the cited examples has been supplied by us.
the next two sections briefly sketch the situation in Middle Indo-Aryan and the middle ages.

4.3. Middle Indo-Aryan: Pâli

Not surprisingly, the phenomenon of joint or complex predication can be found in Middle Indo-Aryan times as well with the tvā gerund. Hendriksen’s (1944) discussion of the gerund in Pâli again reveals a variety of usage that goes hand-in-hand with rather special syntactic properties. As in Sanskrit and in the modern languages of today, the gerund primarily seems to be used for the type of clause chaining above, whereby the embedded gerund takes on a type of adverbial function and describes an action that took place prior to the event of the inflected verb. However, the precise interpretation of the gerund in Pâli is again often dependent on contextual information.

Sometimes the gerund indicates what is simultaneous with the principal verb. In examining these instances it is necessary to proceed with the utmost caution. It must be borne in mind that the gerund is used very frequently in Pâli, and often when other languages would employ the present participle. When in the modern western European languages the present participle or other synonymous constructions are employed, it is because we think primarily of the action as taking place, whereas in Pâli, when the gerund is employed, stress is often laid on the commencing of an action. The example in Ja I 141,13 antevāsitāpaso rājānaḥ disvā pi n’eva vut†hāsi is most naturally translated thus: “although the pupil-ascetic saw the king, he did not stand up,” but the exact meaning is “although the pupil-ascetic had caught sight of the king, he ...”. Sentences where the gerund has an iterative meaning are also treacherous: Vin III 105,16 tam enaḥ gijjhā pi kākā pi kulalā pi anupatītīvā anupatītīvā pāsū–antarikāhi vi-
tudenti: “vultures, crows, and hawks pluck him between the ribs, pursuing him all the time.” On account of the iterative meaning of the gerund-āmrečīta, we are inclined to regard the two actions, that of the gerund and that of the sentence verb, as simultaneous; but in Pâli it is the single, not the iterative action, that is considered, and here anupatītīvā precedes the sentence verb in time.

(Hendriksen 1944: 113–114)
Thus Hendriksen (1944) addresses a question that also characterized much of the discussion of the corresponding Sanskrit construction, as summarized and elaborated on by Tikkanen (1987): what is the precise temporal interpretation of the gerund? Does it denote a simultaneous (overlap) action, or a temporal sequence?

One special morphosyntactic property of the gerund which indicates that some kind of clause union (Aissen and Perlmutter 1983) has taken place is the fact that when the verb marked by the gerund is used as a passive, the actual overt passive morphology may appear on the “matrix” verb instead, i.e., the verb which carries the inflectional morphology (Hendriksen 1944: 130).

With respect to lexical semantics, Hendriksen (1944: 113) observes that while the type of clause chaining seen above involves rather concrete meanings of the verbs, other constructions with the gerund exhibit more subtle interpretations of meanings that range over a temporal or otherwise more abstract or “light” usage of the inflected verb. Hendriksen (1944: 131) furthermore notes that some verbs occur with special frequency with the gerund, so that Geiger (1943: §174), for instance, is tempted to analyze these constructions as periphrastic auxiliary constructions in which the former “main” verb (the inflected one) has been demoted to a type of auxiliary. However, Hendriksen finds this problematic because the inflected verb still retains most of its meaning while also expressing things like the duration of the action – again like the modern situation. Some relevant examples with ‘give’, one of the verbs which occurs frequently in conjunction with a gerund, are shown in (41).

(41) a. ...assamapadamu ãnetvâ ãggîµ
...hermitage.Acc lead.Gd fire.Acc

katvâ adâsi
make.Gd give.Impf.3.Sg

‘... brought her to his hermitage and made a fire for her’
[‘having brought (her) to the hermitage, made a fire (for her)’]

(Pali, Jâtaka Tales, Sri Lanka, Hendriksen 1944: 134)
We agree with Hendriksen in that we would classify the verb ‘give’ in these examples as an instance of a light-verb use, rather than as an auxiliary.

4.4. The modern era: 1575–1600

By the middle ages, after New Indo-Aryan was already well underway, the light verb use was easily identifiable (McGregor 1968), so much so that the auxiliary hypothesis was now explicitly adopted (e.g., Meißner 1964). Two examples from Indrajit of Orchá’s prose (ca. 1600 CE) are given in (42) and (43).

(42) cori letu hai
be.Pres.3.Sg steal take.Impf
‘...he steals’
(McGregor 1968: 209–213)

(43) kâôhi lei
pull out take.Perf
‘(he) pulled out (with effort)’
(McGregor 1968: 209–213)

4.5. Summary

The available diachronic data thus allow us to conclude that the V-V type of complex predicate construction found in the modern language does not
appear to be a recent development, as would be expected under a grammaticalization approach. Rather, a construction that is similar to the modern construction both syntactically (clause-chaining and clause-union effects existing side-by-side) and semantically (“light”, elusive contribution of the tvā-marked verb) appears to have been in place in the language for at least two millennia. In no case was a light usage of a verb identified that was not form-identical to a full verb. This lack of development or grammaticalization stands in stark contrast to the process which produced the Urdu future affix -g- from the verb “go”. We therefore conclude that light verbs should be taken out of the grammaticalization cline entirely, as shown in (44).

(44) Revised grammaticalization cline
    full verb > auxiliary > clitic > affix
    | light verb

We instead propose that the light verb usage should be linked to the full verb reading via an analysis under which light verb usages are considered to be an instance of lexical polysemy. This idea is elaborated on in some detail in the next section.

5. The non-grammaticalization of light verbs: A semantic account

In this section, we offer some additional data that will lead us towards a better understanding of the function of light verbs in Hindi/Urdu. The consideration of the semantic function of light verbs ultimately leads to an alternative of the grammaticalization account that avoids its empirical problems.

5.1. Aspectual vs. lexical meaning

We first have to take issue with the thesis that Urdu light verbs are markers for the grammatical category of aspect, as proposed by Hook (1991) and Abbi (1999). There are reasons to believe that this is not quite correct. Let
Miriam Butt and Wilhelm Geuder

us consult a current definition for the functioning of perfective aspect, namely Comrie (1976), which is also used by Hook (1991):

(45) Perfectivity indicates the view of the situation as a single whole, it denotes a complete situation with a beginning, a middle and an end. Perfective verbs can be used to indicate the beginning (ingressive) or the end (completive) of a situation. (Comrie 1976)

While this definition is rather loose, it is still strong enough to allow us to detect counterexamples to the thesis that light verbs in Urdu are markers for perfectivity. Consider the examples in (46), involving when-clauses:

(46) a. mariam iimeel lik₆
    Miriam.F.Nom e-mail.F.Nom write
    rah-ii ṭ-ii jab viilii
    Prog-F.Sg be.Past-F.Sg when
    Willi.M.Nom
    kamre=kee andar aa-yaa
    room.M.Obl=Gen.Obl in come-
    Perf.M.Sg
    ‘Miriam was writing an E-mail when Willi came into the room.’

b. mariam iiimel lik₆ maar
    Miriam e-mail.F.Nom write
    hit
    rah-ii ṭ-ii jab viilii
    Prog-F.Sg be.Past-F.Sg when
    Willi.M.Nom
    kamre=keeanandar aa-yaa
    room.M.Obl=Gen.Obl in come-
    Perf.M.Sg
    ‘Miriam was dashing off an E-mail when Willi came into the room.’
Example (46a) shows a construction without a light verb. The head clause, containing a progressive marker, describes an extended situation of writing which functions as the time frame for the perfective jab (‘when’)-clause that marks ‘Willi’s coming in’; the latter must be the situation that is “viewed as a whole”, because it is embedded into another ongoing action. Example (46b) shows the same two clauses, but with the light verb ‘hit’ added in the head clause. Now, if the light verb were a perfective aspect marker, its occurrence in (46b) would be paradoxical. It is the when-clause that introduces the event that is viewed as a whole in this construction, but this is not the one that contains the light verb. Moreover, the light verb ‘hit’ not only occurs in the wrong clause, it is even embedded by a progressive, although the progressive is clearly an imperfective construction. Finally, the addition of the light verb changes the content of the event description in the head clause. The perfective variant of ‘to write’ would be expected to express a complete view on the situation of writing, as the definition above suggests; however, the combination ‘write-hit’ amounts to a new situation, namely that of sending off something. This is indeed a bounded situation, but this is due to the new lexical content that has been added. As already indicated, this predicate ‘write-hit’ is then put in the progressive.

We thus have to be careful with the notion of “aspect”. In light of the above data, we claim that what is involved with these light verbs is actually aktionsart (unfortunately also called “lexical aspect” by some) – but not aspect in the temporal sense, i.e., in the sense of locating an event relative to the reference time of the clause (like, for instance, “viewing it as a whole”) (cf. Klein 1994). Aktionsart is a type of lexically specified content of the verb which interacts with the aspectual semantics and which gives rise to effects like the specification of telicity. However, since light verbs perform a function on this lexical level, they can at the same time add lexical features like volitionality, force, benefaction, etc.

While the most common contribution of a light verb in Hindi/Urdu seems to be a notion of “completeness”, inceptive (inchoative) meanings can also be expressed, as in (47):

(47) a. naadyaagaa par5-ii
    Nadya.F.Sg.Nom sing fall-
    Perf.F.Sg
    ‘Nadya burst into singing (fell to singing).’

b. anjum has par5-ii
Note that, under a certain interpretation, the definitions of “perfectivity” and “perfective verb” in (45) above might be understood as contradicting each other: if a perfective verb is “used to indicate the beginning of a situation” (as stated in the second part), then this is apparently not the “complete view” of the situation promised in the first definition of “perfectivity”. The contradiction is only resolved by the fact that the ingressive meaning is taken to be a lexical content (i.e., a situation) of its own, and since this kind of situation is rather punctual, it is the only natural construal to render it as a perfective. This is to say, if a situation has no parts then there is usually no other choice than viewing it “as a whole”. The only alternative might be to force another interpretation. Moens and Steedman (1988) first pointed out that progressives derived from punctual verbs can only receive an aspectual interpretation via coercion: a preparatory phase has to be added to that punctual event. Over and above the aspectual meaning, however, Piñón (1997) points out that the interpretations which such progressives receive in English are unpredictable and conventionalized with respect to the conceptual contents that these preparatory phases are associated with (for instance, in the progressive form to be dying, a feature of ‘suffering’ shows up that is not necessarily associated with either the lexical content of the punctual verb die or with locating the reference time to a time prior to the dying). The example of English die, then, would be another case of a lexical content that gives rise to only one natural aspectual encoding, i.e., the perfective one. Still, the verb die as such is not perfective in the temporal sense, because we can in fact put it in the progressive. In a similar vein, we can assume that the light verbs in our Urdu examples change the lexical properties of the predicate in ways that naturally lead to a temporal encoding as a perfective. To repeat, this is different from the claim that light verbs simply denote a temporal relation.

Incidentally, there is another quirk in the grammar of Urdu that is difficult to reconcile with the aspectual hypothesis: light verbs can themselves be marked for morphological aspect (as already implied in section 3 above, where we said that they interact with the whole morphological paradigm of a verb instead of being defective). The meaning of light verbs with imperfective morphology cannot be derived compositionally in a straightforward manner, since modal or dispositional interpretations arise in this case. A
relevant example as discussed by Butt (1997) for the light verb *lee* ‘take’ is shown in (48b). Contrast this with the simple main verb ‘drive’ in (48a).

(48)  

a. *naadyaa gaarii calaa-tii hE*

   drive-Impf.F.Sg be.Pres.3.Sg

   Nadya.F.Sg.Nom car.F.Sg.Nom

   ‘Nadya drives a car.’

b. *naadyaa gaarii calaa lee-tii hE*

   drive take-Impf.F.Sg be.Pres.3.Sg

   Nadya.F.Sg.Nom car.F.Sg.Nom

   ‘Nadya does/will drive a car (has the ability to drive a car).’

In (48b) the light verb is marked with imperfective morphology and there is no sense of a perfective reading. Rather, the clause expresses a situation in which Nadya is described as being able to drive a car.

If it is true that slight effects of non-compositionality can already be detected in English progressives from punctual verbs (like *die*), then such a finding would not be altogether unexpected; the case of Urdu imperfective light verbs would then just exhibit a difference in degrees. But on the whole, it could fit in with the assumption that light verbs have lexical content (even if it is a very weak one). On the other hand, aspectual morphology on light verbs would have to be viewed as a fossilized morpheme without semantic import, if the light verb as such were an aspectual marker.

Thus, the major problem with the aspectual hypothesis is that the contrasts we get with the presence or absence of light verbs are not purely aspectual in nature; it is always features of the lexical content that make the difference. Besides changes in aktionsart, there are also other semantic notions that can be expressed by light verbs, like benefaction, force, rapidity, or intentionality. The examples with the light verb *maar* ‘hit’ in (46) have already shown that the lexical meaning of the predicate is affected, not just the aspectual semantics: the addition of ‘hit’ to the verb ‘write’ produced a meaning like ‘dash off (a letter)’. Admittedly, the semantic effects of a light verb are sometimes difficult to discern. Still, semantic features can be detected if the effects of adding different light verbs are compared for one and the same context. Take for instance (49) and (50):
Again, the specific meaning contribution of ‘give’ can be identified as a weak lexical feature, something like the “forcefulness” of the action. The light verb ‘take’, in comparison, does not carry this implication. A slight problem is that ‘take’ seems to be just unmarked in this pair. The existence of an unmarked light verb might at first be taken as evidence against the claim that light verbs always convey lexical contents (while it would be the ideal example for the aspectual hypothesis). However, note that there are particular contexts in which even ‘take’ is excluded, and if this happens, it is for reasons of lexical-semantic features:

(51) a. *nadyaa=ne gir lii-yaa
    Perf.M.Sg
    ‘Nadya fell (on purpose).’

    b. nadyaa gir parii-
    Perf.F.Sg
    ‘Nadya fell (suddenly).’
The explanation given in Butt (1995) for this contrast is that the light verb ‘take’ carries a requirement that the action in question results from conscious choice, which is incompatible with the lexical meaning of the Urdu verb ‘fall’. Therefore, the light verb ‘take’ is not void of lexical features, it is simply so general that it does not add new information in many of its occurrences.

On the whole, then, the traces of meanings that can be found with light verbs conform to the findings concerning their syntactic status: with respect to the divide between lexical and grammatical elements, light verbs are on the lexical side.

5.2. Full and light verb meanings

In the preceding section, we reached the conclusion that light verbs convey lexical meanings, and do not have a grammatical (temporal) function. Therefore, light verbs and their form-identical twins, the main verb uses, constitute pairs of lexical variants. We are thus claiming that these pairings are not a product of grammaticalization but represent a case of lexical polysemy. This is, first and foremost, a synchronic relation between two word meanings. The term polysemy implies that this relation is systematic in a certain way, in distinction to cases of accidental homonymy: polysemy can be described in terms of recurrent patterns of meaning extension. In this section, we proceed to trace out this connection in more detail.

A major point of this discussion is that with respect to the lexical semantic relations between full and light verbs, light verbs again differ from cases of grammaticalization. This time, however, the difference is fairly subtle. It would seem that grammaticalization must also start with a stage at which two variants of a verb emerge, and initially these will still be quite similar as well. What, then, is the difference to polysemy? To take an example, we can imagine that the use of a verb of movement as a future marker starts with uses of the verb in which the notion of spatial movement is no longer important, or cancellable, or has become metaphorical. Consider expressions like:

(52) The conference is coming to an end.
This use of *come* is neither spatial, nor does it constitute a tense marker. But we could imagine a state of affairs in which (52) develops the meaning of *The conference will end.*

There is no doubt that the semantic difference between light and full verb can be described as weakening or loss of content, or generalization – or “bleaching”, in some sense. This is much the same description that would be given for the difference between literal *come*, the use in (52) above, and a future marker. So, for describing the differences in content, there appears to be no reasonable alternative to invoking some notion of bleaching; and this, then, is the reason that a grammaticalization account for light verbs has been popular. Nevertheless, semantic bleaching, as the term is used in the literature on grammaticalization, is essentially a historical process. This historical implication, however, is not part of the notion of generalization if it is subsumed under polysemy. In this latter case, the different readings of a word can be grouped in a network, with some central and other more peripheral variants. A weaker, or generalized reading may be seen as dependent on the richer, and more central, reading, in the same way as metaphor is dependent on the concrete reading. But in both these cases, such a dependency relation is a synchronic relation between two closely related senses; it is not something that by definition could only emerge over time. This point is very clear for metaphor; it is less easily visible for generalization mainly because generalization is a somewhat less productive process. Here, then, is one part of the distinction between grammaticalization and polysemy that we want to stress: polysemy need not be thought of as something dynamic; even if we often use dynamic descriptions to render a synchronic dependency relation between two senses of a word.

Another difference is the following: in the long run, grammaticalization typically leads to lexical splits that dissimilate one form from the other. The most typical example of such a splitting process is the prosodic reduction of grammaticalized elements as opposed to their fully lexical sources. In polysemy, on the other hand, there are two (or more) readings for what is formally the same lexical item. Typically, such different readings are connected by definitional interrelationships or by similarity relations, and there are various indications for the fact that such polysemic complexes as a whole are an object of lexical knowledge (even though the connections between the readings are not fully productive). From this background, the history of different uses of the verb for ‘go’ in Indo-Aryan is very revealing. As shown in section 4.1, grammaticalization turned the verb ‘go’ into a
future tense form. At the same time, and in parallel to this development, however, there were light-verb uses of the same original verb for ‘go’. A modern example is repeated in (54), a Sanskrit one in (53).

