

ASSIDUE

Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt

Nr. 2

Texts and communicative practices

An encoding scheme

Stavros Skopeteas

September 2002

Impressum:

Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt

Seminar für Sprachwissenschaft

Philosophische Fakultät

Universität

D - 99105 Erfurt

Herausgeber:

Prof. Dr. Christian Lehmann

© Stavros Skopeteas

ABSTRACT¹

The principle that linguistic activity can only be understood in relation to the surrounding communicative events goes back to the Prague school linguistics and it is taken as evident in most functional studies on language for the last decades. In this respect there is a need for language documentation to deal not only with the archiving of texts but also with the description of their context, i.e. the communicative events and cultural environments, within which they are produced. This paper offers an encoding scheme for this purpose in form of *attribute-value matrices*.

1. Preliminaries

The requirements of a functional textual description are somewhat discussed in the language documentation studies. It is pointed out that a language documentation should consider the properties of texts as parts of communicative events, such properties being e.g. the producer and his purpose, the settings of the communicative situation, the relation among the speech participants, etc. (cf. Lenk 1996; Lehmann 1999:12-14; 2001:92-95). The objectives of the documentation of the communicative environment of texts are manifold. The most relevant issue for the description of the linguistic system is the conditioning of linguistic variation. In many cases variation at the linguistic level can be described as related to variable properties at the communicative level. A further issue is that documentation should render a basis in order to draw an ethnography of communication (Hymes 1974), namely to describe the relation between the linguistic practice and the community's life.

2. Domains

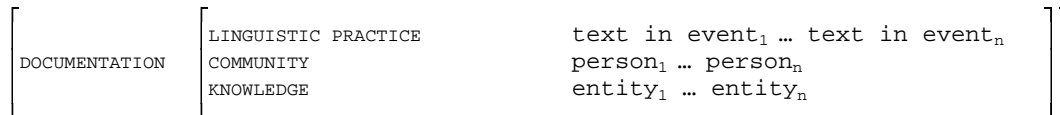
The encoding scheme describes information in three **domains** (cf. Figure 1):

- The domain of **linguistic practice** is the core of the documentation: The unit of this domain is a description of a text with respect to its content and its form, embedded in a description of the communicative event, within which the text is produced.

¹ The present paper has been presented at the *Max Planck Institute for Evolutionary Anthropology*, Leipzig (14.05.2001). A preliminary version of these ideas was the subject of numerous discussions with Wolfgang Kesselheim.

- The domain of the **community** describes **persons** that are involved in the production and reception of texts, e.g. authors, informants, etc. Persons are described in two perspectives: as individuals and as members of groups.
- The last domain includes **knowledge** of the community about entities of any kind, like physical objects, artifacts, places, particular dates or periods etc. and also knowledge about persons or groups of the community.

Figure 1: Domains



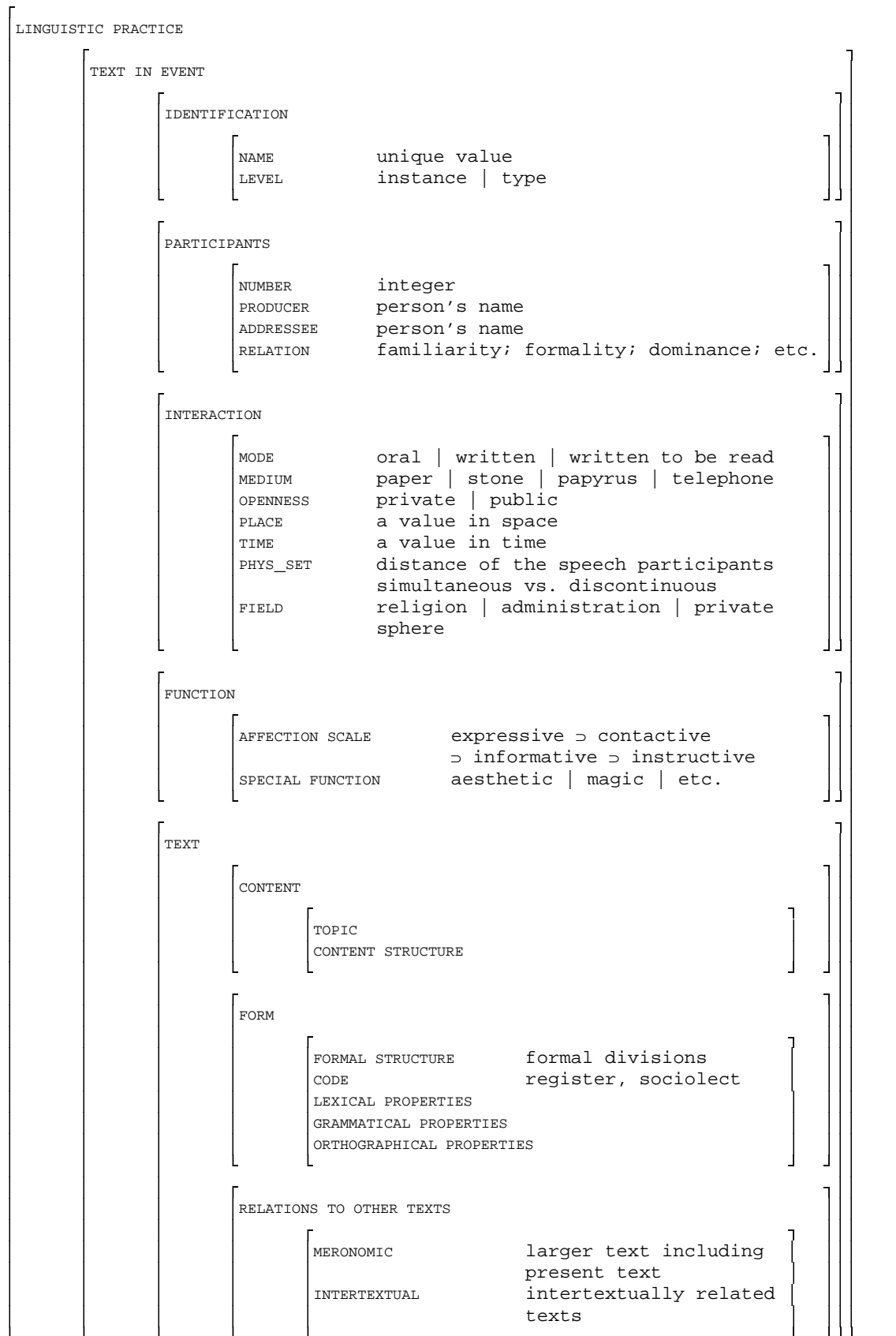
3. Properties

3.1. Linguistic practice

Figure 2 presents the properties used in the domain of linguistic practice. A “text and event” unit is being described according to five sets of properties: The first set **identifies** the unit with a unique name for the text and a characterization with respect to the level of description, if it is an instance and a type (see §4). The second set provides information about the **participants** in the communicative event, the producer and the addressee, and the relation between them (familiar, formal etc.). The third area contains information about the **interaction**, in particular:

