On the linguistic categorization of sounds

Christian Lehmann

University of Erfurt

22.07.2003

Prof. Dr. Christian Lehmann
Seminar für Sprachwissenschaft
Philosophische Fakultät
Universität
D – 99105 Erfurt

E-Mail: Christian.Lehmann@Uni-Erfurt.De
Fax: +49/361/737-4209

Abstract

1. Human beings categorize sounds and noises, in the first place, at the level of perception and cognition. Processes and operations in this domain are the object of such disciplines as acoustics and perceptual psychology. In language, sounds are categorized at various levels. Most prominently, their categorization for the sake of forming significantia of language signs constitutes the object of phonetics and phonology. However, sounds are also categorized in the form of words denoting them. Such words, in turn, constitute structured subsystems at the morphological, syntactic and lexical-semantic levels.

The categorization of sounds at the lexical-semantic level can be researched by a variety of methods, including experiments of psychoacoustics and psycholinguistics, research of corpora into the textual use of sound words and methods of system-linguistic lexical, morphological and submorphemic analysis. In order to bridge the distance from the categorization in the single *langue* to the level of human cognition, comparative-typological research helps to systematize the variation among languages and to establish common principles.

2. Comparative research concentrating on sound verbs (such as *rattle*) and particles (such as *bang*) shows that, at the lexical-semantic level, sounds are categorized by diverse parameters, only some of which are auditory in nature. The auditory parameters, in turn, concern the vertical (simultaneous) and the horizontal (sequential) structure of the sound. The vertical structure comprises such parameters as intensity (which constitutes one of the differences between *roar* and *rustle*), timbre (cf. *click* vs. *clack*) and periodicity (*ring* vs. *hiss*), which are well-known from acoustics. The horizontal structure comprises, first of all, the rate of impulses. This ranges from instantaneous (*bang*) over iterative (*cluck*), crepitative (*rattle*) and vibrating (*buzz*) to continuous (*hiss*). Another parameter of horizontal sound structure is homogeneity: while both *roar* and *hiss* are continuous, the former is heterogeneous, the latter homogeneous. Finally, the close of the single impulse may be checked (*pop*) or fading (*bang*).

Besides these properly auditory parameters, sounds are also categorized by esthetic evaluation, as either pleasant (*hum*) or unpleasant (*creak*). Finally, properties of the sound source play a role in constituting selection restrictions of sound words. Some examples are *chirp* (lower animal), *gnash* (hard object), *click* (technical device). In the conscious categorization of sounds, this latter parameter tends to prevail over the purely auditory categorization. The latter appears to play a largely ancillary role at the level of subconsciousness.

Sounds are intimately bound up with movements. Many lexical items like *knock* designate both a sound and the movement or manipulation that provokes it.

3. When subjects/informants are asked to characterize a sound concept, they prefer to take recourse to selection restrictions and to the movement/manipulation that produces the sound, rather than to a purely auditory characterization. That is, they will characterize *bang* as ‘the sound typically produced by vigorously clapping a door’ and not as ‘a loud, instantaneous, fading noise’. This suggests that human beings do not categorize sounds like they categorize visual percepts, as a goal in itself. Instead, auditory categorization is typically functional in figuring out what the relevance of the sound is for the recipient.

This suggests that human beings are equipped with a perceptual and cognitive apparatus that can analyze a sound minutely and effectively at the subconscious level, in order to interpret it with regard to its possible consequences for the hearer. That is, they are genetically well prepared to use the sound medium for semiotic purposes.

4. This research needs to be extended in various directions:

- To what extent are the parameters that characterize sounds peculiar to this domain, and to what extent do they structure other semantic domains? For instance, are the values of the parameter of the horizontal structure of sounds equal to what otherwise appears as verbal characters and *aktionsarten*?
- An independent characterization of the semantics of sound words is a methodological prerequisite for a responsible analysis of onomatopoetic words, which must now follow.
- The relationship of the parameters of auditory perception to categories of perception by the other senses needs to be clarified in order to develop a theory of ideophony.
- An important theoretical question is what the relation is between those parameters like the above, which categorize sounds at the lexical-semantic level, and those parameters which have been known as phonetic and phonological features. To what extent do they coincide; are they of the same nature? Can we say that at the most general level, human audition proceeds by a set of auditory parameters some of which are sharpened and used for phonetic purposes?

