Manfred Krifka

Humboldt-Universität zu Berlin, Institut für deutsche Sprache und Linguistik

und Zentrum für Allgemeine Sprachwissenschaft, Berlin

Unter den Linden 6

10099 Berlin

krifka@rz.hu-berlin.de

030-20929670

# 13. Varieties of semantic evidence

Outline



Abstract

Meanings are the most elusive objects of linguistic research. The article summarizes the type of evidence we have for them: various types of metalinguistic activities like paraphrasing and translating, the ability to name entities and judge sentences true or false, as well as various behavioral and physiological measures such as reaction time studies, eye tracking, and electromagnetic brain potentials. It furthermore discusses the specific type of evidence we have for different kinds of meanings, such as truth-conditional aspects, presuppositions, implicatures, and connotations.

## 1 Introduction: Aspects of Meaning and Possible Sources of Evidence

### *1.1 Why meaning is a special research topic*

If we ask an astronomer for evidence for phosphorus on Sirius, she will point out that spectral analysis of the light from this star reveals bands that are characteristic of this element, as they also show up when phosphorus is burned in the lab. If we ask a linguist the more pedestrian question for evidence that a certain linguistic expression – say, the sentence *The quick brown fox jumps over the lazy dog* – has meaning, answers are probably less straightforward and predictable. He might point out that speakers of English generally agree that it has meaning – but how do they know? So it is perhaps not an accident that the study of meaning is the subfield of linguistics that developed only very late in the 2500 years of history of linguistics, in the 19th century (cf. article 9 *The emergence of semantics as a discipline*).

The reason why it is difficult to imagine what evidence for meaning could be is that it is difficult to say what *meaning* is. According to a common assumption, communication consists in putting *meaning* into a *form*, a form that is then sent from the speaker to the addressee (the conduit metaphor of communication, see Lakoff & Johnson 1980). Aspects that are concerned with the form of linguistic expressions and their material realization as studied in syntax, morphology, phonology and phonetics; they are generally more tangible than aspects concerned with their content. But semanticists in general hold that semantics, the study of linguistic meaning, indeed has an object to study that is related but distinct from the forms in which it is encoded, from the communicative intentions of the speaker and from the resulting understanding of the addressee.

### *1.2 Aspects of Meaning*

The English noun *meaning* is multiply ambiguous, and there are several readings that are relevant for semantics. One branch of investigation starts out with meaning as a notion rooted

in *communication*. Grice (1957) has pointed out that we can ask what a *speaker* meant by uttering something, and what the *utterance* means that the speaker uttered. Take John F. Kennedy's utterance of the sentence *Ich bin ein Berliner* on June 26, 1963. What JFK meant was is that in spite of the cold war, the USA would not surrender West Berlin – which was probably true. What the utterance meant was that JFK is a citizen of Berlin, which was clearly false. Obviously, the speaker's meaning is derived from the utterance meaning and the communicative situation in which it was uttered. The way how this is derived, however, is less obvious – cf. article 2 *Meaning, Intentionality and Communication*, especially on particularized conversational implicatures.

A complementary approach is concerned with the meaning of linguistic *forms*, sometimes called *literal meanings*, like the meaning German sentence *Ich bin ein Berliner* which was uttered by JFK to convey the intended utterance meaning. With forms, one can distinguish the following aspects of meaning (cf. also 3, *Sense and Reference*, and 4, *Reference: Foundational Issues*). The *character* is the meaning independent from the situation of utterance (like speaker, addressee, time and location – see Kaplan 1979). The character of the sentence used by JFK is that the speaker of the utterance is a citizen of Berlin at the time of utterance. If we find a sticky note in a garbage can, reading *I am back in five minutes* – where we don't know the speaker, the time, or location of the utterance – we just know the character. A character, supplied with the situation of utterance, gets us the *content* or *intension* (Frege's *Sinn*) of a linguistic form. In the proper historical context, JFK's utterance has the content that JFK is a citizen of Berlin on June 26, 1963. (We gloss over the fact here that this first has to be decoded as a particular speech act, like an assertion.). This is a proposition, which can be true or false in particular circumstances. The *extension* or *reference* of an expression (Frege's *Bedeutung*) is its content when applied to the situation of utterance. In the case of a proposition, this is a truth value; in the case of a name or a referring expression, this is an entity.