- if it’s oral or written communication,
- the medium, including channels of oral communication and writing materials,
- a characterization according to the openness of the message – if it is private or public –,
- values for the place and the time of the interaction,
- the settings of the interaction, e.g. the distance between the participants in space, the simultaneity of the interaction etc.
- and finally, the cultural field, within which this interaction is embedded: e.g. religion, administration, private life etc.

Figure 2: Properties of texts/communicative events



The next set of properties in Figure 2 describes the **role**, that the text is supposed to take in this interaction. The first distinction is a hierarchy of the functions of linguistic use. There are numerous classifications for this domain, the one presented here is adopted from Heinemann

& Viehweger (1991:146ff.). According to this hierarchy, the primary function involved with each text is to produce an expression, the next function to the right is to make a contact, the next function is to provide some information, and the most affective function is to give an instruction. Every function in this affection hierarchy includes all functions to the left. Furthermore there are also special functions that cannot be reduced to this first classification; examples of this kind are poetical or magical texts.

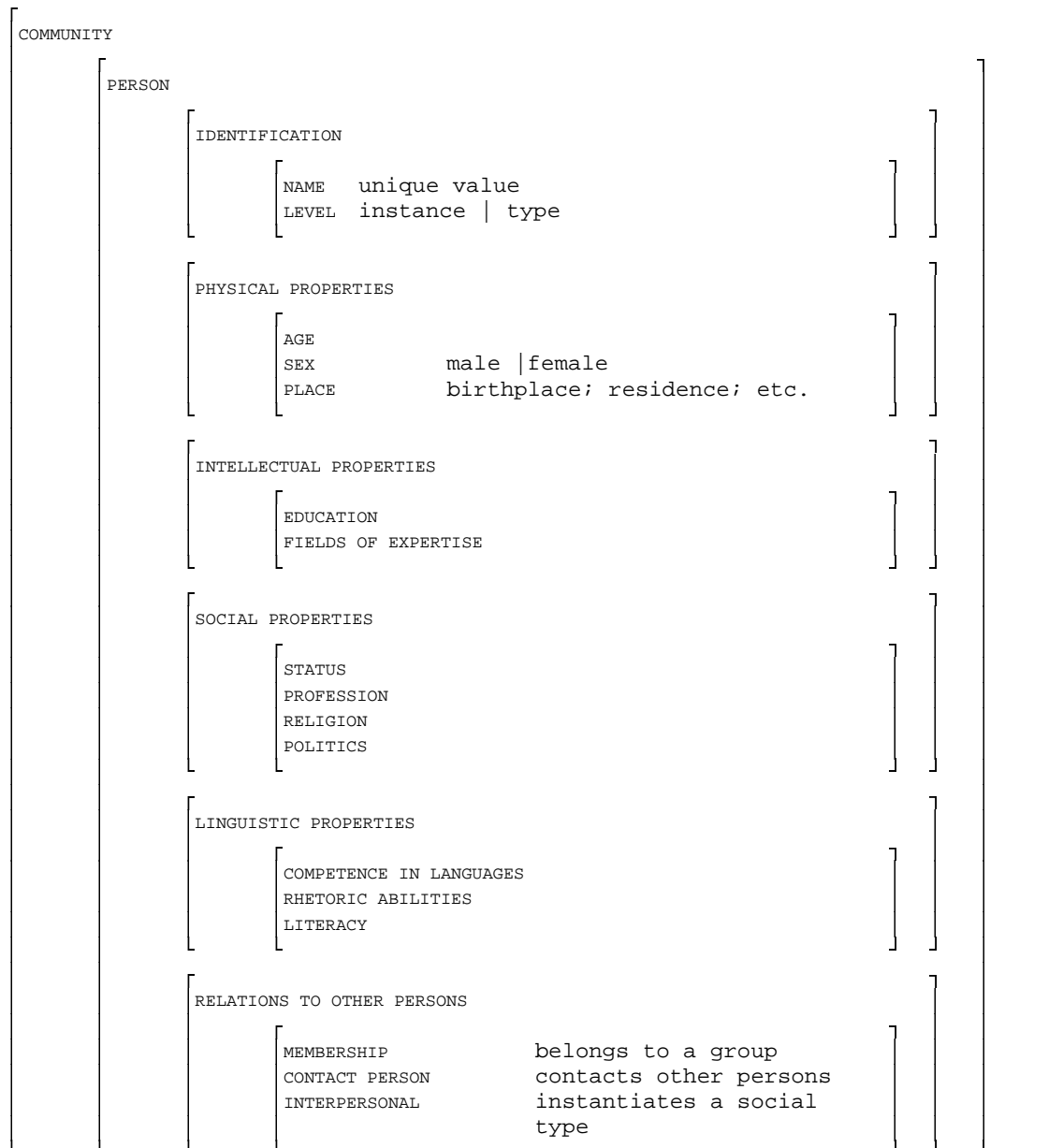
The next set contains properties of **texts** concerning their content (topic and content structure), their form (formal structure, code, lexical, grammatical and orthographical properties) and their relations to other texts. There are two different kind of relations that are included here: firstly meronomic relations that relate textual parts and wholes and secondly intertextual relations to other texts like the original text of a translation, the source of a parody or a citation etc.