1 Introduction

What follows is a project for the investigation of the linguistic categorization of sounds.[1] Sounds in themselves are primarily the object of physics, more in particular, of acoustics. However, not all of the physical qualities of sounds are perceived by human beings; and on the other hand, human beings categorize sounds by parameters and distinctions some of which are not relevant to physics.

Men categorize sounds by criteria some of which are auditory in nature, such as their loudness. Such perceptual criteria are based on acoustic properties of sounds. Apart from these, men employ criteria for the categorization of sounds which have nothing to do with their physical qualities. For instance, the sound made by a clock is designated by *tick-tack*. The expression implies that the sound has a binary rhythmical structure and consists of two different impulses. This apperception, however, relies on the movement which produces the sound, since the physical reality itself is a monotonous sequence of identical impulses.

It may well turn out that auditory categories do, in fact, play a subordinate role in the categorization of sounds and that those properties of sounds which are of most interest to human beings and underlie their linguistic categorization are not at all their auditory properties. While this will be born in mind, we will here primarily be interested in the purely auditory categories of sounds and their linguistic correlates.

My own research has so far been preliminary; no rigorous methods have been applied. The languages investigated include German, Spanish and Yucatec Maya, with occasional glimpses at English and Japanese.

---

[1] Thanks are due to the members of various auditories who commented on previous versions of the present paper, esp. Carmen Pensado, Julián Méndez, Johannes Helmbrecht, Yong-Min Shin, the auditorium of 18.01.02 at University of Pavia.

2 Levels of linguistic categorization

A word designating a sound will be called a **sound word**. Thus, *buzz* is a sound word, but *grind* is not. Although both refer to situations which involve a sound, the latter does not designate the sound produced in the situation, or the production of this sound. The sound is rather a by-product of the designated situation. The distinction is not always easy to make, an issue that we will come back to.

2.1 Semantic categorization

Linguistic categorization of sounds at the purely semantic level, i.e. without reference to expression structure, is their designation by lexemes. Such lexemes are arranged in lexical fields such as 'sounds' vs. 'movements', 'human sounds' vs. 'animal sounds' etc. The meanings of the lexemes are described by lexical-semantic features such as 'loud vs. soft', 'dark vs. bright' etc.; and these are, at the same time, the parameters that structure the lexical field. These features or parameters are, in general, based in cognition; and in the particular case of sound concepts, they are, to a large extent, based in perception.

The structuralist method for the analysis of a lexical field and the assignment of semantic features requires the identification of minimal contrasts. These will be found in some cases. Overall, however, the field will be found to be structured rather irregularly. There are probably far more gaps in it than positions occupied.

2.2 Formal categorization

2.2.1 Morphological-syntactic categorization

2.2.1.1 Syntactic categorization

The most important linguistic categories at the syntactic level of categorization are **word classes**. As we may see from a set of sound words such as *noise, loud, rattle* and *bang*, sound words may, in principle, belong to any of the major classes noun, adjective, verb and particle. The same may be observed in other languages, although no claim is being made that this will be so in all languages.

A superficial collection of sound words in English and neighboring languages, but also in Japanese and Yucatec Maya, immediately reveals that sound words are distributed very unevenly over the word classes. The languages alluded to have dozens or even hundreds of sound words in the categories verb and particle, but only a few sound adjectives and very few basic sound nouns indeed. If these facts are systematized in an adequate way, they will probably follow from universal principles. Word classes may be arranged on a continuum of **time stability** or dynamicity (cf. Lehmann 1991), such that the categorial meanings of nouns, adjectives and verbs are increasingly dynamic ('time-labile'). Particles as a class do not have a position on this continuum, although they may be subclassified according to it. The hypothesis being proposed is that the more dynamic a word class is, the more sound words it contains. This would have its natural explanation in the nature of the meanings involved. We will come back to this in §6.

2.2.1.2 Morphological categorization

Linguistic categorization of sounds at the morphological level is essentially the derivational structure of sound words. Some English sound words, for instance, display ablaut reduplication like *chit-chat, tick-tack, ping-pong, snip-snap* etc. Few derivational processes appear to be peculiar to sound words, but some of the semantic features that are relevant to sound words are identical to features that constitute more general derivational categories, so that the lexical field of sound words is articulated, to some extent, by the same parameters that structure the whole lexicon.