(53) tato mak∑ikoɔi̯ya gatå
    then the fly-fly.Gd
    go.PastParticiple
    ‘then the fly flew away’

(54) naadyaa gir ga-yii
    Nadya.F.Sg.Nom fall go-Perf.F.Sg.
    ‘Nadya fell (down)’

Thus, the grammaticalization process that gave rise to the Urdu future split off one use of the verb ‘go’ that now consists only of the consonant -g-. This type of morphophonological reduction cannot be observed for the light verb ‘go’: it is still form-identical with the lexical verb ‘go’. Effects of a comparable lexical split between full and light verb cannot be documented from Sanskrit down to Urdu/Hindi over two millennia; i.e., there are no cases in which one use underwent formal changes that did not affect the other, or cases in which one variant was lost from the language but the other retained. A form-identical light-verb use always existed in parallel with the main verb, and we have seen that in a number of its grammatical properties, the light-verb still patterns with a lexical verb. The differences among the various lines of development of ‘go’ provide solid evidence that light-verbs are lexical-semantic variants of the full verb, not grammaticalized elements.

What is now required is an account of what the lexical meaning of light verbs consists of. In doing this, we would like to focus on a single example, namely the light verb dee ‘give’ in Urdu. It is a light verb with a particularly broad spectrum of applications. The contribution that it makes to the whole predication is not only elusive, but also highly variable. We would like to illustrate this point by discussing three meaning features that can be detected as a concomitant to the addition of dee ‘give’: agentivity, forcefulness, and directedness to a recipient. A more detailed discussion of the semantics of this light verb can be found in Butt and Geuder (2001), from which the rest of this section is adapted.
First, agentivity is a feature that is present with all occurrences of *dee* to some extent (however, in some examples it is weakened to mere responsibility and/or inadvertent causation). In order to ascertain that this is really a feature that resides in the light verb, compare the use of *dee* with a non-agentive light verb such as par5 ‘fall’. With this light verb, the agentive feature is suspended:

(55) naadyaa=nee gaanaa gaa dii-yaa
    Nadya.F.Sg=Erg song.M.Sg.Nom
    sing give-Perf.M.Sg
    ‘Nadya sang a song.’

(56) naadyaa gaa par5-ii
    Nadya.F.Sg.Nom sing fall-Perf.F.Sg
    ‘Nadya burst into singing (fell to singing).’

Next, consider the feature that we try to render as the “forcefulness” of the action. As already observed in example (49) above, it is *dee* ‘give’ which contributes such a feature, since it is absent if the light verb *lee* ‘take’ is inserted in the same context. The examples are repeated here for convenience:

(57) us=nee dus&man=koo
    paanii=me$e$s
    Pron.3.Sg=Erg en-
    emy.M.Sg=Acc water=in
    d5ubaa lii-yaa
    depress take-Perf.M.Sg
    ‘He/She drowned the enemy in the water.’
Light verbs in Urdu and grammaticalization

(58)  
\[ \text{us}=\text{nee} \quad \text{dus} \& \text{man}=\text{koo} \]
\[ \text{paani}=\text{me}$\text{e}$\text{s}\$ \quad \text{Pron.3.Sg}=\text{Erg} \quad \text{en-} \]
\[ \text{emy.M.Sg}=\text{Acc} \quad \text{water}=\text{in} \]
\[ d5u\text{baa} \text{ dii-yaa} \]
\[ \text{depress} \quad \text{give-Perf.M.Sg} \]

‘He/She drowned the enemy in the water (forcefully).’

But even though this proves that the forcefulness feature resides in the light verb ‘give’ in (58), there are other uses in which it is clearly absent:

(59)  
\[ b^\text{uul}=\text{se} \quad \text{muf}^\text{b} \text{e} \quad \text{apnaa} \]
\[ \text{forget}=\text{Inst} \quad \text{Pron1.Sg.Dat} \]
\[ \text{self.M.Sg.Nom} \]
\[ \text{sahii} \quad \text{naam} \quad \text{bataa} \quad \text{dii-yaa} \]
\[ \text{true} \quad \text{name.M.Sg.Nom} \quad \text{tell} \]
\[ \text{give-Perf.M.Sg} \]

‘He inadvertently told me his real name.’

(Hook 1974: 273)

(60)  
\[ \text{kisii}=\text{nee} \quad \text{bat5uaa} \quad k^\text{oo} \]
\[ \text{someone}=\text{Erg} \quad \text{wallet.M.Sg.Nom} \]
\[ \text{lose} \]
\[ \text{dii-yaa} \]
\[ \text{give-Perf.M.Sg} \]

‘Someone lost a/the wallet.’

(Hook 1974: 310)

Third, there is a feature of directedness of the action, and with this we can again observe the same phenomenon: if there is a recipient for the action, this is preferably marked by the light verb \textit{dee} ‘give’. Still, there are uses of \textit{dee} without the slightest trace of a recipient, as in (62).

(61)  
\[ \text{us}=\text{nee} \quad \text{prezid5ent5}=\text{koo} \]
\[ \text{Pron.3.Sg}=\text{Erg} \quad \text{president}=\text{Dat} \]
\[ x\text{at} \quad lik^k \quad \text{dii-yaa} \]
This variability in the semantic effects of light verbs is what makes the semantic analysis of their lexical content a particularly hard task. The data seem to require an analysis in terms of default reasoning: obviously, the light verb is able in principle to convey all of the meaning features like agentivity, forcefulness, and directedness—however, according to the context in which the light verb is used, a selection is made from this set of semantic features. If a feature is compatible with the meaning of the main verb, the light verb serves to add or enhance this feature in the construction, if a feature is unsuitable, it can be dropped (except for the features agentivity / “responsibility” and completion, cf. the remarks below). If this is on the right track, then we can see an interesting consequence for the issue of polysemy: if there is a dynamic interpretation process for the light verb that has to start out from an underlying representation in which all these meaning features are still united, then this underlying set of semantic features bears more similarity to the literal meaning of the verb ‘give’ than any single one of the surfacing interpretations. This removes some of the semantic distance between full and light verb ‘give’ and thus supports the analysis in terms of polysemy.

By putting together the features that we have collected from single instances so far, we arrive at the following picture of the meaning that underlies the light-verb uses: it is a complete agentive action (i.e., an accomplishment), involving the emission of something and directed at a recipient. We may assume that the notion of emission of an object that is present in literal ‘give’ is the source of the notion of “forcefulness” in the light-verb
use, for the action is now viewed as “emitted” and not just “done” (see again Butt and Geuder 2001 for more discussion, and Newman 1996 for discussion of the verb “give” in general, including the relatedness of emission and force transmission in its various uses). All in all, this underlying representation can be identified as a kind of skeletal meaning of the full verb ‘give’. The light verb meaning thus consists of certain fundamental semantic traits that can also be identified as the structural framework of the full lexical meaning.

While light and full verb meanings are found to be closely related in this respect, they turn out to have a very different status with regard to their compositional semantics. First of all, light verbs do not provide a complete, independent event description. They do not introduce an event of their own but rather, as we have said, they add features to the event description given by the main verb. Therefore, the logical form of a light-verb construction will at least require that they share the same event variable. There is more to be said, however. As a concrete example, consider again sentence (58), repeated below as (63a) and analyzed as (63b):

(63)  a. \textit{us=nee d5ubaa} \textit{dus&man=koo paanii=me$e$e}$\_{\textit{yaa}}$

\textit{Pron.3.Sg=Erg enemy.M.Sg=Acc water=in depress give-Perf.M.Sg}

‘He/She drowned the enemy in the water (forcefully).’

b. \exists \text{drown(e)} \& \text{Agent(e,he’)} \& \text{Theme(e, the enemy’)} \& \text{GIVE(e)}

It is of course unusual for this formula to contain two verbs; the identification of the events of drowning and of giving is doubtful given that events have fairly fine-grained identity conditions. It is only the light semantics of the light verb that appears to leave room for event identification here. A more obvious problem arises with respect to the thematic roles: if GIVE in this formula were a normal predicate of events, there would have to be a thematic relationship between GIVE, on the one hand, and the Agent and Theme arguments, on the other. In view of the agentivity requirement discussed above, it might seem reasonable to posit an Agent relation. However, it is entirely wrong to say that the Theme argument selected by ‘drown’ is also a Theme with respect to GIVE. Hence, event identification cannot work straightforwardly here.
The situation suggests that GIVE should be ascribed the semantics of a neo-Davidsonian modifier and not of a verb (in Geuder 2000 it is shown that the problem pointed out with (63b) typically appears in the logical form of many constructions with event-related adverbs, once the question of their lexical semantics is taken into account). In Butt and Geuder (2001) we therefore proposed an analysis in the style of neo-Davidsonian event modification. Thus, we hypothesize that light verbs in V-V constructions, by their lexical semantics, do not select any arguments at all. They combine with the main verb in the style of a modifier, thus potentially modulating semantic features of the predicate (note that an agentivity requirement also seems to be part of the semantics of many adverbs, e.g., ‘intentionally’ or ‘reluctantly’). Concerning the logical form of light-verb constructions, we can indeed use (63b), provided that GIVE is understood with the right kind of lexical semantics. It could roughly be explicated as follows (to distinguish it from the full verb ‘give’, we write GIVE-TYPE for the modifier variant):

(64)  \[ \exists e \text{ drown}(e) & \text{ Agent}(e,he) & \text{ Theme}(e,\text{the enemy}) & \text{ GIVE-TYPE}(e) \]

\[
\text{with } \text{ GIVE-TYPE}(e) = \\
\begin{align*}
\text{ e involves the force emission/transmission pattern } & \text{AGENT}(e) - \\
\text{ THEME}(e), \\
\text{ e is of the accomplishment type (completive),} \\
\text{ e is agentive/purposeful, etc} 
\end{align*}
\]

Incidentally, the fact that the meaning features contributed by the light verb are abstract and schematic makes light verbs appear somewhat similar to verbal classifiers. An analysis along these lines would not challenge our view on their compositional semantics, though.

From all this, it follows that there are two properties that make a verb “light”: first, in terms of semantic features, its meaning is very weak and general. Second, it is construed in a particular way, i.e., it serves as a modifier and cannot be used as a self-standing event description. Admittedly, the neo-Davidsonian notation alone is not ultimately satisfactory, however, it does express the intersective nature of the construction. The real semantic work must be done by the lexical content that is stipulated for this modifier. Singling out its features in the style of (64) is clearly only a start as the way in which the contribution of the modifier “GIVE-TYPE” is spelled out is, among other things, based on contextual inferences. Nevertheless, we believe that our approach is heading in the right direction in that it will ulti-
mately allow us to capture the semantic contribution of Urdu light verbs in general.

6. The connection between syntactic and semantic properties of light verbs

Light verbs, as we have seen, are characterized by a very special generalized, modifier-like semantics. The special semantic characteristics of light verbs go hand-in-hand with a special type of syntax. As we demonstrated in section 3, light verbs represent a unique syntactic class in Urdu, and can definitely not be analyzed on a par with auxiliaries.

Butt (1995) shows very clearly that the Urdu light verbs discussed here really involve complex predicate formation: The complex predicate does not involve syntactic embedding and there is no evidence for it being a control construction. We therefore propose that co-predication is to be represented by assuming a verbal complex $V'$ under which two verbal heads contribute to the predication, the main verb and the light verb (here, we want to leave open the question of whether the subject is generated in the first maximal projection of the verb or outside in an extended verbal projection; therefore we have indiscriminately labelled the verbal nodes as projections of the category V):

(65)

We thus posit a syntactic structure that expresses the idea that syntactically and semantically light verbs function as co-predicators: they provide additional lexical (not functional) content which modifies the main verb meaning. The structure in (65) shows a complex predicate in which both heads
contribute to the argument structure and co-determine case assignment.\textsuperscript{16} The light verb is to be considered the head of the first verbal constituent “V-complex” and hence carries the inflectional morphology for aspect and agreement if there is no auxiliary. This is the way compounds generally behave with respect to inflection.

The effects on the assignment of subject case mentioned in section 3.2.1. are in line with our analysis since the assignment of different subject cases in Urdu can be shown to depend on the lexical semantics of the predicate (cf. Butt 1995). Likewise, we can now see the deeper reason for the light verbs’ not conforming to the “Heir-Apparent Principle” formulated by Harris and Campbell (1995) (cf. section 2.2.3 above):

\begin{equation}
\text{(66) The Heir-Apparent Principle (Harris and Campbell 1995: 193).}
\end{equation}

When the two clauses are made one by diachronic processes, the main verb governs the syntax of the reflex clause.

Light verbs serve to modulate an event description, and since they contribute lexical information, they can take part in the determination of case and argument structure properties. To repeat, we would not consider the principle as falsified by our findings about light verbs. Rather, the evidence underlines the fact that light verbs function differently from auxiliaries. We would propose to replace the formulation that mentions clause fusion in the above statement by something more specific that makes clear that the principle is tailored to auxiliary constructions that arise from embedding constructions, which are then simplified by reanalysis. Recall that we could not find clear evidence for a similar historic process of reanalysis that brought about the light-verb construction, for even as far back as Sanskrit and Pāli there were examples that already seemed to have an interpretation very much in the style of light verbs.

With an analysis like (65), light verbs contrast with auxiliaries, which we assume to be part of the system of functional projections (relating to I

\textsuperscript{16} Much more could be said about the syntax of complex predicates and much more has in fact been said. Within Lexical-Functional Grammar, detailed proposals for the syntactic analysis of complex predicates can be found in Alsina (1996), Butt (1995), Mohanan (1994). Within the Government-Binding tradition, ideas that are compatible with our general account can be found in, e.g., Rosen (1989), Neeleman (1994), or van Riemsdijk (1998). Discussing the various syntactic issues would take us too far afield so we confine ourselves to pointing out that they are described in detail elsewhere.
or T). Furthermore, the history of light verbs cannot be identified with the historical development of auxiliaries, for auxiliaries tend to originate from earlier embedding constructions and thus develop into functional elements that form a layer of heads in the sentential structure above the lexical VP (cf. the discussion in section 2). The apparent historical stability of light verb constructions can thus be attributed to the particular syntax and semantics identified for modern Urdu. While embedding constructions tend to be vulnerable for a reanalysis of the syntax (or, in other words, a re-bracketing), as proposed by Roberts (1993) and Roberts and Roussou (1999), for example, the co-predication configuration in (65) will not be subject to structural simplification.

We thus conclude that the synchronic and diachronic peculiarities of the light-verb construction in terms of syntax (complex predicate formation) and semantics (addition of generalized, schematic features to another event description) yield a coherent picture when looked at in the right way.

7. Conclusion

In this paper, we have argued against the view that light verbs in Indo-Aryan developed from full verbs by way of grammaticalization. The idea that the source of light verbs is a grammaticalization process had some initial plausibility because light verbs might seem to be semantically bleached variants of full verbs. However, we have adduced evidence that militates against this proposal. Overall, our findings show that light verbs do not fit in with the usual model of grammaticalization from full verbs to auxiliaries and further on to morphological markers for grammatical categories. The following points established this conclusion:

– Diachronically, light verbs are not subject to progressive material reduction in the same way as auxiliaries.
– Light verbs differ from auxiliaries in the synchronic grammar (of Urdu); in particular, they rather pattern with lexical verbs than with auxiliaries in a number of respects.
– Semantically, light verbs fulfill a special function as modifiers of an event description. They have weak and schematic content, but they clearly belong to the realm of lexical elements, not grammatical formatives.
We propose that it is the special semantic function of light verbs that provides an explanation for the fact that they do not undergo grammaticalization: Light verbs serve the purpose of structuring event concepts. To this end, they are reduced to an entirely schematic meaning. The singular property of light verbs is that this schematic information is applied in order to structure a lexical content (namely the full verb), not in order to express a grammatical meaning, as is the case with auxiliaries and other grammatical formatives.

In brief, we propose that the staying power of light verbs results from the fact that they are, in a way, trapped in the lexical domain by virtue of their modifier semantics and co-predicational syntax. Even if there were a gradual development by which light verbs are becoming semantically weaker over time, this would not mean that they are becoming more grammatical: rather, this development would just lead to a more coarse-grained structuring of event concepts.

References


Böthlingk, Otto  

Butt, Miriam  

Butt, Miriam and Wilhelm Geuder  

Butt, Miriam and Tracy Holloway King  

Butt, Miriam and Aditi Lahiri  

Chatterji, Suniti Kumar  

Comrie, Bernard  

Dasgupta, Probal  

Delbrück, Berthold  
1888 *Altindische Syntax*. Halle: Verlag der Buchhandlung des Waisenhauses.

Fitzpatrick-Cole, Jennifer  
Geiger, Wilhelm

Geuder, Wilhelm
http://w210.ub.uni-tuebingen.de/dbt/volltexte/2002/546

Gopalakrishnan, Devi and Anvita Abbi

Harris, Alice and Lyle Campbell

Hendriksen, Hans

Hook, Peter Edwin


Hopper, Paul J. and Elizabeth Closs Traugott

Indurkhya, Bipin

Kellogg, S.H.

Kidwai, Aeysha
1997 Scrambling and Binding in Hindi-Urdu. Ph.D. diss., Jawaharlal Nehru University, New Delhi.
Light verbs in Urdu and grammaticalization

Kielhorn, Franz

Klein, Wolfgang

Lahiri, Aditi

Lahiri, Aditi and Jennifer Fitzpatrick-Cole

Lightfoot, David

McGregor, Ronald S.

Meissner, H. Konrad

Moens, Mark and Mark Steedman

Mohanan, Tara

Newman, John

Neeleman, Ad

Pinón, Christopher

Ramat, Paolo
Ramchand, Gillian

van Riemsdijk, Henk

Roberts, Ian

Roberts, Ian and Anna Roussou

Singh, Mona

Sweetser, Eve

Tikkanen, Bertil

Traugott, Elisabeth Closs

Trenckner, Vilhelm

Vale, Ramchandra

Whitney, William D.