3.2. Community

In the same way the domain of community contains entries concerning individuals or groups. Each such unit is characterized according to five sets of properties, that can be seen in Figure 3.

- physical properties like age, sex and relevant places (birthplace or residence),
- intellectual properties like education, and other fields of knowledge,
- social properties like social status, profession and religion,
- linguistic properties like the competence in different languages, rhetoric abilities and literacy,
- and relations to other persons with means of participation in groups of persons.

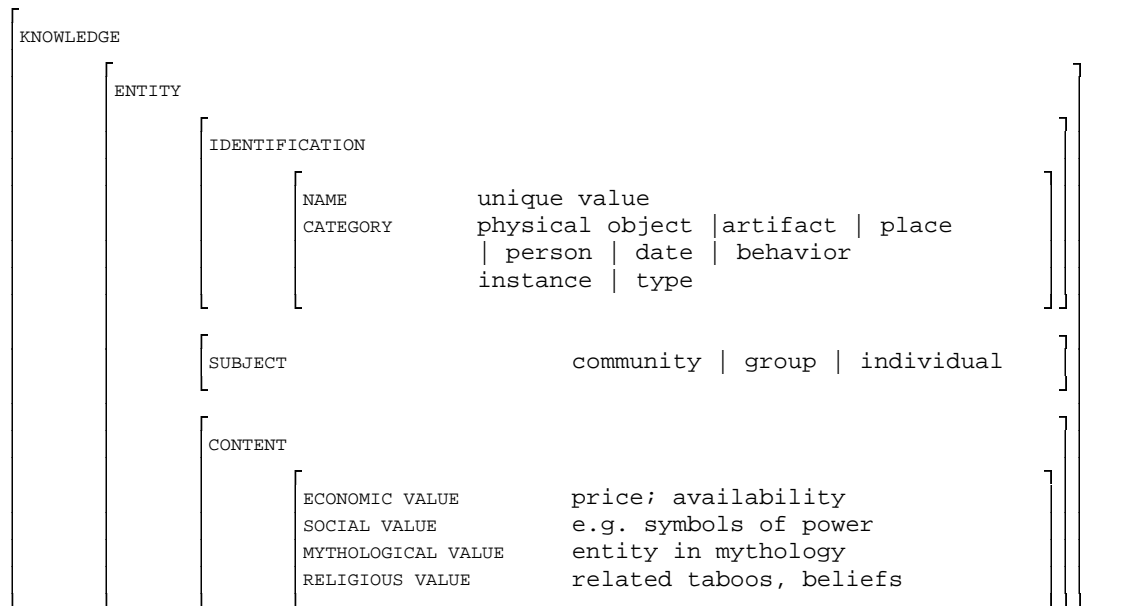
Figure 3: Properties of persons



3.3. Knowledge

The last domain is the domain of knowledge. This domain includes knowledge of all kinds of entities, physical objects, artifacts, places, persons, etc. that is relevant for the understanding of a text. From the point of view of the ontology of communication this component is the symbolic representation of the community's knowledge. However relevant for a documentation program is **knowledge related to communicative practices**, e.g. a prohibition of oral communication in some place, a custom including ritualized text that takes place at a particular date, the price and availability of artifacts that serve as writing materials, etc. The information structure of this domain is shown in Figure 4.

Figure 4: Properties of knowledge about entities



Each unit describes an entity of one category that is specified in the attribute “identification”. This entity is related to the specific culture with a content, that is an economic value like price or availability, a social value like particular objects that serve as symbols of power, and in the same sense a mythological or religious value etc. This knowledge is a property of a community subset: it can either be knowledge shared by the whole community, or knowledge of a group of persons – e.g. knowledge shared in a religious group – or even knowledge of a sole individual.

4. Levels

In general there are two levels of representations that are distinguished in the encoding scheme, as it is illustrated in Table 1: **instances** and **types**. **Instances** are particular occurrences of the units. The instances of the domain of linguistic practice are particular **texts**, and the instances of the domain of community are **persons** and **groups**. The level of instances is exemplified in Table 1 by a simple classical attic dedication, written on three thrones.

Types are abstractions from a set of instances. There are **culture-specific types**, that represent a constellation of properties that have emerged in a particular community and more **abstract types**, that are underspecified with respect to culture-specific properties, and so they apply to the description of texts in different cultures. The different levels of description are exemplified in Table 1 with the discourse type “dedication”. In the attic culture the

prototypical notion of a dedication is a statement that appears as inscription on valuable objects, that are donated mostly to the gods, whereas in the contemporary western culture the prototypical notion of a dedication is a message written at the beginning of a book or a statement made before a play or concert, as a sign of affection or respect for a person. These are culture specific discourse types. At a higher level of abstraction, we can give a description of the type “dedication” with reduced properties, leaving apart for example the property of the medium and writing material, and thus we arrive to a definition, that applies both to the classical attic as to the modern western type of dedication.

Table 1: Levels

instances	<p>e.g. IG II(2) 3108 (written on three thrones):</p> <p>“a citizen of Rhamnous dedicated after winning in the comedy competition”</p>
types	<p>specific <i>culture specific discourse types:</i></p> <p>“a type of attic inscription, written on different objects, that are dedicated typically to the god; they consist of simple sentences including the dedicating person, the addressee of the dedication, sometimes the dedicated object and sometimes the purpose of the dedication” (classical attic discourse type)</p> <p>“a message written at the beginning of a book or a statement made before a play or piece of music performed, as a sign of affection or respect for someone” (contemporary western discourse type)</p> <p>generic <i>abstract discourse types that hold for many languages:</i></p> <p>“a performative statement that specifies the addressee of an offer/donation”</p>

4.1. Instances

The documentation of an instance in the domain of linguistic practice is presented in Figure 5. The instance described is a scratch on pottery dating at about 350 BC. It contains a number of curses concerning different persons as it is specified under “topic”. The “lexical properties” of this instance contain a lexical element that leads to the specification of the code as “colloquial attic”. Under “orthography” are documented some deviations of the orthographic rules of the classic attic, that inform us about the literacy of the anonymous author. The grapheme for the aspirated stop is used instead of the grapheme for the non-aspirated and the graphemes for the semi-open and the semi-closed front vowel are used in a non-consistent way. The latter orthographic confusion serves also as evidence for the evolution of the phonological system

of colloquial attic. The opposition among these vowels is already lost at the innovative varieties of the classic era.