Sometimes *meaning* is used in a more narrow sense, as opposed to reference; here I have used *meaning* in an encompassing way.

Arguably, the communicative notion of meaning is the primary one. Meaning is rooted in the intention to communicate. But human communication crucially relies on linguistic forms, which are endowed with meaning as outlined, and for which speakers can construct meanings in a compositional way (see 6, *Compositionality*). Semantics is concerned with the meaning of linguistic forms, a secondary and derived notion. But the use of these forms in communication is crucial data to re-engineer the underlying meaning of the forms. The ways how literal meanings are used in acts of communication and their effects on the participants, in general, is part of pragmatics (cf. article 99 *Semantics and Pragmatics*).

### *1.3 Types of Access to Meaning*

Grounding meaning of linguistic expressions in communication suggests that there are various kinds of empirical evidence for meaning. First, we can observe the external behavior of the participants in, before and after the act of communication. Some kinds of behavior can be more directly related to linguistic meaning than others, and hence will play a more central role in discovering underlying meaning. For example, commands often lead to a visible non-linguistic reaction, and simple *yes/no*-questions will lead to linguistic reactions that are easily decodable. Secondly, we can measure aspects of the external behavior in detail, like the reaction times to questions, or the speed in which passages of text are read (cf. article 115 *Meaning in Psycholinguistics*). Third, we can observe physiological reactions of participants in communication, like the changing size of their pupil, the saccades of the eyes reading a text, the eye gaze when presented to a visual input together with a spoken comment, or the electromagnetic field generated by their cortex (cf. article 15 *Experimental Methods in Semantics*, and article 118 *Meaning in Neurolinguistics*). Fourth, we can test hypotheses concerning

meaning in the output of linguistic forms itself, using statistical techniques applied to corpora (cf. article 124 *Corpus Linguistics*).

### *1.4 Is Semantics Possible?*

The reader should be warned that correlations between meanings and observable phenomena like non-linguistic patterns of behavior or brain scans do not guarantee that the study of meaning can be carried out successfully. Leonard Bloomfield, a behavioralist, considered the observable effects so complex and interwoven with other causal chains that the science of semantics is impossible:

"We have defined the meaning of a linguistic form as the situation in which the speaker utters it and the response which it calls forth in the hearer. [...] In order to give a scientifically accurate definition of meaning for every form of a language, we should have to have a scientifically accurate knowledge of everything in the speaker's world. The actual extent of human knowledge is very small compared to this. [...]. The statement of meanings is therefore the weak point in language-study, and will remain so until human knowledge advances very far beyond its present state." (Bloomfield 1933: 139f.)

We could imagine similar skepticism concerning the science of semantics from a neuroscientist believing that meanings are activation patterns of our head. The huge number of such patterns, and their variation across individuals that we certainly have to expect, seems to preclude that they will provide the foundation for the study of meaning.

Despite Bloomfield's qualms, the field of semantics has flourished. Where he went wrong was in believing that we have to consider the whole world of the speaker, or the speaker's whole brain. There are ways to cut out phenomena that stand in relation to, and bear evidence for, meanings in much more specific ways. For example, we can investigate whether a specific sentence in a particular context and describing a particular situation is considered true or

false; and derive from that hypotheses about the meaning of the sentence and the meaning of the words involved in that sentence. The usual methods of science – forming hypotheses and models, deriving predictions, making observations and constructing experiments that support or falsify the hypotheses – have turned out to be applicable to linguistic semantics as well.

### *1.5 Native Semantic Activities*

There are many native activities that directly address aspects of meaning. When Adam named the animals of paradise he assigned expressions to meanings, as we do today when naming things or persons or defining technical terms. We explain the meaning of words or idioms by paraphrasing them – that is, by offering different expressions with the same or at least similar meanings. We can refer to aspects of meaning: We say that one expression *means* the same as another one, or its opposite; we say that one expression *refers* to a subcase of another. As for speaker's meanings, we can elaborate on what someone meant by such-and-such words, and can point out differences between that and what the words actually meant or how they were understood by the addressee. Furthermore, for human communication to work it is crucial that untruthful use of language can be detected, and liars can be identified and punished. For this, a notion of what it means for a sentence or text to be true or false is crucial. Giving a statement at court means to know what it means to speak the truth, and the whole truth. Hence, it seems that meanings are firmly established in the prescientific ways we talk about language.