Zbatitel, Dušan
Bilingualism and linguistic interference in the Slavic-Romance contact area of Molise (Southern Italy)

Walter Breu

1. Introduction

This article has grown out of my work on a dictionary of the linguistic interference of the Slavic minority language in the southern Italian region of Molise. The aim of this dictionary is to describe not only the actual state of the vocabulary but also the origin of the single lexicon entries, as well as phonetic, semantic and grammatical changes in a situation of total language contact (see Breu and Piccoli 2000). The main emphasis of the present article will be on language change in the lexical-semantic and grammatical-functional areas. In doing so, I will show that my data support the hypothesis that bilingual speakers do not strictly separate the vocabulary and grammar of two or more languages. Instead, they combine them with each other in the most economical manner possible. Special attention will be paid to the questions, to which extent language change through contact leads to the adaptation of different language systems, which pathways of development are chosen in this process, but also which areas resist adaptation and possibly may even show signs of independent developments distinct from the contact language.

2. The language situation in the Molisian area

Let me start with a few facts about the historical context and today’s linguistic situation: Molisian Slavic (MSL) is a minority language that has gained its own specific character through hundreds of years of Romance influence.¹ Judging from dialectal peculiarities (see Ivić 1958: 262–268),

¹ For a short overview of the linguistic and extra-linguistic situation of Molisian Slavic, see Breu (1997). Résetar (1911) is considered the classical standard work
the ancestors of the contemporary Molisian Slavs emigrated around 1500 from the Hercegovinian Neretva Valley. The reason for their exile was the Turkish expansion on the Balkan peninsula. They settled primarily in areas of Italy that were relatively vacant due to earthquakes and epidemics. Originally, they populated a considerably larger number of villages. Although the Croatian basis of the language is still very well recognizable, contact language interference is evident on all levels.

At present MSL is spoken with significant dialectal differences in three neighboring villages: Acquaviva Collecroce, San Felice del Molise, and Montemitro. The total number of inhabitants of these villages is less than 2500 today, and only a fraction of them still master the Molisian Slavic language. Especially in the second largest community, San Felice, it has completely disappeared from public life as a consequence of the heavy immigration of Italians, and it is nowadays spoken in the homes of only a few families. In both other villages the active usage of the language lies approximately at 60%, passive knowledge at 80%, although these figures vary greatly by generation. The following description is based on the dialect of the larger community of Acquaviva Collecroce.

At the beginning stages of Italian influence, the Slavic language of Molise (at that time still a Croatian dialect) was in a pure adstrate situation; see the left column of diagram (1). This means that elements and structures from the Romance contact language L2 were borrowed by the existing Slavic language system L1. As a result of this process, L1' came into existence:

<table>
<thead>
<tr>
<th>(1) Types of contact:</th>
<th>Adstrate / Superstrate = borrowing</th>
<th>Substrate = language acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 ⇐ L2 ⊃ L1' (MSL)</td>
<td>(primary language is changed)</td>
<td>L1 ⇒ L2 ⊃ L2'</td>
</tr>
<tr>
<td></td>
<td>(secondary language is changed)</td>
<td></td>
</tr>
</tbody>
</table>

---

on this minority language; see also the Italian translation from 1997 with updated comments. Reichenkron (1934: 338f.) deals with the Italian influences from the Romance perspective, yet his data are derived exclusively from Res&etar (1911).

2 Statistics from the years 1911, 1951 and 1991: Acquaviva 2243, 2250, 897, San Felice 1681, 1727, 881, Montemitro 1017, 906, 544.

3 For more on the question of the direction of language contact, see Breu (1994: 45–48) and Thomason & Kaufmann (1988: 37–45).
The special case of “superstrate” should generally be subsumed under adstrate because from a linguistic point of view it is the same type of language-contact alignment. We have to distinguish from them, however, the so-called substrate situation in the right column of diagram (1) in which speakers of one language L₁ give up their original language and change over to the contact language L₂. This corresponds with the situation of learning a foreign language, where interference from the source language – in the end – erroneously leads to L₂ instead of L₂.

Substrate situations with MSL as L₂', which generally result in more profound changes than adstrate situations, first came about in the MSL villages when locals and other Italians who married Slavic immigrants (presumably mostly Italian women) occasionally switched to Slavic. This phenomenon must have been restricted to individual cases, as a mass conversion to Slavic would be out of the question in view of the Italian majority and their political dominance in Molise. It can hence be assumed that substrate changes in Molisian Slavic were understood as errors by individuals, which consequently had only a minor direct influence on the changes in the language system as a whole. We can consequently presume, as already stated, that MSL is a L₁' language. The contemporary Romance elements in MSL have entered the language through borrowing, as is typical for the adstrate situation, and not by the principle of incomplete language acquisition. In this respect, the type (direction) of language contact in Molise is just the opposite of that in the Balkan Slavic languages, e.g. Bulgarian and Macedonian, another case of Slavic-Romance language contact, but with Balkan Romance as a substrate language.

As Weinreich (1964: 71) has already ascertained, the ultimate locus of language contact is the bilingual (multilingual) speaker. The contact languages influence each other reciprocally in his mind (and brain) and change as a result of the mutual interference to L₁' or L₂'. The type of language change that eventually occurs will also depend on the corrective of the language community, including the monolingual speakers. I cannot go into detail on this and other sociolinguistic questions. Instead, I will directly describe the empirically ascertained linguistic changes as the mutual influence of language systems.

In view of the given circumstances of MSL, the abbreviation L₁ in the following always refers to – as shown under (2) – the Slavic primary source, or replica language before the onset of a certain change, L₁ accordingly after the onset of this change. L₂ stands for the Romance contact or adstrate language, which functions as a secondary language and thus the
target language in the language acquisition or the source and model language for borrowings and restructuring in $L_1$. The situation is complicated by the fact that the Romance influence has a double origin, on one hand in the Molisian-Italian dialect, on the other in the standard Italian language:

(2) \[ \begin{align*}
L_1 &= \text{Slavic (Acquaviva): primary, source, or replica language before the change} \\
L_1' &= \text{Slavic: after the change} \\
L_2 &= \text{Romance (Molisian dialect, standard Italian): secondary, target, source, model language} \\
L_2' &= \text{incomplete } L_2 \text{ of the Molisian Slavs}
\end{align*} \]

In diachronic terms, the majority of the bilingual or trilingual Slavic population has experienced three periods of development: Until the middle of the 19th century, there was only the influence from the Molisian dialect, but after the unification of Italy both sources of influence began to compete with each other. Starting with the first few decades of the 20th century though, practically all new influences have come from standard Italian, because the Molisian-Italian dialect in Acquaviva today is known by relatively few people. If the differences between dialect and standard language are relevant in the following, I will make mention of them. For the sake of simplicity, I will otherwise use standard Italian forms when describing the contact.

3. Borrowed meanings (structural changes)

I will begin the discussion of the adstrate alterations with cases of meaning change in the vocabulary. Assuming that the vocabulary of bilinguals is structured economically, it might be expected that corresponding words in both languages cover about the same range of meanings. This would mean that in each case only one (not necessarily the same) form of expression has to be stored for one and the same complex of meanings in both languages and that corresponding word forms also have the same content potential. There are different ways that such parallelisms can arise in the semantic structure of two languages in contact when they differ originally in this respect, and these are dependent on the type of semantic linkage of the words and meanings in question. In the following, a few exemplary cases should suffice to elucidate these various processes. As already stated, I will focus
on structural alterations, i.e., the emergence and disappearance of polysems (homonyms), while sporadic changes such as isolated meaning deteriorations, etc., can only be discussed in passing.

3.1. Meaning expansion (polysemization)

If a given word in the Romance $L_2$, originally had more meanings than the corresponding lexeme in the Slavic primary language $L_1$, a parallelism in the semantic structure could be brought about by extending the meaning of the word in the replica language accordingly, making this polysemous, too. An example for this is (3):  

(3) older state: \( L_2 \) *prima* ‘first’ \( L_1 \) *prvo* ‘earlier’ 
   newer state: \( L_2 \) *prima* ‘first’ \( L_1 \) ‘earlier’  

In Italian *prima* means ‘first’ and ‘earlier’ among other things. In MSL there were conventionally two words for these two meanings, which still are carefully distinguished by several conservative speakers, on the one hand *prvo* ‘first’, on the other hand *prije* ‘earlier’. Today however, *prvo* covers both meanings for most speakers, like the Italian word *prima*. This shows a parallelization of the semantic structures of both languages. As can be seen in diagram (3), complications arise in the fact that the monosemic *prije* continues to exist as a variant in the meaning ‘earlier’ even among more progressive speakers. This, of course, contradicts the hypothesis that

---

4 In this paper I use a very wide concept of polysemy, thus not differentiating between, e.g., the polysemy of a word form due to a continuum of related meanings and the homonymy of clearly separate and historically independent meanings in one form. This does not mean that the degree of relatedness of the meanings of a given word does not have any influence on the effect of structural language contact. But finding the borderline between the two concepts (if there is any) lies beyond the scope of the present work. See Hopper & Traugott (1993: 69–72) for a discussion of this problem.

5 MSL examples are given in a broad transcription without diacritics for quantity, stress, and pitch; murmured (voiceless) vowels are printed in brevier.
bilinguals aim for the most economical semantic vocabulary organisation possible. If one dared to make a prediction, there would be two possibilities of doing this. Either prije also suffers at least a temporary meaning extension, or it falls completely out of use.⁶

Relationship terms constitute a lexical field in which the original strong differentiation of Slavic is being increasingly dismantled. As demonstrated in diagram (4), in the MSL of older speakers, which has nevertheless already undergone reductions in comparison to Croatian, a distinction remains for brothers-in-law and sisters-in-law between those from the husband’s family and those from the wife’s family. Among younger speakers, however, s&urjak/s&urjakica are used to describe both, coinciding with the Italian cognato/cognata, while the terms divar und zava have been lost:

(4) older speakers:

![Diagram showing relationship terms](image)

If we regard meaning extensions that have to do with technical innovations, we find that parallel developments take place in the contact languages L₁ and L₂, which, of course, also is in accordance with the principle of identity of their semantic structure; e.g. nosit in (5a) takes on the meaning of ‘to

---

⁶ As a matter of fact we have already found some speakers using prije as a full synonym of prvo.
drive a car’ in analogy to the regional Italian extension of *portare* from its original meaning ‘carry’ to ‘to drive a car’. Similarly, *potegnit* in (5b), like *tirare*, nowadays not only carries the original meaning of ‘to pull’, but also ‘to shoot’:

(5) a. \( L_2 \) *partare* ‘to carry’ \( \rightarrow \) *nosit* \( L_1 \) ‘to drive a car’

b. \( L_2 \) *tirare* ‘to pull’ \( \rightarrow \) *potegnit* \( L_1 \) ‘to shoot’

Counterexamples for adaptations to less differentiated semantic structures of the contact language exist as well, however. The MSL word *dask* in diagram (6) remains restricted to ‘board’ and cannot be used additionally for ‘(dinner) table’, unlike *tavola* in Italian. This latter is expressed by the word *store*. Hence there are conservative or resistant areas that will have to be examined more closely in the future:

(6) \( L_2 \) *tavola* ‘board’

\( L_1 \) *store*‘table’

Adaptations in the semantic structure that look like simplifications can eventually lead to semantic complications in other areas, as we have already seen in diagram (3). As shown in diagram (7), a complex case of this is the surprising development of *kas&a*, which originally meant ‘mush’ (or ‘porridge’) in Croatian and other Slavic languages,\(^7\) to ‘globe’ and ‘world’ among other things:

(7) older state: ‘mush’ \( \rightarrow \) *kas&a* \( L_1 \)

\(^7\) A relic of this original expression is still available in the MSL compound *kas&kavunisk*, literally ‘Slavic mush’, which was adopted by Molisian dialects as an expression for a cake filling made of grape jam.
In early stages of this change, an internal meaning degradation in \( L_1 \) probably took place, which could perhaps be traced back to cooking-habit differences between Slavs and local Italians. The cooking tradition of the Slavic minority, which has died out for the most part, was probably looked down upon by the Italians and therefore the term \( k\text{as}&a \) was identified in the replica language with ‘wet dirt, mud’, Ital. ‘fango’. Since the hyperonym \( \text{terra} \) ‘earth’ could be used regionally for \( \text{fango} \), \( k\text{as}&a \) was also treated as an equivalent of \( \text{terra} \). For reasons of semantic structure adaptation, this pertained not only to its special meaning ‘wet dirt’, but also to the whole extent of the meaning of \( \text{terra} \), therefore to ‘ground’ and to ‘earth’ or ‘world’ as well. On the other hand, Slavic \( \text{zemlja} \) had existed previously as the general word for ‘earth’, and was now used as a synonym for \( k\text{as}&a \) in the meanings ‘ground’ and ‘earth’. It is however still distinct from \( k\text{as}&a \) in not meaning ‘wet dirt’. Surprisingly, speakers show no tendency to resolve this complicated situation: A solution would only be possible by driving out the word \( \text{zemlja} \), for which there is no evidence at the moment.

3.2. Narrowing of meanings (monosemization)

In the examples above, the adaptation of the semantic structure took place by means of polysemization, due to the larger diversity of lexemes in the original Slavic \( L_1 \), see (3), (4). However, we also find the reverse situation in which a word from the Slavic \( L_1 \) exhibited more meanings than its Italian

---

8 One might compare in a similar matter the development of Slavic \( juch\a \) ‘soup’ to \emph{Jauche} ‘wet manure, sewage’ in German and analogously in the Sorbian minority languages in eastern Germany as well. The diminutive \( jus&ka \) still describes any type of ‘sauce’ in modern Sorbian. In Molise \emph{juha} has also been downgraded to the (once already used) ‘noodle water’.
counterpart(s). These meanings were expressed by several monosemic lexemes in the model language L₂, making it more differentiated in this respect than MSL. An adaptation of the semantic structure could have occurred principally by ways of a narrowing of meaning in L₁, thus reducing it to only one partial meaning; whereas the second partial meaning could have been expressed by a new word either derived by means of internal word formation or borrowed from the model language.

Contrary to these expectations, it is obviously very difficult for bilinguals to carry out the corresponding monosemizations. In some cases a solution is found in a disambiguation through specializing composition, such as in diagram (8a) with neb’ originally having the meanings ‘sky’ and ‘palate’ (vs. Ital. cielo : palato), where the meaning ‘palate’ is (optionally) specialized to neb’ do usti ‘sky of the mouth’. Yet it remains different from Italian in the idea that the palate and the sky are basically the same thing.

(8) a. younger state:

\[
\begin{array}{c}
\text{L}_2 \\
\text{cielo} & \text{‘sky’} \\
\text{palato} & \text{‘palate’} \\
\text{neb’} & \text{L}_1' \\
\end{array}
\]

Even more stably, the MSL word ruka (8b), meaning both ‘hand’ and ‘arm’, has remained polysemic up to this date despite the distinction in Italian between mano and braccio.⁹

b. \[
\begin{array}{c}
\text{L}_2 \\
\text{mano} & \text{‘hand’} \\
\text{braccio} & \text{‘arm’} \\
\text{ruka} & \text{L}_1 \\
\end{array}
\]

Even in cases of polysemy in which L₁ exhibits in other areas a conceptual division similar to that of the adstrate language, no disambiguation occurred. This case is exemplified by the polysemic imperfective verb grem ‘to go, to come’ in (9), which has remained polysemic, although its perfective equivalents pokj ‘to go’ and dokj ‘to come’ have always been monosemic and therefore correspond to the relationship between andare vs. venire in Italian.¹⁰

---

⁹ When the polysemy (homonymy) encompasses different parts of speech, disambiguation does not come about either; e.g. brat₁, ‘brother’, brat₂, ‘to harvest’. This is however less problematic, because ambiguities in the sentence context are out of the question here.

¹⁰ The insensibility towards a semantic structure adaptation goes so far that the (also imperfective) monoseme synonym jijam of grem in the meaning ‘to go’ is
The description of suspected changes is often complicated by the fact that the Romance contact languages themselves differ quite strongly in their semantic structures and that MSL – in the course of five centuries – has therefore been exposed to influences coming from both Italian and Molisian, which we always have to take into consideration. Thus, one could for example consider san ‘sleep, dream, temple (on one’s head)’ to be another case of non-adaptation, given the Italian monosemic lexemes sonno, sogno, tempia. However, one will immediately notice that ‘sleep’ and ‘dream’ in standard Croatian are indeed mononomously expressed by san, but not ‘temple’ which is denoted by sljepoć&nica. The puzzle can be resolved by pointing out that MSL, far from showing an idiosyncratic development, simply mirrored the word sùonn’ in the local Molisian dialect which has all three listed meanings:¹¹

On the other hand, subsequent contact with the Italian standard language has not yet led to a differentiation. In view of the restrictions on disambiguation in L₁ established above, this was not to be expected either.

---

¹¹ Details on Italian dialects discussed in this paper are from Giammarco (1968–85).
In these examples of underdifferentiation with respect to the adstrate language we find, however, that corresponding errors are made in the acquisition of Italian as a second language, i.e., L2 differs from L1. In this case, this is the only possible procedure by which bilinguals can reach the target of an economical – i.e., as uniform as possible – structure of the lexicon, as the possibility of contact-influenced meaning differentiation in L1 seems to be blocked.