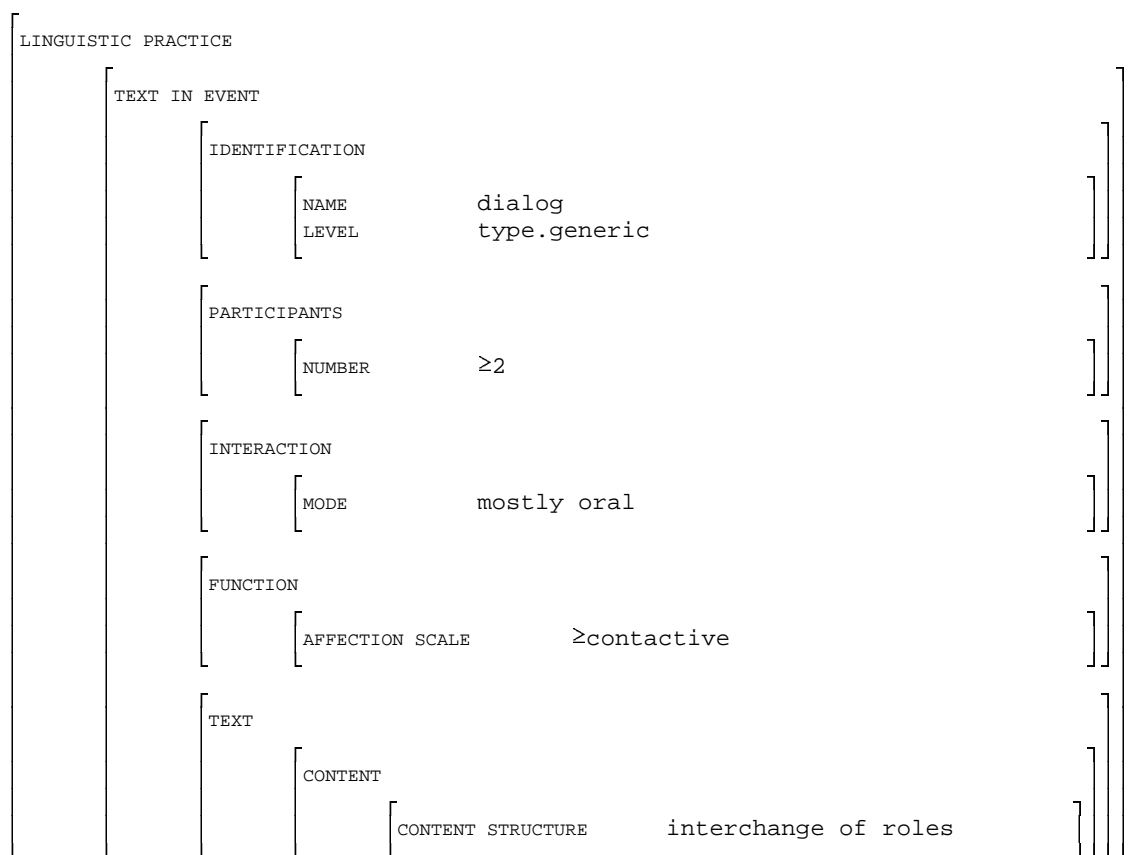
Figure 5: Documentation of a simple instance: An attic curse

LINGUISTIC PRACTICE	
TEXT IN EVENT	
IDENTIFICATION	
NAME	Ker III Att C 9.1
LEVEL	instance
PARTICIPANTS	
NUMBER	2
PRODUCER	anonymous
ADDRESSEE	Hermes, Persephone
RELATION	personal; religious
INTERACTION	
MODE	written
MEDIUM	scratch on pottery
OPENNESS	private
PLACE	Kerameikos
TIME	c. 350 BC
PHYS_SET	(meta-) & physical distance of the speech participants; discontinuous
FIELD	religion
FUNCTION	
AFFECTION SCALE	instructive
SPECIAL FUNCTION	magic
TEXT	
CONTENT	
TOPIC	curse about several persons
CONTENT STRUCTURE	sequence of separate curses for each person
FORM	
FORMAL STRUCTURE	simple sentences
CODE	colloquial attic
LEXICAL PROPERTIES	<i>katadō</i> 'damn:1.sg' (= colloquial)
GRAMM. PROPERTIES	pattern 'I damn + body part + person + before God'
ORTH. PROPERTIES	faults: <p ^h > instead of <p>, <ē> instead of <ē̄>

4.2. Generic types

The generic types are abstract schemes of properties that can be applied across cultures. Such types are for example the tale, the drama, the dialog, the biography, the song, the game instruction, the stone inscription, the personal letter, etc. (cf. Lehmann s. d.). Figure 5 exemplifies the encoding scheme of a generic type, namely the dialog. The generic type of dialog is an underspecified scheme. It presupposes the existence of two or more speech participants and an interchange of the roles of producer and addressee between them. It is mostly an oral discourse type, but it not necessarily so.

Figure 6: Scheme of a generic type



The notion of generic types can also be applied to the domain of community. Generic types of persons that can be applied across cultures are for example types like the “intellectual”, the “father”, the “old man” etc.