We can translate, that is, rephrase an expression in one language by an expression in another while keeping the meaning largely constant. We can teach the meaning of words or expressions to second language learners or to children acquiring their native language – even though both groups, in particular children in first language acquisition, will acquire meanings to a large part implicitly, by contextual clues. The sheer possibility of translation has been

enormously important for the development of humankind. We find records of associated practices, like the making of dictionaries, dating back to Sumerian-Akkadian glossaries of 2300 BC.

These linguistic activities show that meaning is a natural notion, not a theoretical concept. They also provide important source of evidence for meaning. For example, it would be nearly impossible to construct a dictionary in linguistic field work without being able to ask for what a particular word means, or how a particular object is called. As another example, it would be foolish to dismiss the monumental achievements of the art of dictionary writing as evidence for the meaning of words.

But there are problems with this kind of evidence that one must be aware of. Take dictionary writing. Traditional dictionaries are often unsystematic and imprecise in their description of meaning. They do not distinguish systematically between contextual (or "occasional") meaning and systematic meaning, nor do they keep ambiguity and polysemy apart in a rigorous way. They often do not distinguish between linguistic aspects and more general cultural aspects of the meaning and use of words. Weinreich (1964) famously criticized the 115 meanings of the verb *to turn* that can be found in Webster's Third Dictionary. Lexicography has greatly improved since then, with efforts to define lexical entries by a set of basic words and by recognizing regularities like systematic variations between word meanings (e.g. the intransitive use of transitive verbs, or the polysemy triggered in particular contexts of use).

### *1.6 Talking about Meanings*

Pre-scientific ways to address meanings rely on an important feature of human language, its *reflexivity* – we can use language to talk about language. This feature is so entrenched in language that it went unnoticed until logicians like Frege, Russell and Tarski, working with much more restricted languages, pointed out the importance of the metalanguage / object

language distinction. It is only quite recently that we distinguish between regular language, reference to expressions, and reference to meanings by typographical conventions and write things like "*XXX* means 'YYY'".

The possibility to describe meanings may be considered circular – as when Tarski states that '*Snow is white' is true if and only if snow is white.* However, it does work under certain conditions. First, the meaning of an unknown word can be described, or at least delimited, by an expression that uses known words; this is the classical case of a definition. If we had only this procedure available as evidence for meaning, things would be hopeless because we have to start somewhere with a few expressions whose meanings are known; but once we have those, they can act as bootstraps for the whole lexicon of a language. The theory of Natural Semantic Metalanguage even claims that a small set of concepts (around 200) and a few modes of combining them are sufficient to achieve access to the meanings of all words of a language (Goddard 1998).

Second, the meanings of an ambiguous word or expression can be paraphrased by expressions that have only one or the other meaning. This is common practice in linguistic semantics, e.g. when describing the meaning of *He saw that gasoline can explode* as (a) 'He saw an explosion of a can of gasoline' and (b) 'He recognized the fact that gasoline is explosive'. Speakers will generally agree that the original sentence has the two meanings teased apart by the paraphrases. There are variations on this access to meaning. For example, we might consider a sentence in different linguistic contexts and observe differences in the meaning of the sentence by recognizing that it has to be paraphrased differently. For the paraphrases, we can use a language that has specific devices that help to clarify meanings, like variables. For example, we can state that a sentence like *Every man likes a woman that likes him* has a reading 'Every man x likes a woman y that likes x', but not 'There is a woman y that every man x likes and that likes x'. The disadvantage of this is that the paraphrases cannot be easily

grasped by naïve native speakers. In the extreme case, we can use a fully specified formal language to specify such meanings, such as first-order predicate logic; the existing reading of our example then could be specified as $\forall x[man(x) \rightarrow \exists y[woman(y) \wedge likes(x, y) \wedge likes (y, x)]]$.

Talking about meanings is a very important source of evidence for meanings. However, it is limited not only by the problem mentioned above, that it describes meanings with the help of other meanings. There are many cases where speakers cannot describe the meanings of expressions because this task is too complex – think of children acquiring their first language, or aphasics loosing the capacity of language. And there are cases in which the description of meanings would be too complex for the linguist. We may think of first fieldwork sessions in a research project on an unknown language. Somewhat closer to home, we may also think of the astonishingly complex meanings of natural language determiners such as *a, some, a certain, a particular, a given* or indefinite *that* in *there was that man standing at the door* whose meanings had to be teased apart by careful considerations of their acceptability in particular contexts.