3.3. Borrowing forms and structure adaptation

The lexical adaptations in MSL certainly did not take place solely by means of semantic changes, but also by borrowing complete lexemes (in form and content) as well. Here we must differentiate between the replacement of already existing lexemes and borrowing to express new concepts. An example of the first type is the substitution of razumit ‘to understand’ in (11a) by the Romance kapit (Ital. capire). There is still variation with regard to this concrete lexeme within the language community as a whole and individual variation among the older generation. On the other hand, bijat ‘to send’ (from inviare) in (11b) has already completely replaced the previously existing (Common Slavic) poslati. It would be interesting to look beyond the histories of individual lexemes to see which groups or lexical fields have been driven out by borrowing substitutions.

12 If in the case of example (10) a genuine Slavic polysemy in an earlier period of time is actually discovered, one could speculate that the meaning ‘temple’ of sùonn’ was previously carried over from Slavic into the surrounding Italian dialect due to such errors, as the cited polysemy can only be found in the province of Campobasso and the directly neighboring areas, where Italian-Slavic contact existed in former times. The dialects farther to the North have the monosemic form tèmbi’.

13 In the case of the example san or sùonn’ conclusions are also relevant in an inner-Romance context, because the Molisian dialect is also in contact with the standard Italian language.

14 As for technical innovations, polysemic lexemes are generally borrowed with the entire extent of their meaning, possibly by way of a multi-step process, e.g., in the case of magin’, which means ‘engine’ and ‘car’ just like the Italian macchina or in the case of batarij” with both meanings of Italian batteria, i.e. ‘battery’ and ‘drum-set’. In principle, a meaning extension of an existing word is also always possible in the case of new concepts; see the already discussed case of nosit with
Contrary to the principle of semantic structure adaptation, the adoption of words can also lead to meaning specification; e.g. in (12) with \textit{bagnare} ‘to wet, to bathe’ for which a verb with the same meaning in \textit{L$_1$}, i.e., \textit{smoc&it}, already existed. Nevertheless, \textit{banja} was borrowed, but only in the restricted sense of ‘to bathe an animal’.

The reason for borrowing with the specialized meaning could have been that the custom of washing sheep before the clip was adopted from the Italian neighbors and that \textit{smoc&it} belonged to the less central vocabulary. Yet the fact remains that such asymmetries in the semantic structure put a strain on the double lexicon of bilinguals; this is at least partially offset by the fact that as a result of this type of borrowing only one word form must be stored in the double lexicon for the given specific meaning.

3.4. Summary of the adaptation of the semantic structure

These few examples are sufficient to demonstrate that there is a basic difference in dealing with inequalities in the semantic structure of lexemes between contact languages, according to whether the minority language \textit{L$_1$} is more differentiated or less differentiated in a given lexical area than the adstrate language \textit{L$_2$}. If the minority language uses different lexemes to denote distinct concepts and the dominant language doesn’t this fine-grainedness will be abolished or altered. Evidently it does not cause the meaning ‘to drive a car’. As it seems, speakers tend to opt for this solution when a well-integrated word pair is available in the double lexicon of the bilinguals.
bilingual any difficulties to extend the range of one word, following the pattern of the adstrate language, and to sacrifice a specialized expression that has become “superfluous”. If on the other hand the dominant adstrate language is more differentiated, the bilingual usually does not manage to carry over this differentiation into the minority language. In this case, an alignment of the semantic structure can only take place by transferring the characteristics of the mother tongue into the contact language as a secondary language. This corresponds with the substrate type of language contact and leads to errors in the contact language.

Ideally, the additional borrowing of word forms results in the replacement of an existing lexeme in $L_1$ by the corresponding lexeme from $L_2$, with only temporary variation. In this case as well as when a lexeme introduces an entirely new concept, e.g., in the case of technical innovations, the semantic structures of $L_1$ and $L_2$ remain unchanged. But if only a partial meaning of the source lexeme is expressed by the loanword in $L_1$, an asymmetry in the semantic structures of $L_1$ and $L_2$ arises. In the end, the alignment leads to a semantic reduction in so far as the more complex structure is dismantled, whether it be in $L_1$ or $L_2$. This is then the basis of the common semantic structure $L_1'/L_2'$ in the lexicon of bilinguals. Yet the given examples show that the semantic adaptation in $L_1$ does not necessarily lead to a reduction of word forms. A new partial synonym pair can arise, complicating the semantic structure in another area so that the changed source language $L_i'$ under certain circumstances can be more complex than the original $L_i$, without directly reflecting the situation in $L_2$.

At the current state of research on the MSL lexicon, no statement can be made about the percentages of semantic adaptations and the complications that arise in the process.

4. Adstrate influence in the morphosyntax

4.1. Grammatical categories of the noun

It is assumed that four grammatical categories determine the inflection of Slavic nouns: case, number, gender, and declension.\(^{15}\)

\(^{15}\) The language data on the Romance influence in the MSL morphosyntax in the following are more thoroughly described in Breu (1998).
The category of declension (category of the stem class), which exerts only classifying characteristics, is structurally the weakest member. This is confirmed by the observation that these conventional differentiations have completely disappeared in MSL except for minimal residues in individual cases. As a result, a differentiation of paradigms based exclusively on gender has developed. This can be considered a general tendency in the development of Slavic languages that has been strengthened by the total Molisian language contact situation. The adstrate influence is so farreaching that nouns of the former feminine i-declension have been distributed according to gender in the contact language, i.e., by an external factor, to the masculine and feminine declensions. Typical examples for this are:

$$(13) \qquad \begin{align*}
\text{a.} & \quad \text{pec} \& \text{‘stove’ i-fem.} \rightarrow \text{mask.} \quad \text{// Ital. forno mask.} \\
& \quad \text{kost} \quad \text{‘bone’ i-fem.} \rightarrow \text{mask.} \quad \text{// Ital. osso mask.}
\end{align*}$$

$$
\begin{align*}
\text{b.} & \quad \text{stvar} \quad \text{‘thing’ i-fem.} \rightarrow \text{fem. (stvar$^\circ$)} \quad \text{// Ital. cosa fem.} \\
& \quad \text{ric} \& \text{‘word’ i-fem.} \rightarrow \text{fem. (ric$^\circ$)} \quad \text{// Ital. parola fem.}
\end{align*}
$$

In accordance with their conversion to the only remaining feminine declension type, which is based on the historical a-declension, the examples in (13b) take the corresponding new nominative forms stvar$^\circ$ or ric$^\circ$ in brackets.

Expressing the same meaning by two words of different gender in the two languages poses a special burden on the lexicon of bilinguals. Despite this, gender adaptations occur rarely in the morphologically regular vocabulary of MSL, which shows that the gender of a given noun is stored monolingually in a very strict manner. But, as the given examples show, a morphologically weak point, such as the loss of the feminine i-declension, can actually be the starting point for gender adaptation.

There was also a weak point in the category of gender itself, in that the Slavic neuter nouns did not have a corresponding gender in Italian. As a logical consequence the neuter grammeme was abolished and neuter nouns were distributed over the other genders. However, the adstrate influence is only effective for the reduction of the system as a whole, while the choice of the concrete gender in individual cases seems to be more a matter of chance and is generally not based on the gender in the contact language, contrary to the development of nouns of the former i-declension. In the transition of neuter nouns to feminine gender as in (14a) (a rare case altogether) the nominatives with vowels in the final position are preserved,
with an (unstressed) -o changing to -a according to the phonetic rules of the MSL of Acquaviva. The resulting form agrees with the nominative form of feminine nouns elsewhere. The change to masculine gender occurs in non-uniform ways. In most cases the nominative and accusative endings remain intact (in the form of -a), see (14b). Sometimes, though, they drop off; hence the resulting form matches the original masculine nouns, which do not have an ending; see (14c).


nebo ‘sky’ n., gen. neba → f. nebə, gen. Nebe // vs. Ital. cielo m.

b. mle&ko ‘milk’ n., gen. mle&ka → m. mblikə, gen. = nom. // Ital. latte m.

c. more ‘sea’ n., gen. mora → m. mor, gen. morə // Ital. mare m.

The gender of borrowed words remains for the most part identical with the gender in the source language. A notable exception are Italian feminine nouns with final stress: These are carried over into an existing masculine alternation class, unless they remain uninflected like dz&uvindu ‘youth’ in (15a) from the Ital. gioventù. The case is exemplified in (15b) c&ita ‘city’, gen. c&italə, from the Ital. città, which is inflected in analogy to the masculine paradigm of vo ‘ox’, gen. volə. In these cases, the change of gender seems to indicate that the formal expression of a loanword in the replica language is still dominant over the grammatical gender of its model in the source language. This can be viewed as a development within the replica language that is motivated internally in spite of the intimate contact situation.

(15) a. Ital. gioventù f. → dz&uvindu f. indeclinable ‘youth’

The category of case remains generally intact, in contrast to the situation in the contact languages. However, a loss of single grammemes has occurred; see Breu (1995). Thus, nominative, genitive, dative, accusative, and instrumental continue to be used, while the old vocative has been replaced by the nominative and the old locative by the accusative. It is likely that these replacements were directly instigated by the language contact. In the case of the former locative, the lack of a grammatical distinction between location and direction may have been decisive. Contrary to the original difference still existing in standard Croatian and in accordance with the situation in Italian, we now find in MSL the accusative bas” not only in the case (16a) expressing motion towards a location, but also in (16b), which denotes a location:

(16) a. grem dol u bas” // Ital. vado giù in cantina
   ‘I’m going down to the cellar.’

   b. jesa dol u bas” // Ital. sono giù in cantina
   ‘I’m down in the cellar.’

Another respect in which the case system has been reduced as a consequence of language contact is the emerging dominance of the genitive. Formally, we find that the genitive plural ending -i extends today beyond this case. Functionally, in L₁ prepositions often assign genitive case where they originally required the accusative. Even more uses of the genitive case arise by the widespread substitution of prepositions with accusative or locative government by adverbs with genitive government. An example is given in (17) with the construction zgora storce ‘on the table’. In these cases, MSL seems to imitate the widespread use of adverbs of location combined with the preposition di ‘of’ in Italian, especially in its dialects, e.g. dentro della casa for nella casa ‘in the house’.

(17) zgora [adv.] storce [gen.] ‘on the table’
    (instead of na ‘on’ with acc./loc.)

---

16 As a rule, borrowed prepositions from Italian also show genitive government, e.g., senza → sendz” ‘without’, which itself tends to be extended with di in colloquial Italian speech.
As a rule, constructions with Italian di are functionally equated with the genitive in MSL. But there is also a strong tendency to imitate them formally with the help of a prepositional phrase. The preposition do is applied here, which itself appears to have been blended with Italian di, as its original meaning has been changed from ‘until’ to ‘of’. Since it also takes the genitive, an analytical-synthetic double expression of the genitive arises:

(18) vrata do [prep.] hiz&e [gen.] ‘house door’ // Ital. porta di casa

As in Italian, such prepositional constructions play an important role in word-formation, where the genitive relationship is toned down, e.g., in mis& do zidi ‘bat’ (mouse of the walls) or the previously mentioned neb& do usti ‘palate’ (sky of the mouth). The complex construction of do with genitive government also mirrors Italian phrasal constructions with di, e.g., in jimas& nohta kan do mac&ke ‘you have fingernails like those of a cat’ (thus ‘long fingernails’), conforming with the Italian hai le unghie come quelle del gatto ‘you have fingernails like those of a cat’.

The newly constructed genitive with do with an additional inflectional ending shows – as demonstrated in (19) – all signs of a transitional form between the synthetic source form and the analytic construction of the contact languages. As a consequence of the double nature of the construction – both synthetic and analytic – it is in a certain way more complex than either of the two source forms:

(19)

\[
\begin{array}{c}
\text{L}_2 \\
\text{di} \\
\text{do + gen.} \\
\text{gen.} \\
\text{L}_1 \\
\text{L}_1'
\end{array}
\]

Regarding the category of number, the retention of the special paucal forms (surviving form of the dual) after the numbers 2, 3, 4 clearly stands in contrast to the contact language. For example:

(20) dva tic&en & ‘two birds’ (≠ nom. plural tic&enja ‘birds’)

Definiteness does not belong to the general Slavic grammatical categories of the noun. However, the Balkan Slavic languages Macedonian and Bulgarian have developed article systems based on demonstrative pronouns, and in other Slavic languages and dialects, for example in Sorbian, we find
indications that such a development has started. Since all these article systems arose under the influence of language contact, one might also expect an article system to arise in Molisian Slavic, which would reflect the Romance adstrate article system.

First of all, no definite article has been created in MSL. An important reason why it has not come into existence may be the fact that a corresponding formal identification model is missing in the contact languages. In Italian (unlike in German as the contact language of Sorbian), a link between articles and demonstratives is no longer recognizable. The situation is different for the indefinite article. The indefinite article in Italian can still be identified with the number ‘one’ (uno, una) and, following the constellation described in section 3.1, an analogous broadening of the range of meanings of the number word for ‘one’ in MSL was to be expected. Thus, it is no surprise that MSL uses na ‘one’ as an indefinite article in examples like (21):


‘A cow [fem.] is coming by.’

Even though no direct formal expression for the definite article exists, a privative determination category has developed indirectly. The reason for this is precisely the formation of an indefinite article. Since it is obligatory with indefinite countable objects as in (22a), its absence in (22b) symbolizes definiteness. This is why bilinguals constantly identify the zero morpheme in (22b) with the Ital. article *il* in anaphoric function:

(22)  a.  *Si kupij* na libar?  // Ital.  *Hai comprato un libro*
    ‘Did you buy a book?’

   b.  *Si kupij* Ø libar?  // Ital.  *Hai comprato il libro*
    ‘Did you buy the book?’

17 In comparison to Balkan Slavic, it must be stressed once again that there is an adstrate situation in Molise. Therefore the concept of a definite article would have to be newly introduced into the primary language $L_1$, whereas in the Balkan Slavic situation, articles had probably been in existence in the Romance substrate language, and simply had to be identified with the demonstratives of the acquired $L_2$. 
Not being able to differentiate between a definite article and no article certainly puts restraints on the category of definiteness. The proper functioning of this restricted article system must be examined more closely in the future. Whatever the result of such an analysis might be, the zero-article of MSL shows how a definite article could arise without being able to arise directly from polysemy in the model language. A secondary result of this development is a relatively low error rate in article usage in the Italian secondary language by comparison with speakers of other Slavic languages.

4.2. Grammatical categories of verbs

Looking at the verbal system, we find the grammatical categories of aspect, tense, and mood in the Slavic as well as in the Romance languages. Among the individual grammemes there are differences, though. In the category of aspect, Slavic languages originally distinguish between two formal oppositions, firstly the derivative opposition of perfectivity with the grammemes “perfective (pf.)” and “imperfective (ipf.)” expressed by means of aspectual pairs, and secondly the inflectional opposition of “imperfect: aorist/perfect”, which is limited to the preterite and exists in Italian as well. Both Slavic oppositions can still be found in MSL. In diagram (23) the most decisive part of the verb system in this respect is demonstrated, exemplified in the 1. Pers. Sg. of the word partit ‘to depart’, borrowed from Italian partire, which was first integrated as a perfective verb and then internally developed a corresponding imperfective verb parc&ivat:

(23) Ital. partire → partit [pf.] | parc&ivat [ipf.] ‘to depart’

<table>
<thead>
<tr>
<th>inflectional derivative</th>
<th>imperfect</th>
<th>aorist →</th>
<th>perfect</th>
</tr>
</thead>
<tbody>
<tr>
<td>imperfective</td>
<td>+ (parc&amp;ivah^a)</td>
<td>[+]</td>
<td>+ (sa parc&amp;iva)</td>
</tr>
<tr>
<td>perfective</td>
<td>+ (parc&amp;ah^a)</td>
<td>[+]</td>
<td>+ (sa partij^a)</td>
</tr>
</tbody>
</table>

It should be kept in mind, however, that in MSL – as shown by the arrows in (23) – a typical restriction of the inflectional opposition occurred, which resulted in the complete loss of the aorist and in the perfect taking over all
its functions, see Breu (1992: 104–108). Consequently, this opposition is now expressed only by the grammemes imperfect and perfect. Hence, the aspectual system of MSL as a whole provides a counterexample to a diachronic constant of Slavic which states that, in the case of reduction of the inflectional aspect category, the imperfect always disappears first, and only thereafter the aorist. The case is exemplified by conservative Croatian and Serbian dialects or by the process of replacement of the simple past (aorist and imperfect) by the old perfect in Russian.

The reason for this special development appears to lie within the Romance adstrate: Following a widespread development in the Romance languages, the local Molisian dialects of Italian also replaced the aorist (passato remoto) with the perfect, with the imperfect remaining intact. We lack historical evidence to decide whether the loss of the aorist in MSL should be classified as a parallel development in both contact languages after the immigration or instead as a secondary adaptation. In each case, an identification of the inflectional aspect category of Molisian Slavic with that of Romance must have occurred.

Regarding the category of tense, the strong position of the pluperfect in MSL is very evident. In most Slavic languages it has either completely disappeared or its usage is greatly restricted. A decisive factor is of course the identification with the corresponding grammeme in Italian.

The tendency to express time relationships explicitly by using separate tense forms, as observed in the use of the pluperfect, also pertains to the MSL future. The future tense is used even more consistently than in the surrounding Italian dialects and in the regional colloquial language, in which speakers prefer to use the present to express future events. The future appears in two sets of forms, functionally differentiated with respect to mood: We find a probability future like in (24) and a necessity future like in (25):

(24)  
\[c\&u \ po \ rabit\] ‘I will (probably) go to work.’

(25)  
\[mam \ po \ rabit\] ‘I will (definitely) go to work.’