2.2.2 Phonological categorization

Categorization of sounds at the phonological level is essentially done by onomatopeia. The word *bang*, e.g., designates a loud explosion that gradually fades away. The word starts with a plosive, contains an open vowel and ends in a nasal, all of which represent the semantic features mentioned.

3 Research methods

In any scientific research, it is advisable to apply several independent methods that are capable of controlling each other. In the present instance, this is even obligatory, since both sides of the linguistic sign, with their relation to linguistic substance, and all of the levels of the structure of linguistic signs are involved.

When prompted to characterize the sound which is the significatum of a sound word, we are liable to characterize the sound of its significans instead. Thus, if asked to tell the meaning difference between *crack* and *crackle*, we are prone to answer: *crack* is a single impulse, *crackle* is a sequence of impulses, relying in this on the morphotactic shape of the two words. This is the **onomatopoetic circle**, which in itself is theoretically interesting,[4] but which must be avoided in the methodology both of the descriptive linguistic analyses and of the psycholinguistic experiments to be conducted. The dangers of the onomatopoetic circle will be controlled by the multiplicity of methods to be applied.

3.1 Methods of structural linguistics

3.1.1 Lexical field analysis

3.1.2 Analysis of usage

Mannheimer Corpus

[4] It is a decisive argument for structural linguistics to observe that even speakers (not only linguists) form their significatum on the basis of the structure of the significans.

The typical context for a sound verb is the designation of its sound or energy source. To the extent that the use of the verb in question is controlled by selection restrictions on its subject, it is not used to give information on the auditory aspects of the situation, but to predicate on the subject that it makes its default sound.

This is also the basis of metaphors in this domain. When we say of an officer that he grumbled, we are not so much specifying the sound that he made as we are likening him to some animal as a dog.

3.2 Experimental methods

3.2.1 Folk classification

For the similarity groupings in the following two experiments, items are presented in triples, and subjects judge on their relative similarity.

3.2.1.1 Similarity of words

Have subjects group sound words according to semantic similarity.

3.2.1.2 Similarity of sounds

Select a set of sounds that have been determined in previous experiments as focal instances of certain sound words. Have subjects (of different native languages) group the sounds according to similarity.

3.2.1.3 Use of words

Present a sound word to the subject and ask him to associate objects that would be sources producing the sound in question.

3.2.2 Linguistic categorization

3.2.2.1 Apperception of natural noises

Question: What lexical categories are used to classify naturally occurring noises (onomasiological approach)?
Experiment: Subjects hear a naturally produced noise.

- They name it spontaneously.
- They assign it a term from a list shown to them.
- They imitate the sound vocally.
- They guess at the source of the noise they heard.

Their products are analyzed by linguistic methods. In particular, the paralinguistic sounds produced are compared with those presented as stimuli, in search of common features.

3.2.2.2 Guided production of noises

Question: What is the specific (acoustic part of the) meaning of sound-designating words (semasiological approach)?

Experiment: After sufficient introduction to the sound generator, and with technical assistence,

- Subjects receive a linguistic stimulus for a concept (e.g. *it cracks*).
- They compose a noise which they regard as a focal instance for the concept.

Their products are analyzed by acoustic methods. They are compared with each other, and common acoustic features are destilled.

This approach also pays attention to the culture dependence of the meanings in question.

3.2.2.3 Apperception of artificial noises

Question: To what extent are acoustic meanings interpersonal (mutual control of onomasiological and semasiological approaches)?

Experiment: Subjects hear a noise of those produced in experiment 3.2.2.2.

- Same steps as in experiment 3.2.2.1.

The correlation of the names assigned by subjects in experiment 3.2.2.3 with those presented in experiment 3.2.2.2 is analyzed.

4 Parameters of categorization of sounds

In this section, sounds will be classified according to a set of parameters. The approach will be primarily linguistic. This means that the distinctions made will be justified on the basis of linguistic criteria. These concern essentially formal and semantic properties of sound words, viz. their own structure and their distribution, i.e. the structure of their linguistic environment.

4.1 Auditory properties

Sounds are events with an extension in time. The structure observable in their progression will be called their sequential, temporal or **horizontal structure**. For instance, *howl* involves an alternation of growing and decreasing, which is not involved in *hiss*. The structure that a sound has at a given point in time will be called its simultaneous or **vertical structure**. For instance, *click* and *clack* differ in their vertical structure as the sound designated by the former is brighter.