## 2 Fieldwork Techniques in Semantics

In this section we will discuss various techniques that have been used in linguistic fieldwork, understood in a wide sense as to include work on one's own language and on language acquisition, for example. There are a variety of sources that reflect on possible procedures; for example, the authors in McDaniel e.a. (eds.) (1996) discuss techniques for the investigation of syntax in child language, many of which also apply to semantic investigations, and Matthewson (2004) is concerned with techniques for semantic research in American languages which, of course, are applicable for work on other languages as well (cf. also article 13 *Methods in Cross-Linguistic Semantics*, and article 116 *Semantics in First Language Ac-*

*quisition*).

### *2.1 Observation, Transcription and Translation*

The classical linguistic fieldwork method is to record conversations and texts in natural settings, transcribe them, and assign translations, ideally with the help of speakers that are competent in a language that they share with the investigator. In classical American structuralism, this has been the method *de rigueur*, and it is certainly of great importance when we want to investigate natural use of language.

However, this technique is also severely limited. First, even large text collections may not provide the evidence that distinguishes between different hypothesis. Consider superlatives in English; is *John is the tallest student* true if John and Mary both are students that are of the same height and taller than any other student? Competent English speakers say no, superlatives must be unique – but it might be impossible to find out on the basis of a corpus of non-elicited text.

Secondly, there is the problem of translation. Even when we grant that the translation is competent according to usual standards, it is not clear how we should deal with distinctions in the object language that are not easily made in the meta language. For example, Matthewson (2004) shows that in Menominee (Algonquian, Northern Central United States of America) , inalienable nouns can have a prefix *me-* indicating an arbitrary owner, as contrasted with a prefix *o-* indicating a specific 3rd person owner. This difference could not be derived from simple translations of Menominee texts into English, as English does not make this distinction. There is also the opposite problem of distinctions that are forced on us by the meta language; for example, pronouns in English referring to humans distinguish two genders, which may not be a feature of the object language. Hence, as Matthewson puts it, translations should be seen as clues for semantic analysis, rather as its result.

Translations, or more generally paraphrases, are problematic for more fundamental reasons as evidence for meaning, as they explain the meaning of an expression α by way of the meaning of an expression β, hence it presupposes the existence and knowledge of meanings, and a judgment of similarity of meaning. However, it appears that without accepting this type of hermeneutic circle the study of semantics could not get off the ground. But there are methods to test hypotheses that have been generated first with the help of translations and paraphrases by independent means.

### *2.2 Pointing*

Pointing is a universal non-linguistic human behavior that aligns with aspects of meanings of certain types of linguistic expressions (cf. also article 101 *Deixis and Demonstratives*). Actually, pointing may be as characteristic for humans as language, as humans appear to be the only apes that point (cf. Tomasello 2008).

Pointing is most relevant for referring expressions, with names as the prototypical example (cf. article 4 *Reference: Foundational Issues*). These expressions denote a particular entity that is also identified by the pointing gesture, and hence pointing is independent evidence for the meaning of such expressions. For example, if in a linguistic fieldwork situation an informant points to a person and utters *Max*, this might be taken to be the name of that person. We can conclude that *Max* denotes that person, in other words, that the meaning of *Max* is the person pointed at.

Simple as this scenario is, there are certain prerequisites for it to work. For example, the pointing gesture must be recognized as such; in different cultures, the index finger, the stretched-out hand, or an upward movement of the chin may be used, and in some cultures there may be a taboo against pointing gestured when directed at humans. Furthermore, there must be one most salient object in the pointing cone (cf. Kranstedt e.a. 2006) that will then be

identified. This presupposes a pre-linguistic notion of objects, and of saliency. This might work well when persons or animals are pointed at, who are cognitively highly salient. But mistakes can occur when there is more than one object in the pointing cone that are equally salient. When Captain Cook on his second voyage visited an island in the New Hebrides with friendly natives and tried to communicate with them, he pointed to the ground. What he heard was *tanna*, which he took as the name of the island, which is still known under this name. Yet the meaning of *tana* in all Melanesian languages is simply "earth". The native name for Tanna is reported to be *parei* (Gregory 2003); it is not in use anymore.