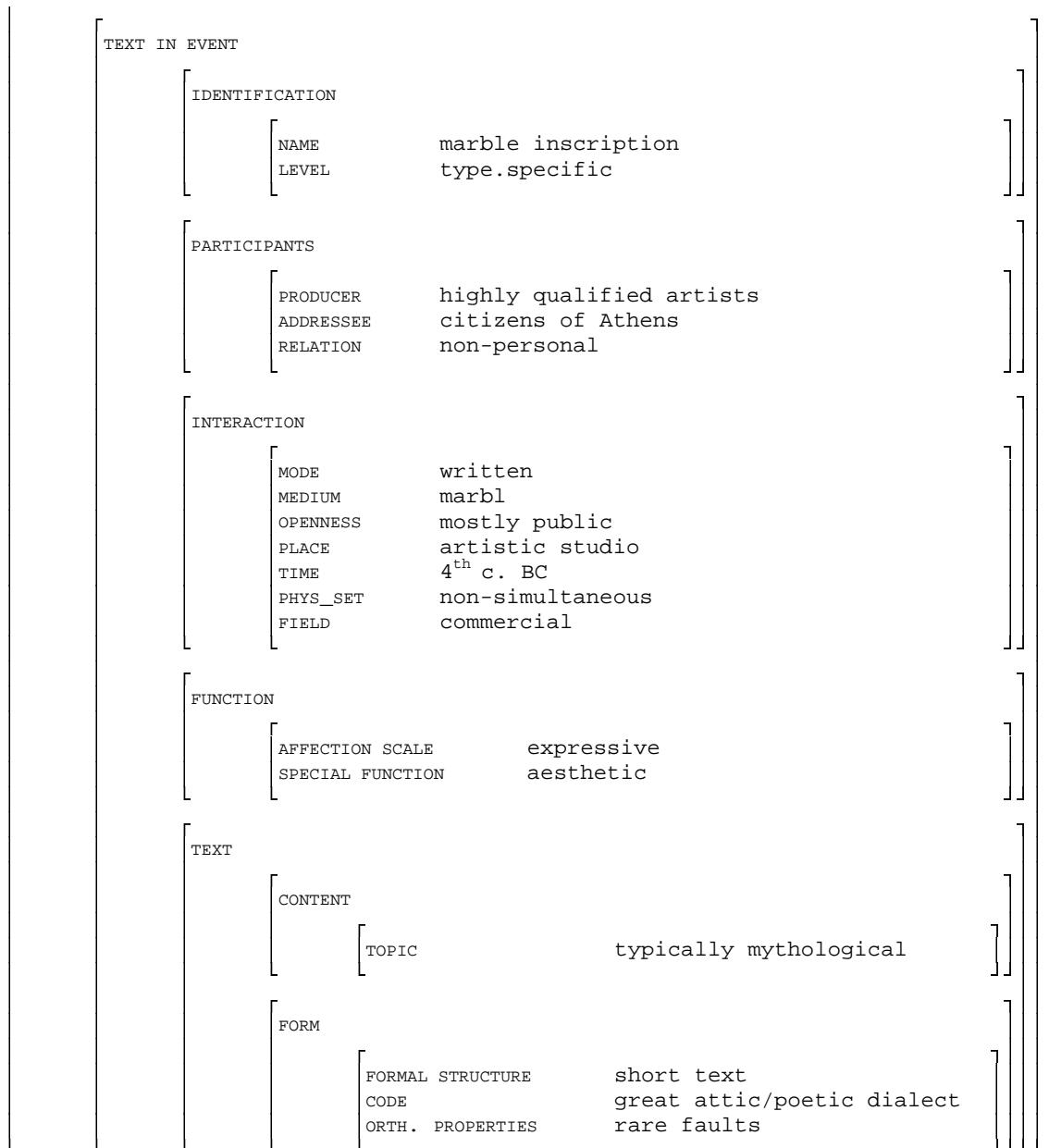
4.3. Specific types

The subject of language documentation is the “specific type”, namely the realization of an abstract cross-linguistic type in a particular linguistic community. In Figure 7 two entries in the domain of language practice illustrate the difference between two specific types of

classical attic. The first entry concerns the specific type “graffiti on pottery” and the second entry the type “marble inscription”.

Figure 7: Specific types: graffiti on pottery vs. inscription on marble in classical Attic

LINGUISTIC PRACTICE	
TEXT IN EVENT	
IDENTIFICATION	
NAME	graffiti on pottery
LEVEL	type.specific
PARTICIPANTS	
PRODUCER	non educated masters, mostly slaves
ADDRESSEE	citizens of Athens
RELATION	non-personal
INTERACTION	
MODE	written
MEDIUM	pottery; graffiti
OPENNESS	private
PLACE	workshops with ca. 40-50 workers
TIME	4 th c. BC
PHYS_SET	non-simultaneous
FIELD	commercial
FUNCTION	
AFFECTION SCALE	expressive
SPECIAL FUNCTION	aesthetic
TEXT	
CONTENT	
TOPIC	typically mythological
FORM	
FORMAL STRUCTURE	short text
CODE	vulgar attic
ORTH. PROPERTIES	common faults



There are some common properties between both text types. Both are produced for commercial purposes, both are instances of written communication, the Athenian citizens are in both cases the addressees, the function is expressive-aesthetic, the topic of the text is in most cases a mythological motif, and the formal structure is in both cases typically a short text.