4.1.1 Vertical structure

The parameters defining the vertical structure of a sound are well-known from acoustics and phonetics. They are enumerated in T1.

T1. *Vertical structure of a sound*

paramete	value	explanation	examples

r			**English**	**German**
intensity	high	wide amplitude	*bang, roar*	*knallen, brüllen*
	low	narrow amplitude	*rustle, crackle*	*gluckern, rascheln, knistern*
timbre	bright	high pitch or high overtones	*click, clink*	*klicken, sirren*
	dark	low pitch or low overtones	*clack, clank*	*klacken, surren*
quality	musical	periodic vibration	*ring, whistle*	*klingen, pfeifen*
	noise	aperiodic vibration	*hiss*	*zischen*

German has a morphological contrast that expresses **intensity** in sound verbs. It makes use of the attenuative verb derivation in *-el* + metaphony, which appears outside sound verbs in such pairs as *tanzen - tänzeln* 'dance - dance lightly'. T2 shows some sound verbs with attenuative derivation expressing an intensity contrast:

T2 *Attenuative derivation expressing low intensity in German sound verbs*

unmarked		**low**	
German	**meaning**	**German**	**meaning**
sausen	sough	*säuseln*	swish
zischen	hiss	*zischeln*	swish
rumsen	thud	*rummeln*	rumble
trappen		*trappeln*	patter
brummen	drone, buzz	*brummeln*	

There are also some pairs such as *klingen – klingeln* which display the same formal relationship without contrasting semantically in the same way. And there are even more verbs such as *dudeln* exhibiting an *-el* formative without being opposed to anything else.

Plätschern as opposed to *platschen* apparently shows dissimilation of the attenuative suffix or some interference with iterative derivation (below).

It is not clear that the acoustic distinction between **pitch** (base frequency) and **timbre** (spectral energy distribution) is relevant, apart from the meaning of the adjective in *high/low tone*. The irrelevance of the distinction obviously follows from the absence in nature of pure musical tones, so that generally pitch and overtones are indistinguishable.

Linguistic correlates of the distinction between **musical sound** and **noise** pertain primarily to the phonological level (thus, to onomotopoeia): words designating musical sounds contain, at an average, more sonorants; words designating noises contain more obstruents.

4.1.2 Horizontal structure

While the parameters of the analysis of the vertical structure of sounds are well-established in science, the analysis of their horizontal structure is still in leadings strings. Three parameters will be used here: rate of impulse, homogeneity and close of impulse.

4.1.2.1 Rate of impulse

4.1.2.1.1 Structure of the parameter

Acoustics only makes a binary distinction between impulse and continuous sound. However, at the linguistic semantic level, this parameter is more finely articulated, as displayed in T3.

T3 *Temporal structure of a sound: rate of impulse*

hertz	**parameter value**	**explanation**	**examples**	
			English	**German**
	instantaneous	single impulse	*crack, bang, clang*	*krachen, knallen,*
	repeated	single impulse, repeated by default	*bark, knock, ring, peep*	*bellen, klopfen, läuten, piepen*
2	iterative	continuous: iteration of distinct instances	*cluck, clack*	*glucken, klacken*
3-6	crepitative	continuous: high iteration frequency	*rattle, jingle, clatter, crackle, tinkle*	*rasseln, klingeln, klirren, knistern, scheppern*
7-16	vibrating	continuous: vibrating	*buzz, drone, creak, whir, purr*	*summen, dröhnen, knarren, schwirren, schnurren*
> 16	continuous	continuous: simplex	*hiss, whistle*	*zischen, pfeifen*

T3 is a continuum with gradual transitions. From top to bottom, we can distinguish six points on this continuum. They are named by the terms in the second column, and approximate repetition rates are given in hertz in the first column. Since we are here dealing with the frequency of an impulse, a remark on acoustics is in order. Every sound whatsoever presupposes the vibration of a physical body (including the air). If the frequency of this vibration is, roughly, between 16 and 16.000 hertz, it is audible as a sound for the human ear. If the frequency is below 16 hertz, then each single impulse is discernible. We will return to this point below. The six stages my be characterized as follows:

Instantaneous: Vibration (above 16 hertz) initiates abruptly, reaches its maximum amplitude in fractions of a second and decreases more slowly.