Pointing gestures may also help to identify the meaning of common nouns, adjectives, or verbs – expressions that denote sets of entities or events. The pointing is directed towards a specimen, but reference is at entities of the same type as the one pointed at. There is an added source of ambiguity or vagueness here: What is "the same type as"? On his first voyage, Captain Cook made landfall in Australia, and observed creatures with rabbit-like ears hopping on their hind legs. When naturalist Joseph Banks asked the local Guugu Yimidhirr people how they are called, presumably with the help of some pointing gesture, he was the first to record the word *kangaroo*. But the word *gangurru* actually just refers to a large species of black kangaroo, not to the marsupial family in general (cf. Haviland 1974).

Quine (1960), in an argument to discount the possibility of true translation, famously described the problems that even a simple act like pointing and naming might involve. Assume a linguist points to a white rabbit, and gets the response *gavagai*. Quine asks whether this may mean 'rabbit', or perhaps 'animal', or perhaps 'white', or perhaps even 'undetached rabbit parts'. It also might mean 'rabbit stage', in which case repeated pointing will identify different reference objects. All these options are theoretical possibilities under the assumption that words can refer to arbitrary aspects of reality. However, it is now commonly assumed that language is build on broad cognitive commonalities about entities and classes. There is

evidence that pre-linguistic babies and higher animals have concepts of objects (as contrasted to substances) and animals (as contrasted to lifeless beings) that preclude a conceptualization of a rabbit as a set of rabbit legs, a rabbit body, a rabbit head and a pair of rabbit ears moving in unison. Furthermore, there is evidence that objects are called with terms of a middle layer of a taxonomic hierarchy, the so-called "generic level", avoiding terms that are too general or too specific (cf. Berlin, Breedlove & Raven 1973). Hence a rabbit will not be called *thing* in English, or *animal*, and it will not be called *English angora* either except perhaps by rabbit breeders that work with a different taxonomy. This was the reason for Captain Cooks misunderstanding of *gangurru*; the native Guugu Yimidhirr people had a different, and more refined, taxonomic hierarchy for Australian animals, where species of kangaroo formed the generic level; for the British visitors the family itself belonged to that level.

Pointing, or related gestures, have been used to identify the meaning of words. For example, in the original study of Berlin & Kay (1969) on color terms subjects were presented with a two-dimensional chart of 320 colors varying according to spectral color and saturation. The task was to identify the best specimen for a particular color word (the focal color) and the extent to which colors fall under a particular color word. Similar techniques have been used for other lexical fields, for example for the classification of vessels using terms like *cup, mug* or *pitcher* (cf. Kempton 1981; see Figure 1).