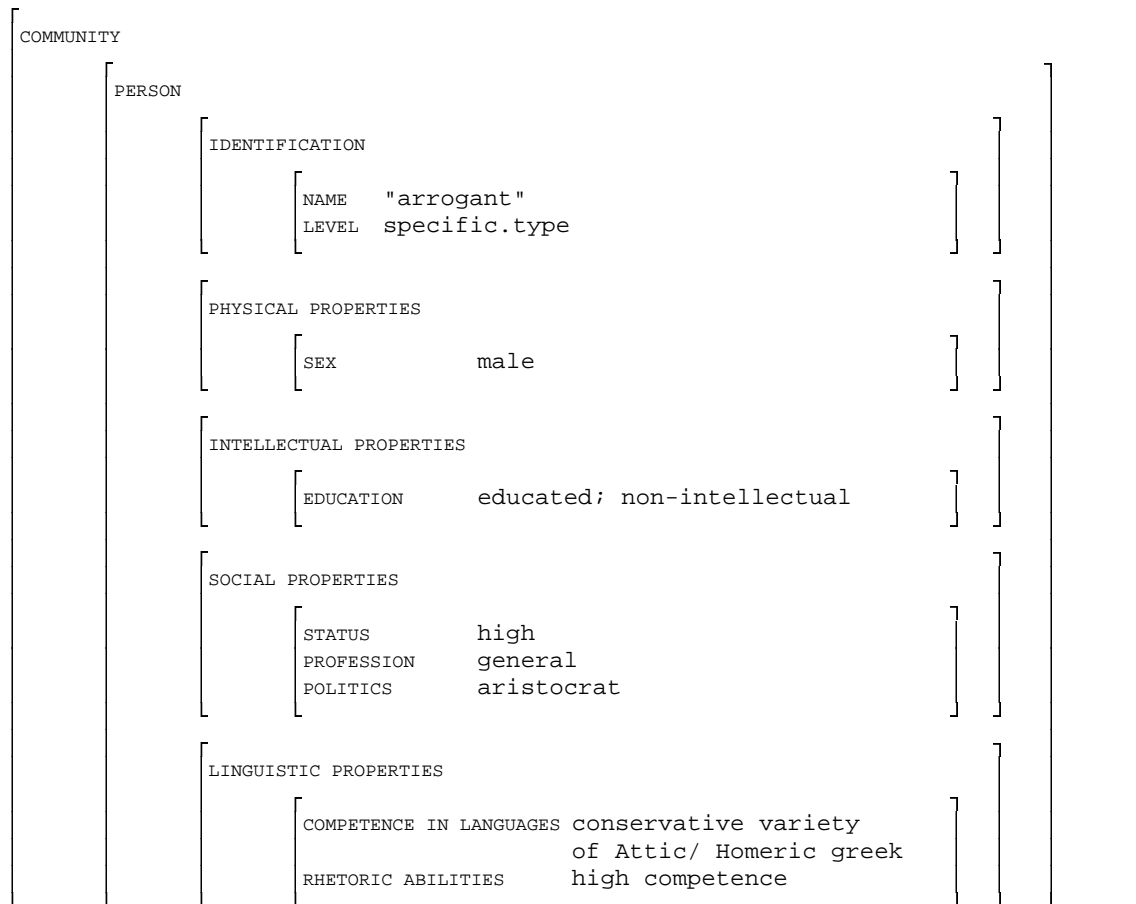
The texts of both types are usually written in different codes. Whereas marble inscriptions are instances of documentation of the great attic or of the poetic dialects used by the Athenian intellectuals of the classical period, the graffiti on pottery is the best evidence for the vulgar attic of this era (s. Woodhead ²1981). At the level of the communicative event, within which both types are produced, there are some crucial differences that condition the variation at the level of the expression.

The material used is quite different. Marble is valuable and pottery is cheap. Marble inscriptions are produced by highly qualified artists, that work in personal studios, whereas pottery is produced in classical Athens in big workshops by non-educated masters, in the majority slaves.

Beside the relevance of this information for the conditioning of linguistic variation this example illustrates the issue about the specific discourse types. They are patterns of combination of features at different levels, from the level of the expression to the level of the participants of the interaction, the used material, etc. that emerge in context of a particular culture. The relation between an instance and a specific type is a **prototypical relation**. Instances can occur in different degrees of deviation with respect to the prototypical properties of the specific type. Since specific discourse types are patterns built in a particular culture and not simple descriptive abstractions they must be considered as conventional linguistic units like the units of morphology and syntax. Furthermore, they can be described in terms of different degrees of **conventionalization**, that emerges partially parallel to their evolution. For example funerary inscriptions of the early period are less complex, as they include verse epitaphs of one or two lines and have less conventionalized lexical properties than the inscriptions of the same type in the classical period. Most funerary inscriptions of the last period are written following less than ten different motifs like ‘here lies A’, or ‘this is tomb of A’ , ‘B set the tomb of A’, or ‘I am the tomb of A’ etc.

Specific types in the domain of persons are the different **characters**. Figure 8 exemplifies one of the four characters that are instantiated through different heroes in the comedies of Aristophanes. The character is the “arrogant” and has the following properties: he is always a man, he is an educated person but not an intellectual, and he has a high social status and aristocratic beliefs. In the comedy he uses instructive texts, his addressee is the “average citizen” – another specific type – and the code he uses is either the conservative variety of attic or Homeric Greek.