Iterative: An iterative sound verb is one which designates an iteration of distinct instances. The value 'iterative' differs from the value 'repeated' in that the iteration is constitutive of the concept. German has an iterative derivation, as shown by the pairs in T4.

T4 *Iterative derivation in German sound verbs*

unmarked		**iterative**	
German	**meaning**	**German**	**meaning**
glucksen		*gluckern*	bubble, guggle
bumsen	thud	*bummern*	
platschen	splash	*plätschern*	plash, dabble, ripple
klappen	flap	*klappern*	clatter, chatter
trappen		*trappeln*	
klingen	ring	*klingeln*	ring

Trappeln has a dissimilated allomorph of the suffix. There is probably some interference with the attenuative suffix as evinced by such cases as *klingen : klingeln*; attenuative and iterative are akin in many languages.

There are also iterative verbs ending in *-er* which lack a non-derived counterpart: *rattern, knattern, poltern, trillern*. However, apparently every sound verb with a bysillabic stem ending in *-er* at least has a durative, if not an iterative meaning.[2]

And again, there are iterative verbs with an *-el* ending, such as *rascheln, rumpeln, prasseln, rasseln*, all of which have an /r/ in the root.

The iterative character of the verb may be tested by inserting it in the frame '___ einmal/once.', as in E1.

E1. a. Es klopfte einmal. 'It knocked once.'
GERM b. Es plätscherte einmal. 'It dabbled once.'
c. Es ratterte einmal. 'It rattled once.'

For non-iterative verbs, such as the verb in E1.a, the insertion in the test frame produces unremarkable results with a compositional meaning. For derived iterative verbs, such as the verb in E1.b, the test frame is semantically deviant. For underived iterative verbs, such as the one in E1.c, the test frame is unusual, but may designate not a single impulse as part of a repeated sound, but a single (instance/occurrence of a) sound which involves internal iteration. This result is typical for a semantic character vs. an aktionsart.

Vibrating (T3, l. 5): As remarked above, any sound whatsoever necessarily involves vibration above 16 hertz. This is not the sense in which *vibrate* is being used in T3. The frequency of the individual impulses in lines 2 through 5 of T3 is below this threshold. For vibrating sounds,

[2] A possible exception is *(sich) räuspern* 'hawk', if one wants to consider it as a sound verb (cf. §4.3).

however, it is so close to the threshold that there is only an indistinct perception of vibration, but no possibility of identifying an individual impulse.

Vibrative words contain more /r/s than usual.

4.1.2.1.2 Role of the parameter

As noted above, if the iteration rate of a repeated impulse exceeds 16 hertz, it is perceived as a simplex sound. Apart from its duration, this has the same internal structure as an instantaneous sound. Insofar, the last value of T3 precedes the first one, and the parameter of T3 has a circular structure. This has also been shown by phonetic experiments: If the (continuous) fricative in a German word such as *Bache* is cut to the length of a stop (appr. 40 ms), subjects hear *Backe*.

Some of the parameter values used, esp. 'iterative' and 'semelfactive', are well-known from the theory of **aktionsarten**. More generally, sound events are, in principle, categorized along the same dimensions as situations are in general. That is, the continuum of time-stability (dynamicity) as set out in Lehmann 1991 applies to them, so that sound words are subsumed under certain predicate classes and their subclasses, and the same kinds of formal processes appear on them.

The concept of iteration rate in sound words is similar to the concept of **granularity** as applied to substances in cognitive semantics (cf.).[3] Here, too, one starts out with aggregations of like objects; sheep in a herd are like repeated sounds. Next come A substance like corn is like a crepitative sound in that the individual item is, in principle, isolable, but is practically never isolated. A mass like salt or sand is like a vibrating sound in that the naked eye barely discerns the grain, but still has the global impression of granularity. Butter, finally, is like a continuous sound in that no constitutive components can be separated.

4.1.2.2 Homogeneity

The second parameter in the horizontal structure is relevant for sounds whose impulse rate is crepitative/vibrating or greater. Sounds with a lower impulse rate are heterogeneous by definition.

T5 *Temporal structure of a sound: homogeneity*

parameter value	**explanation**	**examples**	
		English	**German**
heterogeneous	perceptible segments of sound differ	*bluster, thunder, crow, rumble, roar*	*donnern, krähen, rumpeln, brüllen, heulen, poltern*
homogeneous	perceptible segments of sound are equal	*whistle, hiss, purr*	*pfeifen, zischen, schnurren*

[3] This observation is due to Johannes Helmbrecht (p.c.).