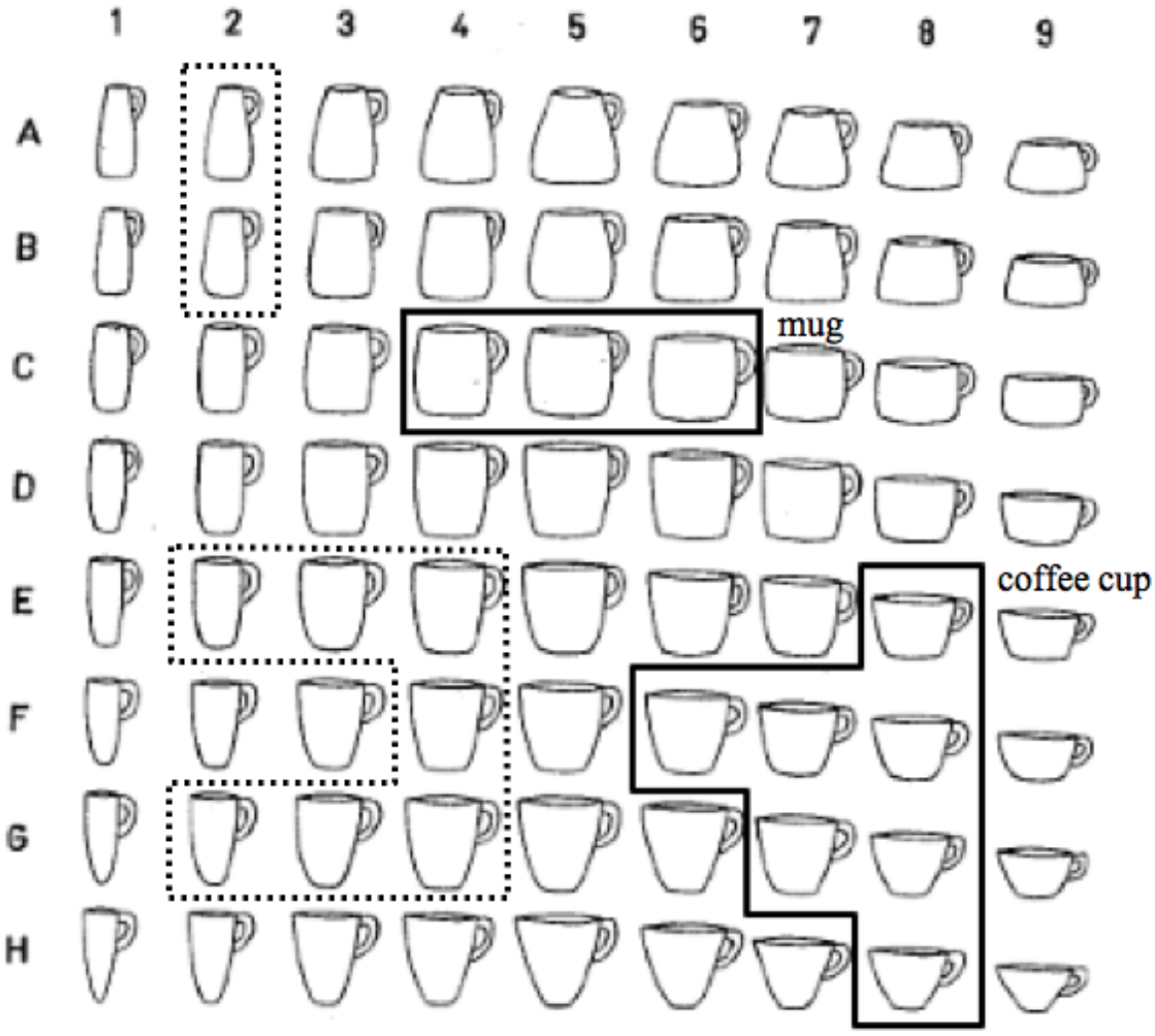


FIGURE 1: Vessel categories after Kempton (1981). Bold lines: Identification of >80% agreement between subjects for *mug* and *coffee cup*. Dotted lines: Hypothetical concept that would violate connectedness and convexity (see below).

Tests of this type have been carried out in two ways: Either subjects were presented with a field of reference objects ordered after certain dimensions; e.g., Berlin & Kay (1969) presented colors ordered after their wave length (the order they present themselves in a rainbow) and after their saturation (with white and black as the extremes). Kempton's vessels were presented as varying in two dimensions: The relation between the upper and lower diameters, and the relation between height and width. When judging whether certain items fall under a

term or not, the neighboring items that already have been classified might influence the decision. Another technique, which was carried out in the World Color Survey (see Kay e.a. 2008), presented color chips in random order to avoid this kind of influence.

The pointing test can be used in two ways: Either we point at an entity in order to get the term that is applicable to that entity, or we have a term and point to various objects to find out whether the term is applicable. The first approach asks an *onomasiological* question; it is concerned with the question: How is this thing called? The second approach asks the complementary *semasiological* question: What does this expression mean?

Within a Fregean theory of meaning, a distinction is made between reference and sense (cf. article 3 *Sinn and Bedeutung*). With pointing to concrete entities we gain access to the reference of expressions, and not to the sense, the concept that allows us to identify the reference. But by varying potential reference objects we can form hypotheses about the underlying concept, even though we can never be certain that by a variation of reference objects we will uncover all aspects of the underlying concept. Goodman (1955) illustrated this with the hypothetical adjective *grue* that, say, refers to green objects when used before the year 2100 and to blue objects when used after that time; no pointing experiment executed before 2100 could differentiate *grue* from *green*. Meaning shifts like that do happen historically: The term *Scotia* refered to Ireland before the 11$^{th}$ century, and after to Scotland; the German term *gelb* was reduced in extension when the term *orange* entered the language (cf. the traditional local term *Gelbe Rüben* 'yellow turnips' for carrots). But these are language changes, and not meanings of items within a language. A meaning like the hypothetical *grue* appears as strange as a reference towards non-detached rabbit parts. We work under the hypothesis that meanings of lexical items are restricted by general principles of uniformity over time. There are other such principles that restrict possible meanings, for example connectedness and, more specifically, convexity (Gärdenfors 2000). In the vessel example above, where potential