Figure 8: Specific types in community: Characters

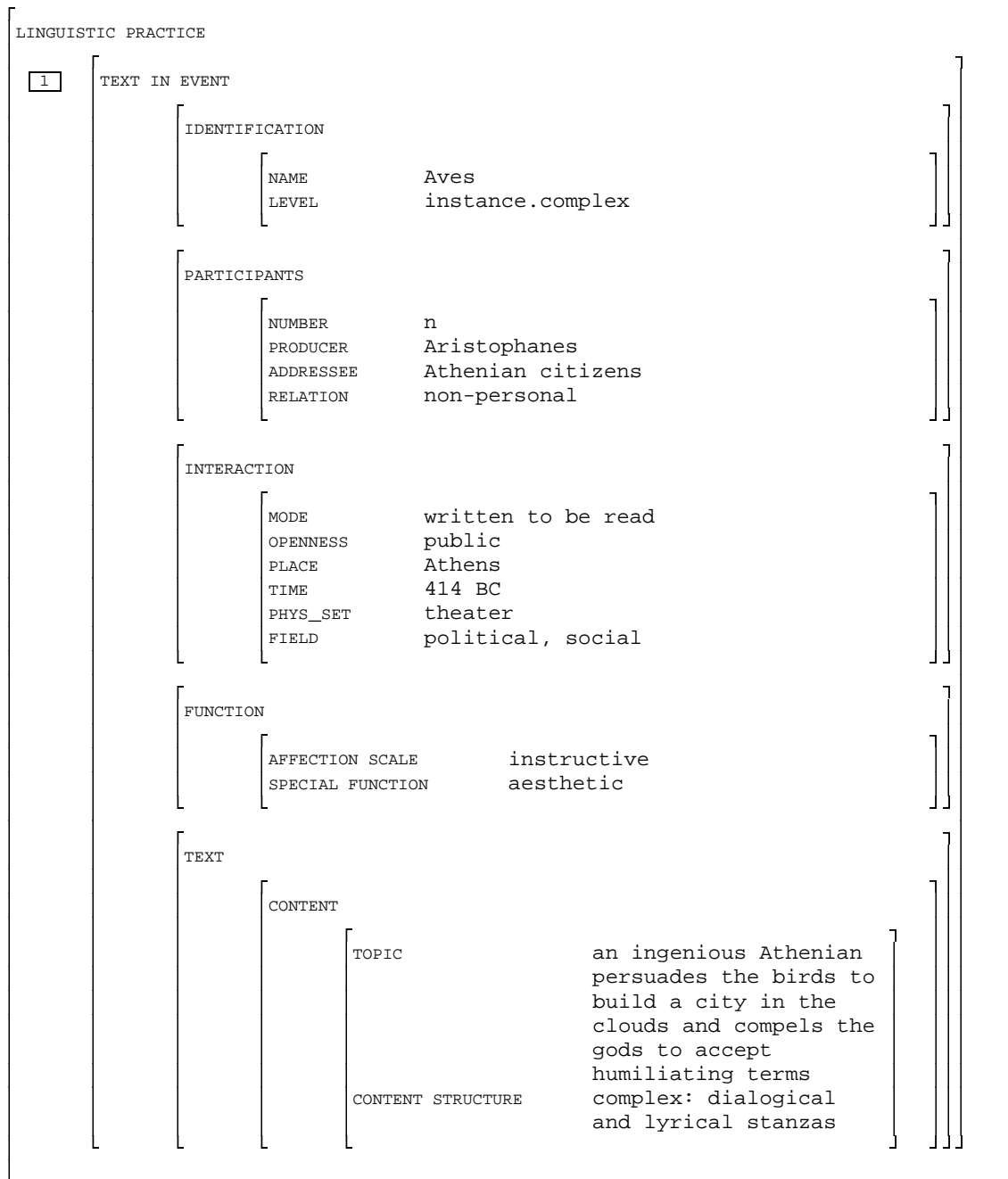


5. Complex units

The units we examined so far are all simple units. However, an encoding scheme concerning texts needs also a technique for the description of more complex units, i.e. texts that include other texts. Figure 9 illustrates this case with a comedy. A comedy is a very complex text, since it contains textual divisions, that could stand also as individual texts: lyric stanzas with different functions (like introducing the piece or breaking the plot etc.), dialogical parts in many scenes, citations from other poets etc.

In Figure 9 a comedy and a part of the same comedy are represented as different instances. The comedy is the *Birds* of Aristophanes and the part is a dialog between the main hero of the comedy Peistheteros and an attic poet Kinesias. The relation among them is notated in the field "relations to other texts" in the part of the comedy, namely the dialog.

Figure 9: Documentation of a complex instance: An attic comedy



TEXT IN EVENT	
IDENTIFICATION	
NAME	[1] 1372-1409
LEVEL	instance.simple
PARTICIPANTS	
NUMBER	2
PARTNER	[2] Peistheteros
PARTNER	[3] Kinesias
RELATION	non-familiar; personal
INTERACTION	
MODE	oral
OPENNESS	private
PLACE	[4] in a city on the clouds
TIME	cf. [1]
PHYS_SET	[4]
FIELD	cf. [1]
FUNCTION	
AFFECTION SCALE	instructive
TEXT	
CONTENT	
TOPIC	[3] wishes to be accepted in the in [4] and to take wings. [2] send him away
CONTENT STRUCTURE	complex
FORM	
FORMAL STRUCTURE	verses; simple sentences
CODE	[2] colloquial attic [3] attic of lyric
RELATIONS TO OTHER TEXTS	
MERONOMIC	part of [1]

The relation between a complex instance and its parts is a relation of entailment. The dialog is part of the whole “comedy”. Analogous distinctions appear also for other domains. The corresponding distinction in community is the distinction between individual and groups. Groups like a **family**, a **parish**, a **village**, a **school class**, a **working team** and a **political party** are populations of individuals that share common knowledge and interact linguistically with each other. Similarly in the domain of the knowledge a simple unit includes an object

like a **throne** and a complex unit a collective where this object belongs to, like a **theater** (as concerns thrones in the particular culture).

6. Relations

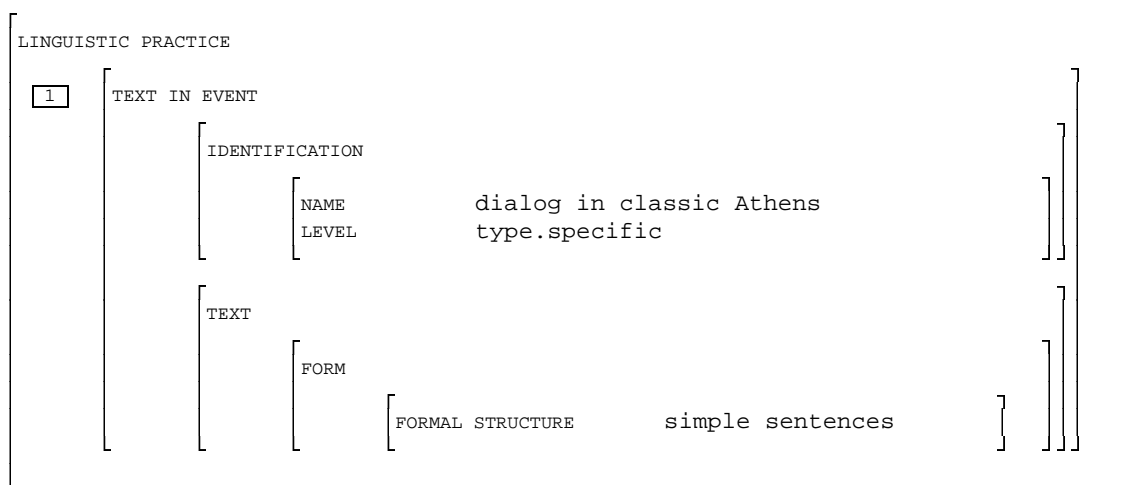
The last section examines the **relations** among different units. Figure 9 illustrates only some of the relations in the presented text. The unit of interest is the text in event nr. 4, that is a part of the dialog of Figure 9, in particular the passages of Peistheteros, who represents the “average citizen”. The form of these passages has the following properties: written in verses, they consist only of short sentences, and the code used is the colloquial attic.