The probability future is formally derived from the Southern Slavic future of the volition type (auxiliary WANT + main verb); compare Croatian \[radit\] c/u ‘I’ll work’. The necessity future is an innovation driven by the adstrate situation and is expressed, as in Southern Italian dialects with the auxiliary
verb *imat* ‘to have, must’. As a result, the MSL tense system is more complex than the original one in L₁ and equally encompasses that of L₂.

Changes in the category of mood can be exemplified by conditional sentences. The genuine Slavic analytical construction with the conjunctive marker *bi-* and the *l*-participle continues to exist in the subordinate conditional sentence (protasis) and the superordinate sentence (apodosis). Along with this exists the Slavic peculiarity that one cannot separate between a potential, feasible condition and an unreal condition, for which there are distinct means of expression in Italian. Thus, the Molisian Slavic sentence (26) is equivalent to the Italian sentences in both (26a) and (26b):

\[
\text{(26)} \quad \text{Si ja bi dosa, ovo ne bi surtila.}
\]

a. Ital. *Se venissi, questo non succederebbe.*
   ‘If I came, that would not happen.’

b. Ital. *Se fossi venuto, questo non sarebbe successo.*
   ‘If I had come, that would not have happened.’

Hence, the MSL *bi*-phrase corresponds to four different forms in standard Italian, namely the potential condition (*venissi*), the irreal condition (*fossi venuto*) and the potential or unreal consequence (*succederebbe* or *sarebbe successo*).

There is still another way of expressing unreal conditions in colloquial Italian, namely by the double-setting of the indicative imperfect. Under the adstrate influence this construction has been transferred to MSL, such that the meaning of (26b) can also be expressed by the MSL sentence (26c):

\[
\text{c. } \text{Si ja dojah}^a \text{ [imperf.], ovo ne surca}^a \text{ [imperf.].}
\]

Ital. *Se venivo, questo non succedeva.*

Due to the interaction of both forms of expression for unreal conditions and the polyfunctionality of the *bi*-construction, confusions can arise in MSL. As a consequence, the indicative imperfect that is employed in the irrealis construction in (26c) can also be applied in the potentialis. This means that speakers cannot express the difference between irrealis and potentialis in this type of construction either. As a further consequence, the indicative imperfect can also be used in the free conditional as in (26d). In Italian this

\[\text{18 For a more thorough description of the conditional in MSL, see Breu (1999).} \]
is just as impossible as using the indicative imperfect in the potentialis itself:

d. *To ne c&inja*.

corresponds with Ital. *non lo facevo* ‘I did not do that’, ‘I would not have done that’, but it is also associated with *non lo farei* ‘I would not do that’.

The existing variations in the system of MSL are erroneously carried over into Italian by bilinguals; thus there is again a parallelization of the structures of languages in contact with the debit of L₂.

The MSL developments in the area of the conditional show once again that an extension of functions in analogy to L₂ can occur unproblematically (see the extension of the imperfect to the irrealis), but the introduction of a new distinction between the potentialis and irrealis in L₁’ in analogy to L₂ is impossible, giving rise instead to an error in L₂’ acquisition. This is in accordance with the situation that we have observed in the field of the lexical-semantic underdifferentiations. The errors showing up in the contact language L₂ demonstrate in both cases that bilinguals, even in the case of underdifferentiation of L₁, attempt to construct a parallel semantic structure in both languages, i.e., parallel ways of connecting forms and meanings or functions. As this is not achievable by means of a change in L₁ (i.e., by borrowing forms or structures which is typical for the case of substrate) it takes place through a change in L₂ by means of incorrect learning, which is typical for the substrate.

5. Conclusion

In conclusion, the Molisian Slavic linguistic situation confirms the hypothesis that bilinguals tend to build a uniform semantic (and functional) structure for both contact languages. The respective developments always lead to a parallel underdifferentiation of existing lexemes or grammatical categories. As a result, we find either a corresponding analogous change in MSL or the improper use of the Italian contact and target language. Generally the absorption of additional meanings for a given lexeme or the borrowing of a whole lexeme does not pose any difficulties whatsoever. Nor do functional extensions of already existing grammemes lead to any difficulties. Even the creation of new grammemes in analogy to the model language occurs. However, the singling out of a partial meaning of a given lexeme or a partial function of a grammeme represents the absolute excep-
tion to the rule. It is thus very difficult for bilinguals to detach concepts combined in a common form.

It has been ascertained at several points that there is a tendency to form transitional structures between \( L_1 \) and \( L_2 \) in MSL, which are in some senses more complex than the initial structure of each single contact language. This tendency can also be found on other levels of the MSL language that could not be discussed here. As an example, in phonology an extremely complicated system has emerged with quantity oppositions, murmured (voiceless) vowels, at least four degrees of vowel height, and with the preservation of clear reflexes of the Croatian tonal system as well; see Breu and Piccoli (2000: 385–389). In all these cases, the complex structure allows bilinguals to treat the contact language \( L_2 \) as a subsystem of \( L_1' \). In summary, MSL as the product of an intimate contact situation with totally bilingual speakers is a demonstration of the integration of two languages in one system that bilinguals generally strive for.

References

Breu, Walter


Walter Breu


Lexical-grammatical variation and development: The use of conjunctions as discourse markers in everyday spoken German

Susanne Günthner

1. Introduction

The subject of this study lies at the interface of lexicon and grammar. Based on conversational data from colloquial German collected from 1983 to 1999, it aims at uncovering how “new” uses for “old” words have developed in spoken colloquial German. The analysis will focus on the growing tendency over the last 20 to 25 years to use the traditionally subordinating conjunctions weil ‘because’ and obwohl ‘although’ with main-clause syntax.

German, which has “verb-second” as its basic word order in independent sentences, requires final positioning of the finite verb in subordinate clauses. Thus, clauses introduced by subordinating conjunctions, such as weil ‘because’ and obwohl ‘although’ – according to German grammar – must display “verb-final” ordering:

(1)  
Er kommt nicht, weil er Grippe hat.  
‘He comes not, because he flu has.’ (word-by-word gloss)

Sie kommt, obwohl sie Grippe hat.  
‘She comes, although she flu has.’ (word-by-word gloss)

The weil- and obwohl-clauses show final positioning of the finite verb “hat” (has).

German thus provides a clear signal for the grammatical incorporation of one clause into another. However, over the last 20 to 25 years, in spoken colloquial German as well as in certain written genres that reproduce colloquial language (e.g., interviews, dialogue in advertising, e-mail-

interaction), speakers increasingly use these traditional subordinate conjunctions to introduce clauses with verb-second ordering; i.e., main-clause syntax:²

(2) \textit{Er kommt nicht, weil (-) er hat Grippe.}\n
‘He comes not, because (-) he has flu.’ (word-by-word gloss)

\textit{Sie kommt, obwohl (-) sie hat Grippe.}\n
‘She comes, although (-) she has flu.’ (word-by-word gloss)

These “new” uses cause the connectors \textit{weil} and \textit{obwohl} to lose grammatical features, which identify them as subordinators.³

On the basis of an analysis of conversational data I will address the following questions:

– Is there a trend to reanalyze the traditional subordinating conjunctions \textit{weil} and \textit{obwohl} and treat them as coordinators?
– Do the different uses of \textit{weil} and \textit{obwohl} (with verb-final and verb-second positioning) reveal different discourse functions?
– Are these “new” functions and meanings of \textit{weil} and \textit{obwohl} related to their traditional semantics?
– Is there a general tendency to give up subordinate clause order in German?

The following analysis is based on 67 everyday informal German conversations (30 to 180 minutes in length) among friends and family members (dinner conversations, chats over coffee, telephone interactions) and 9 radio-phone-in conversations. These data were collected from 1983 to 1999 in Baden-Wuerttemberg, Bavaria, Brandenburg, Hesse, and Thuringia. Methods of conversation analysis and discourse analysis are employed to study the lexical-grammatical features of spoken language.

³ For a discussion on the increase of \textit{weil} and \textit{obwohl} with verb-second ordering during the last 30 years cf. Schlobinski (1992); Keller (1993); Pasch (1997); Uhmann (1998); Günthner (1999).
2. Traditional uses of weil and obwohl

2.1. Weil with verb-final positioning

The standard, unmarked word order for weil-clauses in German is verb-final position; and thus “integrative word order” (König and van der Auwera 1988), marking the weil-clause as a subordinate one.

The following segment is taken from a family dinner conversation. Karl tells his parents about his trip to China:

(3) RIKSHA IN CHINA

7Karl: und da hatten wir uns dann (-) ja eigentlich vom kaiserpalast aus (-)
8 mit so ner rikscha da rüber fahrn lassen.
9Klara: ja.
10Karl: was auch (wieder) ne ziemlich BLÖDE erfahrung war, eh (-)
11 weil die uns (-) mit dem PREIS für die rikscha (-)
12 ziemlich übers ohr haun [wollten.]

7Karl: and then we (-) well to get a ride from the imperial palace (-)
8 we took a riksha.
9Klara: yeah.
10Karl: which (again) turned out to be a bad experience, eh (-)
11 because they (-) with the price for the riksha (-)
12 they wanted to do us out of a lot of money. (-)

The causal clause “weil die uns (-) mit dem PREIS für die rikscha (-) ziemlich übers ohr haun wollten.” (because they (-) with the price for the riksha (-) they wanted to do us out of a lot of money) provides the reason for the main-clause proposition “was auch (wieder) ne ziemlich BLÖDE erfahrung war,” (which (again) turned out to be a bad experience ). Thus, the two clauses (the main-clause and the subordinate weil-clause) are closely connected by “real-world causality” (Sweetser 1990). As the finite verb of the weil-clause is in final position (“weil die uns (-) mit dem PREIS für die rikscha (-) ziemlich übers ohr haun wollten.”), this clause is clearly marked as subordinate.
As subordinate clauses, *weil*-clauses can be in final position, following the main-clause (as in our transcript RIKSHA IN CHINA); they can precede the main-clause (“Weil die uns mit dem Preis für die Rikscha ziemlich übers Ohr hauen wollten, war das auch wieder ne ziemlich blöde Erfahrung”), or they can be integrated into the main-clause (“Das war auch wieder – weil die uns mit dem Preis für die Rikscha ziemlich übers Ohr hauen wollten, ne ziemlich böde Erfahrung”).

Furthermore, typical subordinate clause formulations are possible; e.g., in questions, the *weil*-clause is within the scope of the main-clause: “War das auch ne ziemlich blöde Erfahrung, weil die uns mit dem Preis für die Rikscha ziemlich übers Ohr hauen wollten?” (“did it turn out to be a bad experience because they wanted to do us out of a lot of money?”), and we can change the *weil*-clause into an adverbial phrase: “Wegen deren Versuch, uns mit dem Preis für die Rikscha übers Ohr zu hauen, war das auch ne ziemlich blöde Erfahrung.” (Because of their intention to do us out of a lot of money, it turned out to be a bad experience). Syntactically, *weil* as a subordinate conjunction fills the C-position of a verb-final clause. The *weil*-clause is a constituent of the preceding main-clause’s VP.

In general, *weil*-clauses with final positioning of the finite verb are used to mark a tight causal link often operating in the content domain (Sweetser 1990), indicating a causal relation between two propositions or events. The *weil*-clause is tied to the scope of the main-clause’s illocutionary force. Syntactic integration can be emphasized by means of prosodic integration; i.e., the *weil*-clause can be prosodically integrated into the main-clause. However, as the *weil*-clause in RIKSHA IN CHINA reveals, prosodic integration is not a prerequisite for subordinate word order in *weil*-clauses.

2.2. *Obwohl* with verb-final positioning

As with *weil*-clauses, the standard unmarked word order for *obwohl*-concessives is final positioning of the finite verb (i.e., syntactic integra-

---

5 Cf. also Wegener (1993).
Lexical-grammatical variation and development

The following transcript segment is taken from an informal conversation about people in East Germany who had internalized their former ways of behaving under the socialist regime. Thea gives the example of an East German woman who – even after German reunification – still stood to attention when she met a former Stasi official:

(4) ANPASSUNGSSTRATEGIEN (adaptation strategies)

34Thea: und als der dann wieder kam,
35 stand die STRAMM,
36 obwohl sie nicht mehr stramm stehn mußte. (-)
37 also so verhaltensmuster,
38 die ma drin warn,
39 kommen dann auch wieder;

34Thea: and when he came back,
35 she stood to attention,
36 although she didn’t have to stand to attention any longer. (-)
37 well such patterns of behavior,
38 which were once inculcated,
39 do come back;

As Quirk et al. (1985: 1098) state, “Concessive clauses indicate that the situation in the matrix clause is contrary to expectation in the light of what is said in the concessive clause”. Thea’s utterance “und als der dann wieder kam, stand die STRAMM,” (and when he came back, she stood to attention) is presented as contrary to expectation raised by the concessive clause “obwohl sie nicht mehr stramm stehn mußte.” (although she didn’t have to stand to attention any longer). The concessive construction thus expresses – as König and van der Auwera (1988: 107) point out – that there is a “relationship of ‘normal incompatibility’ or dissonance between the two component propositions”. This can be stated as presupposing a conditional: If p, then normally ¬ q. Accordingly, the discourse presupposition of the example ANPASSUNGSSTRATEGIEN is: if you don’t have to stand to attention, you do not stand to attention.

Logically, concessive clauses entail both their component clauses, i.e., both “p” (she didn’t have to stand to attention any longer) and “q” (she stood to attention) are treated as valid. As subordinate weil-clauses, obwohl-clauses with verb-final positioning can follow the main-clause (as in
ANPASSUNGSSTRATEGIEN), precede the main-clause (“obwohl sie nicht mehr stramm stehn mußte, stand sie stramm”) or they can be inserted into the main-clause (“sie stand – obwohl sie nicht mehr stramm stehen mußte – dennoch stramm”).

Furthermore, typical subordinate clause transformations are possible with subordinate obwohl-construction; e.g., in the case of a question, the obwohl-clause is within the scope of the main-clause: “Stand sie stramm, obwohl sie nicht mehr stramm stehen mußte?” (Did she to stand to attention, although she didn’t have to?), and we can transform the obwohl-clause into an adverbial phrase: “Trotz der Tatsache, daß sie nicht mehr stramm stehen mußte, stand sie stramm” (In spite of the fact that she didn’t have to stand to attention, she stood to attention).

The obwohl-clause is a constituent of the preceding main-clauses’ VP and fills the C-position of the verb-final clause.

3. “New” uses of weil and obwohl

Over the last 20 to 25 years there has been an increasing tendency to use weil and obwohl to introduce clauses which show all the characteristics of independent sentences.\(^8\) In these cases, the traditional subordinate conjunctions weil and obwohl no longer fill the C-position of a verb-final clause, but introduce a sentence which is separately assertable and carries its own illocutionary force.

\(^8\) Cf. Selting (1999) on the question of a possible historical continuity of weil with main clause syntax. She argues that – in spite of the fact that weil with verb-second position can be found in non-standard varieties of spoken German back to the beginning of the last century – it is not possible to show a continuity of weil with verb-second positioning from Middle High German to the 20th century. Cf. also Arndt (1956) and Eroms (1980) on the historical development of causal conjunctions in German. Up to the 17th century weil was used with subordinate as well as coordinate verb positioning. In the 17th century, weil with verb-final positioning finally was standardized (Maurer 1926).
3.1. *Weil* with verb-second positioning

Especially in colloquial spoken German, speakers increasingly use main-clause order in *weil*-utterances. Weil with second positioning of the finite verb is mainly employed in cases in which there is only a loose causal connection between the preceding utterance and the *weil*-utterance; i.e., (i) in cases in which the causal connection operates in the epistemic or speech-act domain; (ii) in cases in which there is no direct causal relationship between the *weil*-utterance and the preceding clause.

(i) **weil operating in the epistemic and speech-act domain**

The following example shows the use of *weil* with verb-second positioning. Here, *weil* operates in the “epistemic domain”. Arne talks about his neighbor (Klaus):

(5) **FRÜHSTÜCK** (breakfast)

| 12Arne | der hat sicher wieder getrunken. (–) |
| 13     | weil (–) sie läuft total deprimiert durch die Gegend. |

| 12Arne | he must have been drinking again. (–) |
| 13     | because (–) she is walking around looking totally depressed |

Arne’s *weil*-clause (line 13) shows second positioning of the finite verb *läuft*, and thus reveals main-clause order. The *weil*-clause does not introduce the reason for Klaus’ drinking; instead it provides the basis for Arne’s assumption that Klaus must have been drinking again. Thus, the causality of this epistemic *weil*-construction is that between the premise (*she is walking around looking totally depressed*) and the conclusion in the speaker’s

---


10 For causal connectors operating in the epistemic and speech-act domain cf. Sweester (1990). In these cases *weil* functions on the “internal plane of conjunctive relations” (Halliday and Hasan 1976: 241–243). Cf. also Eroms (1980) on different functions (“Äußerungs begründungen vs. Sachverhalts begründungen”) of causal conjunctions in “new high German”.
mind (*he must have been drinking again*). The modal adverb “sicher” (*definitely*; line 12) functions as a lexical indicator of this epistemic reading. In the case of verb-final positioning, the utterance would get a different reading altogether. The subordinate syntax “der hat sicher wieder gesoffen, weil sie total deprimiert durch die Gegend läuft” in combination with prosodic integration of the two clauses would suggest that her walking around looking totally depressed is the reason for his drinking again, and the *weil*-relation could be interpreted as operating in the content domain.

In the next segment, the *weil*-utterance provides the reason for the preceding speech-act. Again, main-clause order is used in the *weil*-clause, i.e., the finite verb *reichen* is in second position:

(6) **MOTTEN** *(moths)*  

1Anna: warum kauft ihr denn keine größeren müslipäckchen.(-)  
2 weil (-) DIE reichen doch nirgends hin.  
1Anna: why don’t you buy bigger packages of müsli.  
2 because (-) these don’t get you anywhere.