reference objects were presented following certain dimensions, we expect that concepts do not apply to discontinuous areas and have the general property that when x is an $\alpha$ and y is an $\alpha$, then everything in between x and y is an $\alpha$ as well. The dotted lines in Figure 1 represent an extension of a concept that would violate connectedness and convexity.

In spite of all its problems, pointing is the most elementary kind of evidence for meaning without which linguistic field work, everyday communication and language acquisition would be impossible. Yet it seems that little research has been done on pointing and language acquisition, be it first or second. Its importance, however, was recognized as early as in St. Augustin's *Confessions*, where he writes about his own learning of language:

"When they [the elders] called some thing by name and pointed it out while they spoke, I saw it and realized that the thing they wished to indicate was called by the name they then uttered. And what they meant was made plain by the gestures of their bodies, by a kind of natural language, common to all nations […]" (Confessions, Book I:8).

### 2.3 *Truth Value Judgments (TVJ)*

Truth value judgments do the same job for the meaning of sentences as pointing does for referring expressions. In the classical setup, a situation is presented with non-linguistic means together with a declarative sentence, and the speaker has to indicate whether this sentence is true or false with respect to the situation. This judgement is an linguistic act by itself, so it can be doubted that this provides a way to base the study of meaning wholly outside of language. But arguably, agreeing or disagreeing are more primitive linguistic acts that may even rely on simple gestures, just as in the case of pointing.

The similarity between referential expressions – which identify objects – and declarative sentences – which identify states of affairs in which they are true – is related to Frege's identification of the reference of sentences with their truth value with respect to a particular situa-

tion (even though this was not the original motivation for this identification, cf. Frege 1892). This is reflected in the two basic extensional types assumed in sentence semantics: Type e for entities referred to by names, and type t for truth values referred to by sentences. But there is an important difference here: There are many distinct objects – $D_e$, the universe of discourse, is typically large; but there are just two (basic) truth values – $D_t$, the set of truth values, standardly is {0, 1}, falsity and truth. Hence we can distinguish referring expressions more easily by their reference than we can distinguish declarative sentences. One consequence of this is that onomasiological tests do not work. We cannot present a "truth value" and expect a declarative sentence that is true. Also, on presenting a situation in a picture or a little movie we cannot expect that the linguistic reactions are as uniform as when we, say, present the picture of an apple. But the semasiological direction works fine: We can present speakers with a declarative sentence and a situation or a set of situations and ask whether the sentence is true in those situations.

Truth values are not just an ingenious idea of language philosophers to reduce the meaning of declarative sentences to judgments whether a sentence is true or false in given situations. They are used pre-linguistically, e.g. in court procedures. Within linguistics, they are used to investigate the meaning of sentences in experiments and in linguistic field work. They have been particularly popular in the study of language acquisition because they require a rather simple reaction by the child that can be expected even from two-year olds.

The TVJ task comes in two flavors. In both, the subjects are presented with a sentence and a situation, specified by a picture, an acted-out scene with hand puppets or a movie, or by the actual world provided that the subjects have the necessary information about it. In the first version, the subjects should simply state whether the sentence is true or false. This can be done by a linguistic reaction, by a gesture, by pressing one of two buttons, or by ticking off one of two boxes. We may also record the speed of these reactions in order to get data about

the processing of expressions. In the second version, there is a character, e.g. a hand puppet, that utters the sentence in question, and the subjects should reward or punish the character if the sentence is true or false with respect to the situation presented (see e.g. Crain 1991). A reward could be, for example, feeding the hand puppet a cookie. Interestingly, the second procedure taps into cognitive resources of children that are otherwise not as easily accessible.

Gordon (1998), in a description of TVJ in language acquisition, points out that this task is quite natural and easy. This is presumably so because truth value judgment is an elementary linguistic activity, in contrast to, say, grammaticality judgments. TVJ also puts less demands on answers than wh-questions (e.g. *Who chased the zebra?* vs. *Did the lion chase the zebra?*) This makes it the test of choice for children and for language-impaired persons.