This parameter is operationalized as follows: Given a sound of some duration with impulse rate 'crepitative' or greater, and two perceptible segments of equal length selected at random, then the sound is homogeneous if any two segments are alike; and it is heterogeneous otherwise. Thus, a sound whose constitutive impulses alternate with segments of silence is heterogeneous.

Some German ideophones have a complex morphological structure which reflects heterogeneity: *klirrdibirr, holterdipolter*.

Words designating heterogeneous sounds also have a complex syllable structure: *rumpeln, poltern, brutzeln*.

From homogeneity we must distinguish **monotony**. A homogeneous sound is necessarily monotonous. A heterogeneous sound is monotonous if its single impulses are all alike and follow each other rhythmically; otherwise they are varied. For instance, the tick-tack of a clock is heterogeneous and monotonous; a thunder is heterogeneous and varied.

4.1.2.3 Close of impulse

The last parameter of horizontal structure refers either to an impulse sound or to a single impulse of an iterated sound.

T6 *Temporal structure of a sound: close of impulse*

parameter value	**explanation**	**examples**	
		English	**German**
checked	sudden, "dry", "blunt" close	*thud, pop, rattle*	*glucken, rattern*
fading	dwindling close	*bang, clatter*	*klirren*

The parameter is often reflected in the phonological structure of the root of the sound word: Checked sounds are expressed with an obstruent, fading sounds with a sonorant in the rhyme. Some relevant German examples are contained in T7.

T7 *German onomatopeia of 'checked vs. fading'*

checked		**fading**	
rattern	rattle	klirren	clink, clank, jingle, tinkle
knacken	crack(le)	krachen	crack, crash
gluck	gurgle	knallen	bang, crack
klappern	clatter	pummern	bump
ticken	tick	puffen	puff, chuff

4.2 Evaluation

We are now coming to the non-auditory parameters of the categorization of sounds. The first of these is the positive vs. negative evaluation of sounds, as shown in T8.

T8 *Evaluation of a sound*

parameter value	**explanation**	**examples**	
		English	**German**
pleasant	word connotes a positive evaluation	*hum, burble, purr*	*summen, surren, schnurren*
unpleasant	word connotes a negative evaluation	*crack, creak*	*krachen, quietschen, kreischen*

There are relatively few words for pleasant sounds, negative evaluation predominates. This fact is the basis for a whole subdiscipline of psychoacoustics, viz. noise research.

4.3 Properties of the sound source

In acoustic terms, the sound source is a vibrating physical body. In addition, an energy source is required which puts the sound source into vibration. If the sound is represented linguistically by an intransitive verb, then the sound source typically appears as its subject, as in E2.

E2. a. The whip cracked.
b. The creek murmured.

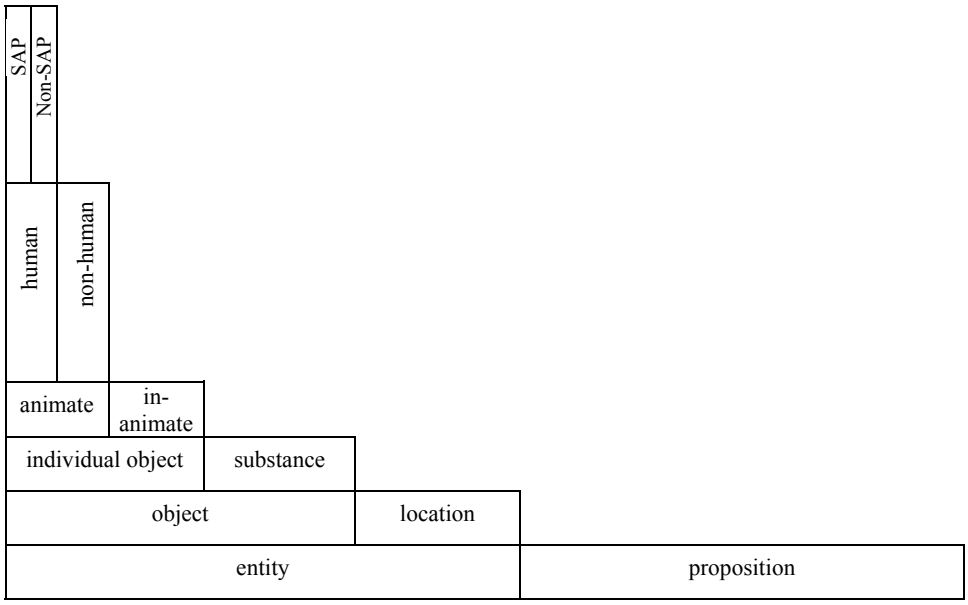
For many sound verbs, the subject does not represent the sound source, but instead the energy source, as in E3.