The property of the “verses” is inherited from the specifications of the including complex text: comedies are written in verses. The relation is indicated with a cross-reference to the including text, the comedy *Aves* (unit nr. 3), which inherits its formal properties from its type, namely the attic comedy as a specific type (unit nr. 2).

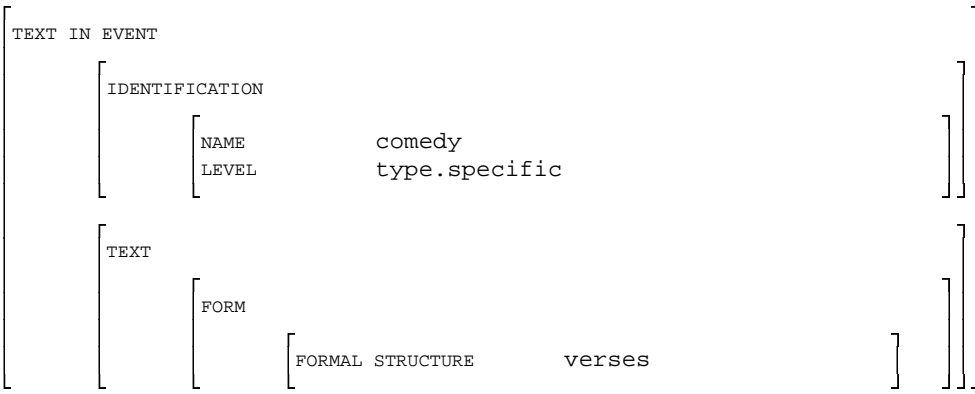
The second property of the passages of Peistheteros concerns the “short sentences”. This feature is inherited from the discourse type of this passage, the attic dialog, an instantiation of the universal dialog.

Finally, the last property of these passages is the “colloquial code”. This feature is inherited from the character of the hero. This property is inherited from the producer of the text, and furthermore from his culture-specific type, the Athenian average citizen and his linguistic habits.

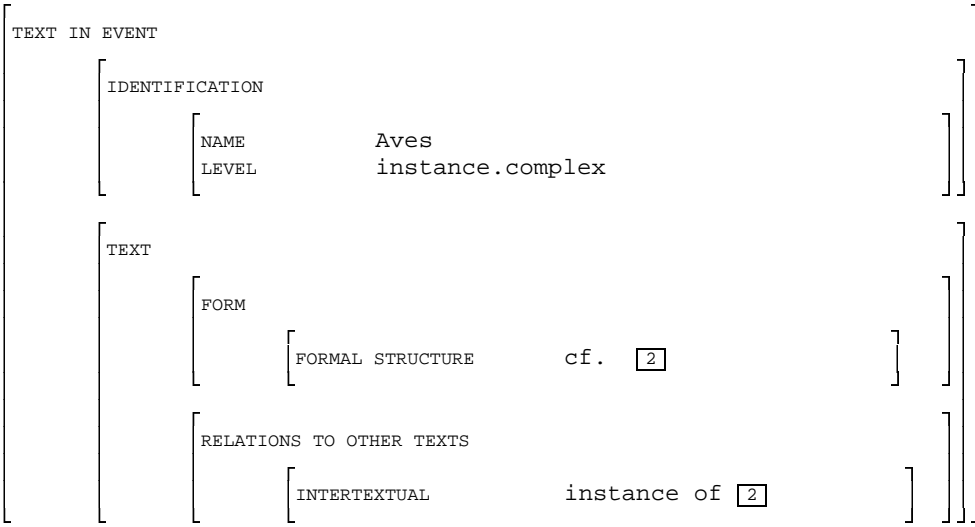
Figure 10: Relations



2



3



4	TEXT IN EVENT		
	IDENTIFICATION		
	NAME	3	verse 1375
	LEVEL		instance.simple
	PARTICIPANTS		
	NUMBER		2
	PRODUCER	5	
	ADDRESSEE		Kinesias
	RELATION		non-familiar; personal
	INTERACTION		
	MODE		oral
	OPENNESS		private
	PLACE	6	city on the clouds
	TIME		cf. 3
	PHYS_SET	6	
	FIELD		political, social
	FUNCTION		
	AFFECTION SCALE		instructive
	FORM		
	FORMAL STRUCTURE		cf. 3 (verse)
			cf. 1 (simple sent)
	CODE		cf. 5 colloquial
	RELATIONS TO OTHER TEXTS		
	MERONOMIC		part of 3
	INTERTEXTUAL		instance of 1
	TEXT		
	FORM		
	FORMAL STRUCTURE		cf. 3 verse
			cf. 1 simple sent .
	RELATIONS TO OTHER TEXTS		
	MERONOMIC		part of 3
	INTERTEXTUAL		instance of 1



7. Conclusion

Generic types are neither the subject of language documentation nor of language description, since they are not culture-specific entities. The relevance of generic types, like dialog or narrative, is that they provide a basis to create a number of templates for the documentation of texts. This is the level of **abstract universal structures**. Simple and complex instances are to be based on these generic templates. This is the stage of **language documentation**. The possibility to draw correlations between specific text types and their communicative foundations is the stage of **language description**.

8. References

- Heinemann, Wolfgang & Vierweger, Dieter 1991, *Textlinguistik: Eine Einführung*. Tübingen: Max Niemeyer
- Hymes, Dell 1974, *Foundations in sociolinguistics: an ethnographic approach*. Philadelphia: Univ. of Pennsylvania Pr.
- Lehmann, Christian, *Structure of a language description*. Manuscript.

- Lehmann, Christian 1999, *Documentation of endangered languages: a priority task for linguists*. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSIDUE 1).
- Lehmann, Christian 2001, "Language documentation: a program". In: Bisang, Walter (ed.), *Aspects of Typology and Universals*. Berlin: Akademie Verlag, 83-97.
- Lenk, Elena 1996, *Erstellung eines Textkorpus für eine Sprachdokumentation mit LDS*. Univ. Bielefeld: Bielefeld (Allgemein-Vergleichende Grammatik, Arbeitspapier Nr.13)
- Woodhead, A. Geoffrey ²1981, *The study of Greek inscriptions*. Cambridge: Cambridge Univ. Pr.