As with epistemic *weil*, the *weil*-clause in line 2 does not provide the reason behind a previous assertion, but rather it provides a causal explanation for the speech-act (the question) performed in the preceding clause.¹¹ The reading is something like “I’m asking why you don’t buy bigger packages of müsli, because the ones you have don’t get you anywhere”.

A characteristic feature that speech-act *weil*-clauses share with epistemic ones is that in both types the *weil*-construction is “separately assertable”, i.e., it occurs as a separate assertion;¹² both clauses are rhematic, and they both have their own illocutionary forces. The separate assertion is not only revealed on the syntactic level by means of main-clause order, but also on the prosodic level: Epistemic and speech-act *weil*-clauses are prosodically independent: They indicate non-integration with the preceding clause by their own intonation contours, pauses, and other prosodic contextualization cues. They are hence independent from the material they elaborate on. One often finds a little pause after the *weil*:

¹¹ Cf. Sweetser (1990) on causal relations in the speech-act domain.  
Lexical-grammatical variation and development

FRÜHSTÜCK: 13 weil (-) sie läuft total deprimiert durch die Gegend.

MOTTEN: 2 weil (-) DIE reichen doch nirgends hin.

The pause after weil, however, is not obligatory.13

Contrary to subordinate weil-clauses, those with verb-second position are restricted to final positioning, following the “main”-clause. Furthermore, typical subordinate transformations (e.g., questions, adverbial phrases, etc.) are no longer possible, or would cause a different reading. The question “Hat der wieder gesoffen, weil sie total deprimiert durch die Gegend läuft?” (Did he drink again, because she is walking around looking totally depressed?) would be interpreted as a causal relation operating in the content domain; i.e., the weil-clause provides the reason for his drinking. In weil-clauses with verb-second position, main-clause phenomena (Green 1976) are possible, e.g., left-dislocations: “Der hat sicher wieder gesoffen, weil seine Frau (-) die läuft total deprimiert durch die Gegend” (he must have been drinking again, because his wife (-) she walks around looking totally depressed). As the example MOTTEN illustrates, the speech-act realized by the weil-clause is independent of the illocutionary force of the preceding utterance: The weil-clause follows a question. Thus, in case of weil with main-clause syntax, the weil-clause is no longer a constituent of the preceding matrix clause, but an independent clause introduced by the connector weil.14

(ii) weil-clauses which are not directly related to the preceding clauses

There are weil-clauses displaying main-clause order, which can neither be attributed to the epistemic nor to the speech-act level. Weil no longer pro-

---

13 Cf. also Schlobinski (1992: 335) and Uhmann (1998: 104). In written texts which reconstruct spoken language, this pause is often reflected by means of dashes, commas, or colons: In a reader’s letter published in the weekly magazine SPIEGEL 1989 issue of July 24th, we find the following passage: “...weil, man will ja zeigen, was man hat” (“...because, one wants to show, what one has”).

14 In the cases at hand one can talk of weil as a coordinate conjunction introducing an independent clause, which carries its own assertion. Cf. parallels between the coordinate use of weil and traditional uses of the causal coordinate conjunction denn (Eroms 1980; Pasch 1997; Uhmann 1998).
vides a reason for the preceding proposition, conclusion, or speech-act; instead it has mainly discourse-organizational functions: It is used to introduce additional explanatory information, to initiate a topical change, and to function as a conversational “continuation marker” signaling that the present speaker does want to continue talking (Günthner 1993; Gohl and Günthner 1999).

The following text segment is taken from a radio phone-in counseling session. The caller talks about the problems she has had dealing with her mother’s death and with the fact that her aunt denies that the speaker’s mother (the aunt’s sister) is dead. The caller mentions that her relationship with her mother was not good. The counselor then assumes that the caller feels guilty about this bad relationship with her mother:

(8) BELASTUNG (burden)

Beraterin: [and this] you hold now against yourself.

(0.5)

Anruferin: no I don’t hold it against myself.

(0.5)

Beraterin: well then (. .) it would actually be possible,

(1.0)

Anruferin: =well but only I mean now this puts (well) this woman now burdens my conscience

because I mean I would have done things a lot differently,

but (-) well here (-) my mother had to die first-
in order for me to think about this;

you know what I mean.

In line 6 the caller changes the subtopic from talking about the problematic relationship with her mother and the question of whether she holds this against herself to the topic of her aunt’s behavior. Then she switches to her subjective assumption (line 9) and provides a new thematic aspect: “ich hätte sicherlich manches anders gemacht” (I mean I would have done things a lot differently). At the same time a change of perspective is taking place: the caller now focuses on her own behavior. This change of topic is initiated by weil (line 9) in combination with the epistemic phrase ‘I mean’ (line 9). It is apparent that this weil-clause cannot be analyzed as a constituent of the structures around it; i.e., the utterance “because I mean I would have done things a lot differently” is not offered as the reason for “this woman now burdens my conscience”, nor does it warrant the conclusion that this woman burdens the speaker’s conscience. The weil-clause is but very loosely connected with the preceding.

In the next transcript weil is used as a floor-holding device (Sacks, Schegloff, and Jefferson 1974). The speaker makes use of the projective potential of weil in order to keep the floor. Tom complains to his friend Eli about having had to wait for two and a half hours at the dentist:

(9) ZAHNARZT (dentist)

23Tom: zweieinhalf stun[den.]
24Geli: [des] isch doch die HÖH!
25Tom: hm. der müller hat nachher auch NIX
[(gsagt)];
26? [ hm ]
27Tom: s=intressiert doch DEN nicht.
28 weil (..) eh ja=also, (-)
29 des=des is denen doch storzegal.
23Tom: two and a half [hours.]
24Geli: [that’s] incredible!
25 Tom: hm. afterwards müller didn’t say [a word];
26 [hm]
27 Tom: HE is simply not interested.
28 because (,) eh yeah=well, (-)
29 they just don’t care at all.

Once again, weil does not introduce a causal explanation for the preceding statement (”es interessiert doch den nicht, weil des denen doch storzegal is”; HE is simply not interested, because they just don’t care at all). Instead, weil indicates that the speaker intends to continue talking.

In both cases (BELASTUNG and ZAHNARZT) the utterance introduced by weil shows verb-second position and has all the characteristics of a main-clause: It carries its own illocutionary force and it has its own prosodic contour. Typical subordinate transformations are no longer possible, instead main-clause phenomena may appear. In both transcript examples (BELASTUNG and ZAHNARZT) weil does not provide an explanation for a preceding proposition, conclusion, or speech-act; its causal meaning has faded. Thus, its function is to connect longer discourse units – for lack of a (clear) causal relation.

In the case of BELASTUNG and ZAHNARZT, it is difficult to still classify weil as a conjunction. It does not hold the connecting position between two symmetrical clauses, instead it is placed at the periphery of the following syntagma and is subject to semantic bleaching. It functions as a metacommunicative element, indicating the textual relationship between the preceding and the following text segment.

To summarize: weil-constructions with verb-second position share the following characteristics:

– they are restricted to final positioning of the weil-utterance;
– the weil-clause is separately assertable and rhematic, i.e., the weil-clause carries its own assertion;
– the weil-clause has its own prosodic contour;

---

16 Cf. also English because and cos (Stenström 1998).
17 Cf. also Ortner (1983) on syntactically prominent connectors in German; Schiffrin (1987: 202) on because as a “discourse marker”; and Schleppegrell (1991) on paratactic because.
– they can be used to indicate causal relations in the epistemic and speech-act domain, and to introduce utterances which are only loosely connected to the preceding utterance.\(^\text{18}\)

Thus, there are structural and functional distinctions between the uses of *weil* with verb-final and with verb-second positioning.

### 3.2. Main-clause order in *obwohl*-clauses

 Speakers of colloquial German not only use *weil* to introduce utterances which show characteristics of main-clauses, but they also use the subordinate concessive connector *obwohl* with main-clause syntax.\(^\text{19}\) The following *obwohl*-constructions, which appear frequently in colloquial German, cannot be accounted for within traditional categories of concessive relations. These non-standard variations, which display main-clause order, have – as the analysis will show – specific discourse-pragmatic functions: They introduce restrictions, corrections, and disagreements. The corrections introduced by *obwohl* can range from a partial repair up to a total negation of the previous utterance.

 In the following segment Ulla is talking about a neighbor who suddenly became seriously ill. She comes to the conclusion that one can consider oneself lucky to have been healthy so far (lines 36–37):

(10) **KRANKHEITEN** *(illnesses)*

<table>
<thead>
<tr>
<th>36</th>
<th>Ulla: do: kann man bis=jetzt=no=eigentlich (–)</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>TOI. TOI. TOI (. ) no ganz FROH sei. gell?</td>
</tr>
<tr>
<td>38</td>
<td>(0.5)</td>
</tr>
<tr>
<td>39</td>
<td>OBWOHL man weiß jo gar net was in oim SCLUMMERT.</td>
</tr>
<tr>
<td>40</td>
<td>(1.5)</td>
</tr>
<tr>
<td>41</td>
<td>vielleicht sen mir au scho bald mol DO.GWESE.</td>
</tr>
</tbody>
</table>

\(^{18}\) It is not possible, however, to postulate a strict correlation between (a) *weil* with verb-final positioning and a causal relation operating in the content domain, and (b) *weil* with verb-final positioning and a causal relation operating in the epistemic and speech-act domain (Scheutz 1998, Pasch 1997, Gohl and Günthner 1999).

\(^{19}\) For *obwohl* with main-clause syntax cf. also Métrich (1980); Gaumann (1983); Günthner (1996, 1999, 2000).
In lines 36–37 Ulla states that she is glad that she has always been healthy up to now. However, after a short pause, she produces an obwohl-clause with main-clause syntax: “OBWOHL man weiß ja gar net was in oim SCHLUMMERT” (although one never knows what’s sleeping inside of you). In contrast to concessive obwohl-clauses, the obwohl-construction at hand does not yield the traditional presupposition for concessive relations; i.e., “if p’, then normally ¬ q’” cannot be applied: “If one never knows what’s sleeping inside of you, then up to now one cannot actually be glad”. Instead, the obwohl-utterance restricts the validity of the preceding conclusion that Ulla is healthy.

Since the restriction of a preceding statement or of a possible conclusion from a preceding statement has the character of an afterthought, it is not surprising that the linear order is such that the obwohl-clause is generally post-positioned: Once the “matrix” clause has been stated, the speaker sees the necessity of restricting the validity of this statement or of a possible conclusion which may be drawn from it.

In the following example, Rolf uses obwohl to revise his preceding proposal. The context is as follows: Sara is trying to persuade her brother Rolf to come along to the thermal bath:

(11) THERMALBAD (thermal bath)

\[\begin{align*}
32 & \text{Rolff:} \quad \text{okay. na komm i halt mit.} \\
33 & \text{Sara:} \quad \text{gut.} \\
34 & \text{(-)} \\
35 & \text{Rolff:} \quad \text{obwohl (. ) du i hab echt koi [Zei-] ZEIT für sowas.} \\
36 & \text{Sara:} \quad \text{[( )]} \\
37 & \text{Sara:} \quad \text{hm. "schade." ("wär nett gwese")} \\
32 & \text{Rolff:} \quad \text{okay. I’ll come along then.} \\
33 & \text{Sara:} \quad \text{great.} \\
34 & \text{(-)}
\end{align*}\]
In line 32 Rolf states his decision to come along. However, after a pause, he corrects his former decision and introduces a sudden switch in his perspective: “obwohl (.) du i hab echt koi [Zeit] ZEIT für sowas.” (although (.) I really don’t have [time] for that kind of thing). Thus, the two constructions at hand – the obwohl-clause and the preceding one – are produced in two steps. The obwohl-construction shows main-clause syntax (the finite verb hab is in second position). Non-integrative word order is supported prosodically: The two clauses have their own intonation contours, and the prosodic non-continuity between the clauses is further marked by a second speaker’s turn as well as a pause.

This type of obwohl-clause contradicts one of the central semantic characteristics of concessives: the entailment of “p” and “q”. Concerning its discourse function, this use of obwohl can no longer be subsumed under the traditional definition of concessive relations. Contrary to a concessive relation, in which “the speaker acknowledges the apparently contrary information, but then advances the nucleus anyway, showing that s/he does not regard the two as genuinely incompatible” (Mann and Thompson 1992: 39), the speaker in corrective obwohl-constructions treats the two utterances as incompatible.

Syntactic integration of the two utterances (with final positioning of the finite verb in the obwohl-clause) would suggest a different – i.e., a concessive – interpretation “Okay, dann komm ich halt mit, obwohl ich echt keine Zeit für so was hab” (okay, I will come along, although I really don’t have time for that kind of thing). However, the meaning of the corrective use of obwohl is more like “as against what I just said”.

Corrective obwohl-utterances no longer represent a “satellite” or “a minor act”; on the contrary, they take on the status of a “nucleus” or “major act”, rectifying the preceding statement. The syntactic organization (the main-clause syntax) as well as the prosodic design (prosodic non-integration) iconically reflect this increase in status as a “major act”. Con-

---

21 Cf. Grote, Lenke, and Stede (1997: 92) for “major” (nucleus) and “minor acts” (satellite) in concessions.
sequently, subordinate-clause transformations (e.g., transformations into adverbial phrases) are no longer possible in the case of corrective uses of *obwohl* (i.e., they would change the meaning of the construction): “Trotz meines Vorbeikommens komm ich runter”. Iconicity is also at work in the linear ordering of corrective *obwohl*-utterances: the linear organization “p, *obwohl* q” corresponds to the fact that the *obwohl*-construction is added after the speaker realizes that the preceding utterance must be rectified. Syntactic and prosodic non-integration go hand in hand with the pragmatic independence of the *obwohl*-utterance and the preceding one: both utterances have their own illocutionary forces. Thus, in case of *obwohl* with main-clause syntax, the *obwohl*-clause is no longer a constituent of the preceding matrix clause, but an independent clause introduced by the connector *obwohl*.

*Obwohl* in spoken interaction is not only used by a speaker to correct her or his own utterance, but second speakers frequently use *obwohl* to display their disagreement with their co-participant’s preceding assessments.

In the following example, Lulu and Ela are discussing the movie “Titanic”:

(12) TITANIC

82Lulu: ich fand den gut.
83
84Ela: *obwohl* (. ) der war doch TOTAL KITSCHIG.
82Lulu: I liked it.
83
84Ela: although (. ) it was total kitsch.

After Lulu’s positive assessment about the movie (82), a short pause appears, which may already indicate an upcoming disagreement. Then, Ela uses *obwohl* to introduce her disagreeing second assessment: “*obwohl* (. ) der war doch TOTAL KITSCHIG.” (although (. ) it was total kitsch). *Obwohl* is used here as a disagreement token, in the sense of “well”, “no”, etc. The reading of this *obwohl*-utterance is something like “As against what you just said...”.

---

22 Cf. Givón (1990: 826) for iconicity and syntactic, semantic and pragmatic independence.
Syntactic integration (“ich fand den gut, obwohl er TOTAL KITSCHIG war”) would cause a concessive reading: “I liked it although it was kitch”.

Thus, by using syntactic non-integration, *obwohl* introduced by a second speaker functions similarly to a non-integrative *obwohl*-clause uttered by a single speaker: it limits the validity of the preceding assessment.

To summarize: In colloquial spoken German speakers reinterpret the traditional concessive conjunction *obwohl* as a correction marker to limit or correct the validity of the preceding utterance. Thus, as with *weil*, there are structural and functional distinctions between the uses of verb-final and verb-second positioning with *obwohl*. *Obwohl*-constructions with verb-second position share the following characteristics:

- they are restricted to final positioning of the *obwohl*-utterance;
- the *obwohl*-clause is separately assertable and rhematic (i.e., the *obwohl*-utterance carries its own assertion);
- the *obwohl*-clause has its own prosodic contour (it is often separated from the preceding clause by a pause);
- *obwohl*-clauses with verb-second position are used to correct the preceding utterance or disagree with the first speaker’s preceding assessment.

Thus, there are structural and functional distinctions between the uses of *obwohl* with verb-final and with verb-second positioning.

### 4. From subordinate conjunction to discourse marker?

Variants of *weil*- and *obwohl*-constructions in contemporary spoken German reveal the following features:
<table>
<thead>
<tr>
<th>syntactic features:</th>
<th>weil- and obwohl-clauses with verb-final positioning</th>
<th>weil- and obwohl-clauses with verb-second positioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>subordinate clause transformations:</td>
<td>possible</td>
<td>not possible</td>
</tr>
<tr>
<td>Main-clause phenomena:</td>
<td>not possible</td>
<td>possible</td>
</tr>
<tr>
<td>positioning of the weil- and obwohl- clauses:</td>
<td>initial and final positioning of the weil- and obwohl- clauses</td>
<td>final positioning of the weil- and obwohl- clauses</td>
</tr>
<tr>
<td>prosodic features:</td>
<td>prosodic integration and non-integration</td>
<td>only prosodic non-integration (there is generally a pause between the preceding clause and the weil- or obwohl-utterance); there is often a short pause after weil and obwohl</td>
</tr>
<tr>
<td>semantic features:</td>
<td>• weil: indicating a tight causal relation, (mainly) operating in the content domain; • obwohl: indicating a concessive relation: p and q are valid</td>
<td>• weil: indicating a loose causal connection, often used for causal relations operating in the epistemic or speech-act domain, or used in cases in which the weil-clause is not directly related to the preceding statement; • obwohl: restriction / correction of the preceding utterance; introducing an upcoming disagreement</td>
</tr>
</tbody>
</table>
Lexical-grammatical variation and development

<table>
<thead>
<tr>
<th>Pragmatic features:</th>
<th>the weil- and obwohl - clauses are in the scope of the illocutionary force of the main-clause</th>
<th>the weil- and obwohl - clauses are separately assertible and carry their own illocutionary forces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope of connection:</td>
<td>connection of two clauses</td>
<td>connection of two utterances or longer discursive episodes</td>
</tr>
</tbody>
</table>

If *weil* and *obwohl* with main-clause syntax are losing those syntactic, semantic, pragmatic and functional properties which are characteristic of causal and concessive subordinate conjunctions, the following questions arise: Can these “new” uses of *weil* and *obwohl* in colloquial German cause *weil* and *obwohl* to be treated as coordinate conjunctions? Or are *weil* and *obwohl* developing as discourse markers?