E3. a. The rooster crowed.
b. Philip murmured.

In the situations designated by the sentences of E3, the sound sources are the articulatory apparatuses of the beings involved, while the beings themselves – represented by the sentence subjects – are the energy sources. The essential difference between E2 and E3 is, obviously, that the subjects of the latter are animate, whereas those of the former are inanimate.

The most important parameter by which sound and energy sources may be classified is, in fact, the hierarchy of participant properties. This hierarchy, called *animacy hierarchy* (cf. Comrie 1981, Ch. 9) or *empathy hierarchy* (cf. Kuno 1987), is represented in F1.

F1. *Empathy hierarchy*



The empathy hierarchy is pervasive in structuring linguistic systems. T9 shows its role in the selection restrictions of sound verbs.

T9 *Selection restrictions of sound verbs*

sound/energy source	**examples**	
	English	**German**
human	*cough*	*husten*
mammal	*howl*	*heulen*
bird	*peep*	*piepen*
lower animal	*chirp*	*zirpen*
hard objects	*gnash*	*knirschen*
technical device	*click*	*klicken*
loose material	*rustle*	*rascheln*

What is perceptually the sound or energy source of a sound materializes grammatically as a selection restriction that the verb designating the sound possesses for its subject. These are, thus, semantic components of a different kind that could, in principle, be disregarded in an analysis focused on auditory categorization.

Metaphor in the usual sense, whereby a verb is used with a subject that does not fit its literal selection restrictions, obviously plays an important role in this domain. This is sometimes true to the extent to render a decision on the primary sound/energy source difficult. Does *whistle* take primarily an animate or an inanimate subject?

We saw before that the higher the subject of a sound verb on the empathy hierarchy, the more likely it will designate the energy source rather than the sound source. Verbs like *whisper* or *groan* designate a dynamic situation whose only participant is a human being. Human beings are not primarily conceived as sound sources, but as active controllers of a situation. For sound verbs with animate subjects, this perspective on the situation in question is more important than the purely auditory perspective. Whispering and groaning means primarily speaking in a certain way or giving expression to strain or suffering; the aspect of producing a sound with certain auditory qualities remains subordinate.

The nature of the energy source may completely prevail over the auditory properties of the sound. For instance, *neigh* is first and foremost the sound that a horse makes; and conversely, if one wants to say of a horse that it makes its default sound, one has no choice but to use the verb *neigh*. The consequence is that the language user is aware of no specific auditory qualities of the sound designated by *neigh* and consequently is put at a loss to make them explicit. A methodological consequence of this is that verbs with animate, and especially human, subjects are not typical sound verbs if one is interested in sound verbs as a lexical categorization of auditory percepts.

A more refined hypothesis would be the following: The less strict and homogeneous (the looser and the more heterogeneous) the selection restrictions of a sound verb are, the stronger is its purely auditory semantic core. For German *knirschen*, e.g., selection restrictions are very heterogeneous (teeth, gravel, snow).

4.4 The system of the parameters

The paradigmatic structure of the lexical field of sound words has so far been reduced to a set of parameters that define binary or multiple contrasts. However, just as in other lexical fields, diverse kinds of lexical relations are conceivable between sound words. They might be in hyponymy or part-whole relations. There might be a classificatory hierarchy of sound words. One would then expect, among other things, that if a language only has a small set of sound words, these express the basic categories, just as it has been shown in the domain of color terms.

The parameters and their values have been introduced as an unordered set. However, they are probably of unequal weight. There are markedness and dependency relations among them. These would emerge if the lexical fields are structured in the ways indicated.

As always in linguistics, the features are defined in relational terms. Nowhere is an absolute physical value relevant; it is always by its paradigmatic oppositions and its syntagmatic contrasts that a feature gets its identity (Jakobson 1968).