One might argue that in instances in which *weil* with main-clause syntax is used to introduce a causal relation – either in the content or in the epistemic or speech-act domain – it can be treated as a coordinate causal conjunction.24 By combining two different syntagmas, *weil* still functions on the sentence level, and still marks a causal relationship (which can operate in the content, epistemic, or speech-act domain) between these two syntagmas. However, in cases in which *weil* does not express a causal relation between two clauses, and thus does not connect two clauses (or two parts of a clause) to build a complex sentence (Eroms 1980: 76; Lang 1991: 598), but has mainly discourse-organizational functions, it is problematic to talk of *weil* as a “conjunction” in the traditional sense. The increased scope of *weil*, its bleached causal meaning, and its predominant discourse-pragmatic functions (e.g., indicating that the speaker wants to hold the floor, marking a topic shift, adding information) open its classification to debate.

In the case of *obwohl*, it is no longer possible to treat those uses of *obwohl* which introduce a correction sequence as concessive conjunctions. The semantic properties of concessives are lost (the preceding p is no longer valid), and the scope of *obwohl* can extend from that of connecting two clauses to the connection of longer discursive units. Can we, thus, still...

---

24 Cf. Günthner (1993); Pasch (1997); Uhmann (1998). For Northern German varieties, one might argue that *weil* with verb-second order takes over the role of the causal coordinating conjunction *denn*. 

categorize the corrective use of *obwohl* as a conjunction? With the exception of the fact that *obwohl* operates on the discourse level and can connect longer sequences of discourse, what would the semantic relation of this “conjunction” be?

A functional approach to the question of classifying the discourse organizational uses of *weil* and *obwohl* suggests treating them as “discourse markers”, which according to Levinson (1983: 87f.) are part of “discourse deixis”:

...there are many words and phrases in English, and no doubt most languages, that indicate the relationship between an utterance and the prior discourse. Examples are utterance-initial usages of *but, therefore, in conclusion, to the contrary, still, however, anyway, well, besides, actually, all in all, so, after all*, and so on. It is generally conceded that such words have at least a component of meaning that resists truthconditional treatment... What they seem to do is indicate, often in very complex ways, just how the utterance that contains them is a response to, or a continuation of, some portion of prior discourse.

Although the definitions of “discourse markers”, “discourse particles”, “pragmatics markers”, or “discourse connectives” (Schiffrin 1987; Kroon 1995; Brinton 1996; Jucker and Ziv 1998; Fraser 1990, 1996, 1999) vary a great deal, there are a number of characteristics which most of the studies of discourse markers refer to:25

(i) Discourse markers are predominantly a feature of oral discourse;26
(ii) they are drawn primarily from the syntactic classes of conjunctions, adverbials, and prepositional phrases; (iii) because of their frequency in oral language, discourse markers are stylistically stigmatized and negatively evaluated (especially in written or formal discourse); (iv) they are usually “short items”; (v) they appear utterance-initially; (vi) they occur either outside the syntactic structure or loosely attached to it and hence have no clear grammatical functions; (vii) they are optional rather than obligatory in the sentence; (viii) they are marginal grammatical categories;

---

26 This does not imply that discourse markers do not appear in written texts (cf. Kroon 1995 on the use of discourse markers in Latin texts). However, even if we find them in certain written genres (especially in those of “secondary orality”, e.g., e-mail-communication, interviews, etc.), they seem to be associated mainly with spoken language.
Lexical-grammatical variation and development

(ix) they are multifunctional, operating simultaneously on the local (semantic, syntactic, etc.) and global (pragmatic) levels; (x) they can operate on longer discursive sequences, i.e., beyond the scope of clauses.

Discourse-organizational uses of weil and obwohl seem to be instances of discourse markers. They operate on the discourse-organizational level and contextualize (Gumperz 1982) how the following utterance or longer episode is to be interpreted in relation to the preceding sequence. In both cases, I would argue that the type of relation weil and obwohl contextualize, is not arbitrary, but closely connected to the semantics of their traditional uses as causal and concessive subordinate conjunctions: Even in their “new” uses as discourse markers, weil and obwohl function as connecting and cohesive elements (Schiffrin 1987; Lenk 1998). The discourse marker weil still carries (more or less extended) traits of a causal meaning, although it has become bleached. In the case of the correction marker obwohl, we still find the element of dissonance and incompatibility, which marks the concessive obwohl, even if it has lost its conceding meaning. Thus, contrary to Fraser’s (1990: 389) thesis that the pragmatic meaning or function of discourse markers is clearly separate from the semantic meaning of their “homophonous forms”, I would argue that the meaning and function of the discourse markers weil and obwohl are still related to their traditional meaning/function as causal and concessive conjunctions.27 Concerning their syntactic organization, discourse markers weil and obwohl have moved from C-position to the periphery, and are only loosely attached to the preceding syntagma. They take over the position of the “pre-front field”, i.e., a peripheral position in the topology of the German sentence, which is highly relevant for spoken German, even if it does not accord with the canonical written-sentence pattern in German (Auer 1996, 1997; Fiehler 1998). This loose syntactic attachment of weil and obwohl to the following syntagma is underlined by the fact that there are often short pauses after weil and obwohl.

One should add, however, that the distinction between the different uses of weil and obwohl is not clear-cut. Instead there is a continuum ranging from the traditional subordinating conjunctions weil and obwohl via the coordinating use of weil and the restrictive use of obwohl to weil and obwohl as discourse markers.

27 Cf. also Gohl and Günthner (1999).
<table>
<thead>
<tr>
<th>subordinating causal conjunction <em>weil</em></th>
<th>coordinating causal conjunction <em>weil</em></th>
<th>discourse marker <em>weil</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>mainly used to express tight causal relations between clauses in the content domain</td>
<td>mainly used to express loose causal relations (often in the epistemic and speech-act domain);</td>
<td>used for discourse-organizational functions: introduction of additional explanatory information; initiation of a topical change; or used as a “floor-holding device”;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the <em>weil</em>-utterance carries its own assertion.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>subordinating concessive conjunction <em>obwohl</em></th>
<th>hybrid uses of <em>obwohl</em></th>
<th>discourse marker <em>obwohl</em></th>
</tr>
</thead>
</table>
| used to express a concessive relation between two clauses | used to restrict the validity of the preceding clause, utterance or speech-act; | • used to correct the preceding utterance and speech-act;  
• used by second speaker to introduce a disagreeing second assessment; |
|                                              |                         | the *obwohl*-utterance carries its own assertion. |

The meanings or functions of both connectors have changed from the content level to the discourse level, and from connecting propositions to connecting passages of discourse.\(^{28}\) In these new functions, the connectors *weil* and *obwohl* mark how the following utterance or episode is to be interpreted in relation to the preceding utterance or episode.\(^{29}\)

The question arises: Is there a systematic reinterpretation of German subordinating conjunctions to coordinating and finally to discourse mark-


\(^{29}\) Whereas Erman and Kotsinas (1993) talk about “pragmaticalization” for this kind of development from conjunction to discourse markers, Traugott (1995) claims that - due to decategorialization, increase in pragmatic function, polysemy - we have an instance of grammaticalization. Cf. also Auer (1996) and Günthner (1999).
ers? Or is this development limited to weil and obwohl? Besides weil and obwohl, we find similar processes only with adversative während (while) (e.g. “wenn ich mir Köln anschau, dann stimmt, während wenn ich mir Bonn anseh, dann hab ich Probleme damit”; if I look at Cologne, it’s right, while if I look at Bonn, I have problems with it) and with the relative adverb wobei (whereby). Wobei (whereby) has been interpreted as a concessive conjunction: “ich wollte aus der Beziehung heraus, wobei ich sie unheimlich gern mag” (I wanted to get out of the relationship, whereby I really like her a lot), and – similarly to obwohl – wobei has recently been used as a correction marker:

(13) ESSENSEINLADUNG (invitation for dinner)
11Bert: ja KÖNNT Ihr?
12Karl: ja. (-) wobei ich hab am frühen Abend ne Univeranstaltung
13 und weiß nicht genau wann die zuENDE ist.
11Bert: is it okay by you
12Karl: yeah (-) whereby I have a seminar at the university in the early evening
13 and I don’t know when it exactly ends.

The tendency to use main-clause syntax with subordinate connectors is restricted to the causal, concessive, and adversative connectors weil, obwohl, wobei, and während. What is striking is that in these cases the functions of the connectors are transferred from the propositional level to the discourse level. This functional development – which goes hand in hand with syntactic change – is limited to weil, obwohl, wobei and während. Thus, in the case of certain causal, concessive, and adversative connectors there appears to be a tendency in colloquial spoken German to indicate a change in function from the propositional to the discourse level by way of a syntactic

---

30 This example is a simplified version of an actual occurring construction in my data. Cf. also Gaumann (1983) on adversative während (while) with main clause syntax.
32 This example is a simplified version of an actual occurring construction in my data.
33 For wobei as discourse marker in contemporary German cf. Günthner (2001).
change from final positioning of the finite verb (syntactic integration) to second positioning of the finite verb (syntactic non-integration).\textsuperscript{34}

5. Conclusion

Although \textit{weil} and \textit{obwohl}-constructions with verb-second positioning are widespread in spoken German, there is no general tendency to give up subordinate clause order in German. Besides \textit{weil} and \textit{obwohl}, main-clause order is limited to adversative \textit{während} and corrective uses of \textit{wobei}. Instead of a general trend to give up subordinate clause order in German, the analysis shows that in the case of \textit{weil} and \textit{obwohl} (the same holds for adversative \textit{während} and corrective \textit{wobei}) the opposition between verb-second and verb-final positioning is used as a resource for indicating different functions of these connectors; i.e., to differentiate between \textit{weil}- and \textit{obwohl}-clauses which are part of the matrix clause assertion vs. those which carry their own assertion, between tight vs. loose causal connections, and between the concessive connector \textit{obwohl} and the discourse marker \textit{obwohl}.

In general, there seems to be a tendency in spoken German to use verb-second position – and thus, main-clause syntax – with traditional subordinate causal and concessive connectors to indicate that the causal relationship (in case of \textit{weil}) or the incompatibility (in case of \textit{obwohl}) is not between propositions but between utterances and discourse units. Thus, we can observe a functional change from meaning based on the referential/propositional level to meaning based on the discourse level; or from sentence grammar to discourse grammar.

Appendix

Transcription Conventions

(The transcription conventions follow GAT “Gesprächsanalytisches Transkriptionssystem”, Selting et al. 1998):

\textsuperscript{34} Cf. also Pasch (1997).
Lexical-grammatical variation and development

[ja das] finde ich conversational overlap;
[du ab] short pauses of less than 0.5 sec.;
( - ) pauses of 0.5 sec. and longer;
( ) unintelligible;
(gestern) uncertain transcription;
= latching;
? intonation phrase-final: rising;
. intonation phrase-final: slightly rising;
, intonation phrase-final: falling;
, intonation phrase-final: slightly falling;
<<↑>word word> global high pitch;
<<↓>word word> global low pitch;
a: lengthening;
"no" soft voice;
NEIN loud voice;
mo((hi))mentan laugh particles within the utterance
hahaha laughter;
((hustet)) nonlexical phenomena (e.g. coughing).

References

Arndt, Erwin

Auer, Peter

Brinton, Laurel J.

Couper-Kuhlen, Elizabeth

Crevels, Mily
Erman, Britt and Ulla-Britt Kotsinas

Eroms, Hans-Werner

Fiehler, Reinhard

Fraser, Bruce

Gaumann, Ulrike

Givón, Talmy

Gohl, Christine and Susanne Günthner

Green, Georgia M.

Grote, B., Nils Lenke, and Manfred Stede

Gumperz, John J.

Günthner, Susanne
2000 From concessive connector to discourse marker: The use of obwohl in everyday German interaction. In *Cause, Condition, Concession*,


Maurer, Friedrich

Métrich, René

Ortner, Hanspeter

Pasch, Renate

Pomerantz, Anita

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson

Scheutz, Hannes

Schiffrin, Deborah

Schleppegrell, Mary J.

Schlobinski, Peter

Selting, Margret


Subject index

a posteriori 179–180, 182
a posteriori truth 182
a priori 175, 179–190, 194
a priori truth 182
a priori/a posteriori 182, 190
Abbasids 120–124, 126–127, 129
ǂabd allâh 125, 128
Abu Bakr 120
account of baptizing 229
actual thinking 181, 191
adaptation 351, 357–359, 361–364, 369
adaptation of the semantic structure 358, 362
AEIEW 84, 110
Afghanistan 119
AHD 79–81, 83, 85, 112
Aktionsart 254, 259, 261, 263–264, 269, 279, 330, 332
Aleppo 127, 130
Almoravids 121, 126
‘Ali 123, 126
Altenglisches Etymologisches Wörterbuch
see AEEW
American Heritage Dictionary
see AHD
American Heritage Electronic Dictionary
see AHD
anîr al-mu‘minîn 118, 120–121, 126
anîr al-muslimîn 121, 126
analogical levelling 25
analytic 3, 5, 8, 19–20, 175, 178–195, 199–201
analytic and synthetic judgments 19, 179, 190
analytic judgment 175, 178, 180, 182–183, 186, 188–190, 193, 199
analytic sentences 199
analytic statement 186, 200–201
analytic truth 175, 182–183, 188, 200
analytic/synthetic 182, 188, 190
analytical philosophy 19, 22
analyticity 178, 181–182, 186–191, 194, 199–201
Ancient Greek 3
Andalus 119, 122, 124, 126–127
Anglo-Saxon Dictionary
see ASD
anthropology 1–2, 14–15, 28
antonymy 150, 166
aorist 368–369
apodosis 370
Apollonius Dyscolus 252
apriority 189–190
apriority 212, 215
Arabic 16, 115–119, 126, 128–129, 133
Aramaic 117
arbor porphyriana 40, 42
argument structure 24
arithmetic 181, 183–184, 191
articles 13, 18, 22, 26, 28
artificial languages 188
ASD 84, 90–91, 109
aspect 300–301, 307–308, 313, 329–330, 332, 342, 368–369
Baghdad 120–124, 126–127
baptizing 226, 230–231, 238–239, 241
belief change 198
belief revision 18, 20, 193–194
Bengali 297, 310, 316–317, 319–320, 322, 346–347, 349
Bible 117
big 192–193, 196
Bilingualism 25
BNC 81, 102, 109
British National Corpus see BNC
Brown 81, 102
Bukhara 119
Bulgarian 353, 367
caliphate 116–117, 121, 123, 125
127–129, 132
case 353–354, 357, 359, 360–365, 369, 371
case marking 308–309
categorization 6–7, 21, 28, 227, 229–230, 236, 238, 242
categorization theory 7
causal theory of reference 225, 229–230, 235
causality 377, 381
change in beliefs 191
change in language 191, 200
character 207, 211–214, 216–219, 221
character of natural-kind terms 213
choice 185, 191, 195
classical philology 4, 22–24
classification criteria 236
clause chaining 316, 318, 320, 325–326
clause fusion 305, 343
clause union 305, 321–322, 326
cognitive abilities of humans 6
cognitive linguistics 7, 14, 37, 43–44, 136
coins 120–122, 125
colloquial language 177
compounding 46–47, 49
conceptual framework 185, 191
conceptual knowledge 178, 193
conceptual shift 158
conceptual structure 14
conceptual thinking 184
conceptualization 8, 12, 14, 44–45, 52, 56–58, 62
conditional 370–371
conditions 371
confirmation 180, 191, 198
confirmed 179–180, 186
conjunction 26, 375–376, 378, 380, 381, 383, 386, 391, 393–396
conjunctive 370
conjunctive participle 316
contact language 26
count of baptizing 237
count of use 236–237, 239–242
contingency 212
contiguity 42–43, 46, 48–51, 53–57, 63
contrast 46–48
coordinate 380, 383, 393
coordinating 393, 395–396
coordinator 376
conversion 46, 49, 51, 53
corpus 27
correction 391–393, 395, 397

criterion to diagnose meaning change 228
Croatian 352, 356–357, 360, 365, 369–370, 372
cultural history 3, 10
cultural science 5, 7, 11, 28
culture-specific 40
decision 363
default knowledge 225, 244
‘Defender of the Faith’ 118
definite article 367–368
definiteness 368
definition 178, 182, 184, 186, 188, 192, 195–196
deliberate definition 236
DEPN 91, 110
diagonal 212, 215–216, 221
Dionysius Thrax 252
discourse analysis 13, 22
discourse grammar 398
discourse marker 26–27, 375, 386, 391–400, 403
discourse particle 394
division of labor 18
division of linguistic labor 135, 137–139, 167
double lexicon 361–362
Dutch 118
EDD 79, 84–91, 113
empirical knowledge 180
encyclopedic knowledge 136
Encyclopédie 145, 160, 170
Engadinian 40, 53, 57
English 9, 14–15, 23, 25, 28, 32, 40–45, 50–51, 56, 62
English Dialect Dictionary see EDD
English Place-Name Society see EPNS
engynomy 42
entrenchment 42
epistemic aspects of meaning 210–211
epistemic domain 381
epistemic necessity 212
epistemic notion of meaning 208, 212, 221
EPNS 76, 84–85, 87–88, 90–91, 109, 112
ergative 309
Ethiopic 117
expert language 177
Extended Now 257, 274, 277, 279–288, 293
Extended-Now-analysis 253
Fâtîma 123
Fatimids 123–124
formal languages 188
formalization 185–186, 188
framework 185
Frege-analyticity 200
French 45, 50–51, 118
function 226, 231–237, 239, 242
functional projection 255, 257
future 357, 368–370
future tense 303, 313–315, 321, 328, 334–336
gender 363–365
generalization 23
genuine 178
gold 175–176, 195
grammaticalization cline 24, 297, 301–302, 304, 313–314, 328
grammaticalization phenomena 7–8
grammemes 365–369, 372
Greek 23–24, 57, 59
Greek perfect 23–24
‘Guardian of the Two Holy Cities’ 118
Hebrew 117–118
Heir-Apparent Principle 305–306, 342–343
hermeneutic method 17
hidden indexical 209–215
hidden indexicality of natural-kind terms 213
Hindi 24
historical pragmatics 6, 11
historical semantics 136, 139
history 1–2, 10, 12, 16, 20-24, 28
history of words 136–137, 166
Holism 188, 191, 196–197, 204
holistic 176, 194
homonyms 355
humanity 1, 4–5, 13, 28
hyponymy 150, 166, 358
hyperonymy 162, 166
idiom 37, 39, 45–49, 52
inān 118, 120
imperfect 368–371