5 Onomatopoeia

Since most of the sounds around us are noises, or at any rate aperiodic vibrations prevail in them, speech sounds whose features involve aperiodic vibrations prevail in their onomatopoetic

designation. This is why consonants are more important in onomatopoetic sound words than vowels.

6 Sound and movement

Most sound words designate a movement or an action together with the sound it produces: *gnash/knirschen, knock/klopfen*.

Yucatec Maya even has a generic verb *pek* which means both 'move' and 'sound'.

Folk definitions of sound words rely not on acoustic features but on the source of sound production and the act or movement which produces it. This is most obviously the case for sounds produced by animate beings, such as *crow* and *bark*, but even for inanimate sounds such as *jingle* and *clatter*.

It is not always easy to distinguish sound words from movement/manipulation words. Expressions such as E4 are frequent.

E4. There was a scratching sound.

However, one should not conclude from them that the verbs in question designate sounds.[1] To decide such questions, one needs test frames that differentiate between sound and movement/manipulation.

7 Towards a theory of the linguistic categorization of sounds

The categories of sounds are, for the most part, subconscious. We are capable of reacting appropriately, within fractions of a second, to a perceived sound, but we are essentially incapable of describing it in general and intersubjective terms.

On hearing a sound, people tend to produce a diagnosis which is not a description of the auditory percept but an inference concerning its source and the way it was produced. Thus, they say:

E5. Someone is coming up the stairs.

And not:

E6. There is an iterated creaking sound coming from outside.

Attention to the action – movement or impact on an object etc. – prevails over attention to the sound as such. Hearers are not concerned with the acoustic nature of the sound but with what it means to them. They take the sound as an **indexical sign**.

Here we are faced with an important difference between visual and auditory perception: **Visual perception** as such is indifferent to the dynamicity of the perceived situation, i.e. the situation may be static as well as dynamic. Visual perception functions importantly in spatial and social

[1] One must also not be too quick in assigning such a word purely to the class of motion and manipulation. Yucatec Maya *boboh*, obviously an onomatopoetic word, means 'sound hollow, knock'. Thus, metaphoric transfers in both directions are possible.

orientation. **Auditory perception** is restricted to dynamic situations.[2] It functions importantly in the adaptation of the behavior of the perceiver to a stimulus.

This gives occasion to the following **hypothesis**: Human auditory perception is prepared to interpret sounds functionally. The leading question in categorizing an auditory percept is: What does this sound mean to me? Human auditory perception is thus well prepared for the perception of speech sounds.

The sounds of language are categorized according to the same parameters as non-linguistic sounds.[3] The subconscious, implicit nature of the phonological organization of human language is prefigured in the use that we make of non-linguistic sounds.

References

Blaubergs, Maija S. 1980, "Are structural features of word meaning reflected in judgements of semantic similarity and difference?" *Language Sciences* 2:1-31.

Gelfand, Stanley A. 1998, *Hearing. An introduction to psychological and physiological acoustics.* New York etc.: Marcel Dekker (3. ed.).

Hinton, Leanne & Nichols, Johanna & Ohala, John J. (eds.) 1994, *Sound symbolism.* Cambridge etc.: Cambridge University Press.

Jakobson, Roman 1968, "The role of phonic elements in speech perception." *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 21:9-20.

Kulemeka, Andrew Tilimbe 1995, "Sound symbolic and grammatical frameworks. A typology of ideophones in Asian and African languages." *South African Journal of African Languages* 15(2): 73-84.

Lehmann, Christian 1991, "Predicate classes and PARTICIPATION." Seiler, Hansjakob & Premper, Waldfried (eds.), *PARTIZIPATION. Das sprachliche Erfassen von Sachverhalten.* Tübingen: G. Narr (LUS, 6); 183-239.

Pilch, Herbert 1978, "Auditory phonetics." *Word* 29:148-60.

Wissemann, Heinz 1954, *Untersuchungen zur Onomatopoiie. 1. Teil: Die sprachpsychologischen Versuche.* Heidelberg: C. Winter (Bibliothek der Allgemeinen Sprachwissenschaft; 2. Reihe).

[2] This is trivially true, because the production of a sound presupposes a vibration. Although the eye cannot discern a movement in every case, the immediate inference if we hear a sound is that something must be going on; i.e. we suspect a dynamic situation.

[3] Speech sounds are categorized primarily in auditory, not in articulatory, terms; see Jakobson 1968.