Subject index 407
imperfective 262–263, 307, 330, 332, 359, 368–369

implicit definitions 178
incommensurability 196
indefinite article 367–368
Indo-Aryan 24
informativity 212–213, 216
intensifier 9
interdisciplinary 2, 31
interest-dependence 194
irreal 370
irrealis 371
Ishmael 116
Ismâîl 116, 132
Italian 25–26, 118, 351–373

jude 16, 22, 228–229
Jordan 121
judgment 175, 178–195, 199, 201

khalîfa 16–17, 115–120, 128
khâlifat Allâh 120
Khilafat 128
Khorazm 119
knowledge a priori 179–180
knowledge increase 12
knowledge of facts 186, 199
knowledge of language 199
knowledge of meanings 178, 186, 193

language 177–196, 199, 201
language change - by social interaction 8
language acquisition 352–354
language and theory formation 187
language change 193–194, 351, 353
language contact 22, 25, 351–355, 362–367
language decay 3, 9
language development 3, 8, 27
language external 15
language innovation 10
language-external forces in meaning change 22
language-internal perspective 6
Late Latin 9
Latin 3, 10, 12, 50, 57–59

law cluster concept 189
law of identity or contradiction 180
legitimacy 122–124, 127
levelling 26
lexical change 39, 45, 50
lexical field 6–7, 14–15
lexical innovation 39, 45–49, 59
lexical knowledge 20, 22, 136
lexical semantic relation 150, 159, 162
lexical structure 14
lexicology 38–39, 44–45, 48–49
light verbs 24, 26, 28, 295–315, 323, 328–344
linguistic division of labor 238
linguistic turn 5
literary writing 10
loanword 39, 45, 362, 365
logic 181, 184, 187–189, 193

Macedonian 353, 367
Madîna 118
main-clause order 381–383, 387
malîk 118
Mamluks 127
Marathi 298–299, 316
mathematics 179, 181–185, 202
meaning by definition 235
meaning by ostension 235, 243
meaning change 1–2, 5–7, 12–22, 26, 28, 192–199
meaning constitutivity 195
meaning postulates 243
measure adverbial 278
Mecca 118
metaphor 37, 39, 45–48, 302–303, 335
metaphorical 37–39, 46–53, 155
metaphysical 207, 212–215, 221
metonymic 39, 150, 152
metonymy 6–8, 14, 45
minority language 351, 358, 362
model language 354, 358, 368, 372
model theory 20
Molise 25
Molise 351–353, 358, 367
Molisin Slavic 25–26
Molitian Slavic 351–353, 367, 369–370, 372
Mongols 127
mood 368, 370
MuhSammad 116, 120–121, 123
murmured (voiceless) vowels 355, 372
Muslims 116, 123, 128–131
natural kind term 16–21, 207, 209–211, 213, 217, 226–229
natural languages 188
necessity future 370
negative interpretation 179, 186
negative judgment 27
negative notion 189
nonstandard logics 20
normative sense 181–182
North Africa 123

observation sentences 198
OED 75, 79-81, 84–91, 112
Old Greek 23
onomasiology 14, 37–38, 40, 43–45, 48–53, 59
onset uses 10
Ottomans 118, 127

Páli 315
paradigm change 19
paradigm examples 216
paradigmatic application 214, 218
paradigmatic referents 215–219
paradigms 209, 214–216, 219
paucal 367
pejorative connotation 141, 164–165, 168
pejorization 164
perfect 368–369
perfective aspect 297, 300, 329–330, 347
perspectivity 194
philological evaluation 17

philology 1–2
philosophical realism 227, 243
philosophy 1–2, 5, 19, 21, 32
philosophy of science 19, 21, 186
phonology 372
physiology 5
pluperfect 369
politics 16
positive interpretation 179, 182
potential 354, 370
potentialis 371
pragmatic inferencing 23
Present Perfect-puzzle 288
primary language 352, 355, 367
probability future 370
proper names 210, 213
prosodic 378, 382, 386, 389–392
prosodically 378, 382, 389
protasis 370
psychological sense 181–182
psychology 5–7, 11, 28
qua question 230–231
Qur’án 117

realistic semantics 208
reanalysis 305–306, 343
reduplication 307–312
reference 178, 181, 195
reference change 16, 18
referential theory of meaning 21
replica language 353–355, 365
resultativity 252–253, 263, 269–270, 274, 288
retractability 195
revisability 189–190, 199–201
revision strategies 187
revisions 187–189, 191–193, 195
Rheto-Romance 40–41
Subject index

Romance 3, 9, 37, 57, 59, 63, 351–355, 359–363, 367–369
root, Arabic 116
Samarkand 119
samples 226–235, 237, 240–242, 244
Sanskrit 295, 307, 313–326, 336, 343, 349
Saudi Arabia 121
schema transfer 303
schematic meaning 303–304, 344
secondary language 352, 353, 362, 368
SED 74, 78, 81, 88–90, 102, 104, 111
self 9–10, 30
semantic bleaching 296–298, 303, 335
semantic change 136, 152
semantic loan 45
semantics 1–2, 5–8, 11, 14–17, 20–28, 31, 207–211, 222
semasiology 14, 38–41
’servant of God’ 126
servus servorum dei 128
‘Shadow of God on Earth’ 118
Shape 235–236, 246
Shī‘īs 123
similarity 46–55, 231, 236–239, 241–242, 244–245
similarity relation 229–231, 235, 242, 244
Slavic 25–26, 351–364, 367–370
small 192
source language 353, 363, 365
Spain, Islamic 119–124, 132
Spanish 46, 50–51, 118
speaker community 227, 237–241
specialization 135, 137, 167–168
spoken German 26
stereotype 22, 210–211, 225, 244–245
stipulative definitions 201
structuralist tradition 7
subordinate 375–400
subordination 47, 49, 51, 53, 56, 375–376, 395–396
substance 226, 228, 237, 239, 241
substrate 352–353
suffixation 45, 49, 51, 56–57
superordination 47
superstrate 352
Survey of English Dialects see SED
Swedish 23
synonymy 150, 166, 186–200
syntactic change 304–305
syntactic fusion 304
syntax-semantic interface 8, 24
synthetic 178–190
synthetic a priori 179
synthetic judgments 178, 190
synthetic judgments a priori 178, 181, 190
synthetic truth 182
target language 354, 372
taxonomy 42
tense 368–370
tense and aspect markers 23
tense systems 24
text analysis 17
The American Heritage Talking Dictionary see AHD
The Oxford English Dictionary see OED
theoretical concepts 176, 201
theoretical terms 176, 178, 193, 195, 199
theory 175–178, 182, 185–201
theory change 20, 175, 192–199
theory changes 208
theory community 187
theory revision 193, 194, 199
tonal system 372
Turkestan 119
Turkish 119, 127, 132
Turkish Republic 127
Twin Earth 209, 213, 221
two-dimensional construction 211
two-dimensional semantics 20, 212
Umayyads 120–126, 133
underdifferentiation 360, 371, 372
universalist 4, 6, 11
Subject index

Urdu 24, 26, 295–345
usage 207, 209–210, 214–221

Uthmān 120

vagueness 194
Vedic 315, 318

verb-final position 377–387, 398

verb-second order 376, 393
verb-second position 376, 380–381, 383, 386–387, 391, 398
Vicarius Christi 128, 131

weil 26, 375–387, 391–403

word formation 39, 45–49, 55
word histories 241
## Author index

<table>
<thead>
<tr>
<th>Author</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gohl, Christine</td>
<td>376, 381, 384, 387, 395</td>
</tr>
<tr>
<td>Goodenough, Ward H.</td>
<td>72</td>
</tr>
<tr>
<td>Goodman, Nelson</td>
<td>230</td>
</tr>
<tr>
<td>Greenberg, Joseph H.</td>
<td>69, 71</td>
</tr>
<tr>
<td>Günthner, Susanne</td>
<td>13, 23, 26–28, 381, 384, 387, 389–397</td>
</tr>
<tr>
<td>Haas-Spohn, Ulrike</td>
<td>13, 18–20</td>
</tr>
<tr>
<td>Hansson, Sven Ove</td>
<td>193</td>
</tr>
<tr>
<td>Harman, Gilbert</td>
<td>201</td>
</tr>
<tr>
<td>Harris, Alice</td>
<td>304–306, 312, 342–343</td>
</tr>
<tr>
<td>Heim, Irene</td>
<td>259, 261</td>
</tr>
<tr>
<td>Heine, Bernd</td>
<td>8–10</td>
</tr>
<tr>
<td>Hendriksen, Hans</td>
<td>322, 325–327</td>
</tr>
<tr>
<td>Hinds, Martin</td>
<td>120</td>
</tr>
<tr>
<td>Hinzen, Wolfram</td>
<td>175</td>
</tr>
<tr>
<td>Hook, Peter Edwin</td>
<td>296–301, 316, 318, 323, 327, 329, 338</td>
</tr>
<tr>
<td>Hopper, Paul</td>
<td>8–9, 296, 301, 304, 314, 355</td>
</tr>
<tr>
<td>Hoyningen-Huene, Paul</td>
<td>175, 181</td>
</tr>
<tr>
<td>von Humboldt, Wilhelm</td>
<td>1–4, 9, 25</td>
</tr>
<tr>
<td>Hünnemeyer, Friederike</td>
<td>8, 10</td>
</tr>
<tr>
<td>Indurkhya, Bipin</td>
<td>8</td>
</tr>
<tr>
<td>Iviç, Pavle</td>
<td>351</td>
</tr>
<tr>
<td>Jackendoff, Ray</td>
<td>243</td>
</tr>
<tr>
<td>Jackson, Frank</td>
<td>213</td>
</tr>
<tr>
<td>Janda, Josef</td>
<td>8, 27</td>
</tr>
<tr>
<td>Johnson, Mark</td>
<td>8, 37, 43</td>
</tr>
<tr>
<td>Jonsson, Hans</td>
<td>89</td>
</tr>
<tr>
<td>Kadmon, Nirit</td>
<td>6</td>
</tr>
<tr>
<td>Kamp, Hans</td>
<td>6</td>
</tr>
<tr>
<td>Kant, Immanuel</td>
<td>19, 175, 178–184, 189–191</td>
</tr>
<tr>
<td>Kaplan, David</td>
<td>20, 207, 211–215, 221</td>
</tr>
<tr>
<td>Kaufmann, Terrence</td>
<td>352</td>
</tr>
<tr>
<td>Kay, Paul</td>
<td>68–69, 72</td>
</tr>
<tr>
<td>Keller, Rudi</td>
<td>6</td>
</tr>
<tr>
<td>Kempton, Willet</td>
<td>69, 72</td>
</tr>
<tr>
<td>Kiefer, Ferenc</td>
<td>7</td>
</tr>
<tr>
<td>Kitcher, Philip</td>
<td>189, 220</td>
</tr>
<tr>
<td>Koch, Peter</td>
<td>37–39, 41–42, 45, 52, 56–57</td>
</tr>
<tr>
<td>König, Ekkehard</td>
<td>8, 11, 375, 377, 379, 382</td>
</tr>
<tr>
<td>Krefeld, Thomas</td>
<td>44, 58–59</td>
</tr>
<tr>
<td>Kripke, Saul A.</td>
<td>176, 179, 195, 210, 213, 215, 225</td>
</tr>
<tr>
<td>Kristensson, Gillis</td>
<td>91</td>
</tr>
<tr>
<td>Kronenfeld, David</td>
<td>12, 14–15, 17, 21, 28, 67, 69, 71, 72</td>
</tr>
<tr>
<td>Kuhn, Thomas S.</td>
<td>19, 196, 208</td>
</tr>
<tr>
<td>Kuteva, Tania</td>
<td>9</td>
</tr>
<tr>
<td>Lakatos, Imre</td>
<td>196</td>
</tr>
<tr>
<td>Lakoff, George</td>
<td>8, 72</td>
</tr>
<tr>
<td>Lakoff, John</td>
<td>37, 43</td>
</tr>
<tr>
<td>Landman, Fred</td>
<td>264</td>
</tr>
<tr>
<td>Langacker, Ronald W.</td>
<td>72</td>
</tr>
<tr>
<td>Lazarus</td>
<td>4</td>
</tr>
<tr>
<td>Lehmann, Christian</td>
<td>8</td>
</tr>
<tr>
<td>Lehrer, Adrienne</td>
<td>68</td>
</tr>
<tr>
<td>Leibniz, Gottfried Wilhelm</td>
<td>von 175–177</td>
</tr>
<tr>
<td>Levi, Isaac</td>
<td>175</td>
</tr>
<tr>
<td>Lewis, D.</td>
<td>189</td>
</tr>
<tr>
<td>Lightfoot, David</td>
<td>305</td>
</tr>
<tr>
<td>Locke, John</td>
<td>175, 195, 200</td>
</tr>
<tr>
<td>Löfvenberg, Mattia T.</td>
<td>84, 86, 91</td>
</tr>
<tr>
<td>Louonsbury, Floyd G.</td>
<td>69, 72</td>
</tr>
<tr>
<td>Lyons, John</td>
<td>67–68</td>
</tr>
<tr>
<td>Makinson, David</td>
<td>193</td>
</tr>
<tr>
<td>Mandilaras, Basileios G.</td>
<td>282, 288</td>
</tr>
<tr>
<td>McCooard, Robert W.</td>
<td>253–254, 272, 274, 278</td>
</tr>
<tr>
<td>Meillet, Antoine</td>
<td>6</td>
</tr>
<tr>
<td>Meinschaefer, Judith</td>
<td>12, 17–18</td>
</tr>
<tr>
<td>Meißner, Konrad</td>
<td>300, 327</td>
</tr>
<tr>
<td>Mey, Jacob L.</td>
<td>6</td>
</tr>
<tr>
<td>Moens, Marc</td>
<td>331</td>
</tr>
<tr>
<td>Montague, Richard</td>
<td>6, 225, 232</td>
</tr>
<tr>
<td>Nahin, Paul J.</td>
<td>184</td>
</tr>
<tr>
<td>Nerlich, Brigitte</td>
<td>73</td>
</tr>
</tbody>
</table>
Pagin, Peter 200–201
Pagliuca, William 8
Pasch, Renate 376, 381, 383, 387, 393, 398
Paul, Hermann 4
Piccoli, Giovanni 351, 372
Piñón, Christopher 331
Plank, Frans 3
Popper, Karl 196
Pustejovsky, James 6
Quine, Willard Van Orman 19, 178, 180, 185–191, 193, 199–200
Reichenbach, Hans 185, 252–254, 260, 280
Reichenkron, Günther 352
Reyle, Uwe 6
Roberts, Ian 304–305, 343
Romney, A. Kimball 72
Rosch, Eleanor H. 69, 72
Rott, Hans 13, 18–20, 175, 183, 193, 196
Roussou, Anna 304–305, 343

de Saussure, Ferdinand 67–69, 72, Schiffrin, Deborah 384, 386, 394–395
Schlegel, Friedrich von 3–4
Schleiermacher, Friedrich 190
Schlick, Moritz 178
Schulze, G.E. 190
Shapere, Dudley 196
Stalnaker, Robert 212
von Stechow, Arnim 13, 23–24, 28, 251, 257, 274, 278, 285
Steedman, Mark 331
Stegmüller, Wolfgang 176
Steinthal, Heyman 4
Sterelny, Kim 13, 227, 229–230
Stern, Gustaf 6, 12
Sweetser, Eve 8, 302–303, 377–378, 382
Tagliavini, Carlo 52–54
Thomason, Sarah Grey 352
Tikkanen, Bertil 316–319, 321–324, 326, 336
Trabant, Jürgen 3
Traugott, Elizabeth 8–9, 11, 296, 301–302, 304, 314, 355, 396
Trendelenburg, Friedrich Adolf 190
Trier, Jost 7, 14, 67–69
Tyler, Stephen 69
Ullmann, Stephan 6, 67–68
Vendler, Zeno 259
Verspoor, Marjolijn 6
Wasserstein, David 12, 16–18
Weinreich, Uriel 353
Wischer, Ilse 8
Wittgenstein, Ludwig 177–178, 187, 189