But there are potential problems in carrying out TVJ tasks. For example, Crain e.a. (1998) have investigated the phenomenon that children seem to consider a sentence like *Every farmer is feeding a donkey* false if there is a donkey that is not fed by the farmer. They argue that children are confused by the extra donkey and try to reinterpret the sentence in a way that seems to make sense. A setup in which attention is not drawn to a single object might be better; even adding a second unfed donkey makes the judgments more adult-like. Also, children respond better to scenes that are acted out than to static pictures. In designing TVJ experiments, one should consider the fact that positive answers are given quicker and more easily than negative ones. Furthermore, one should be aware that unconscious reactions of the experimenter may provide subtle clues for the "right" answer (the "Clever Hans" effect, named after the horse that supposedly could solve arithmetic problems). For example, when acting out and describing a scene, the experimenter may be more hesitant when uttering a false statement.

### *2.4 TVJ and presuppositions / implicatures*

There are different aspects of meaning beyond the literal meaning, such as presuppositions, conventional implicatures, conversational implicatures and the like, and it would be interesting to know how such meaning components fare in TVJ tasks. Take presuppositions (cf. also article 102 *Presuppositions*). Theories such as Stalnaker (1974) that treat them as preconditions of interpretation predict that sentences cannot be interpreted with respect to situations that violate their presuppositions. The TVJ test does not seem to support this view. The sentence *The dog is eating the bone* will most likely be judged true with respect to a picture showing two dogs, where one of the dogs is eating a bone. This may be considered evidence for the ease of accommodation, which consists of restricting the context to the one dog that is eating a bone. Including a third option or truth value like "don't know" might reveal the specific meaning contribution of presuppositions

As for conversational implicature (cf. article 103 *Implicature*) we appear to get the opposite picture. TVJ tests have been used to check the relevance of scalar implicatures. For example, Noveck (2001), building on work of Smith (1980), argued that children are "more logical" than adults because they can dissociate literal meanings from scalar implicatures. Children up to 11 years react to statements like *some giraffes have long necks* (where the picture shows that all giraffes have long necks) with an affirmative answer, while most adults find them inappropriate.

### *2.5 TVJ Variants: Picture Selection and Acting Out*

The picture selection task has been applied for a variety of purposes beyond truth values (cf. Gerken & Shady 1998). But for the purpose of investigating sentence meanings, it can be seen as a variant to the TVJ task: The subject is exposed to a declarative sentence and two or more pictures and has to identify the picture for which the sentence is true. It is good to in-

clude irrelevant pictures as filler items, which can test the attention of the subjects. The task can be used to identify situations that fit best to a sentence. For example, for sentences with presuppositions it is expected that a picture will be chosen that does not only satisfy the assertion, but also the presupposition. So, if the sentence is *The dog is eating a bone*, and if a picture with one or two dogs is shown, then presumably the picture with one dog will be preferred. Also, sentences whose scalar implicature is satisfied will be preferred over those for which this is not the case. For example, if the sentence is *some giraffes have long necks*, a picture in which some but not all giraffes have long necks will be preferred over a picture in which all giraffes are long-necked.

Another relative of the TVJ task is the Act Out task in which the subject has to "act out" a sentence with a scene such that the sentence is true. Again, we should expect that sentences are acted out in a way as to satisfy all meaning components – assertion, presupposition, and implicature – of a sentence.

### *2.6 Restrictions of the TVJ Methodology*

One restriction of the various TVJ methodologies appears to be that they just target expressions that have a truth value, that is, sentences. However, they allow to investigate the meaning of subsentential expressions, under the assumption that the meaning of sentences is computed in a compositional way from the meanings of their syntactic parts (cf. article 6 *Compositionality*). For example, the meaning of spatial presuppositions like *on*, *on top of, above* or *over* can be investigated with scenes in which objects are arranged in particular ways.

Another potential restriction of TVJ as discussed so far is that we assumed that the situations are presented by pictures. Language is not restricted to encoding information that can be represented by visual stimuli. But we can also present sounds, movie scenes or comic strips

that represent temporal developments, or even olfactory and tactile stimuli to judge the range of meanings of words (cf. e.g. Majid (ed.) 2006 for verbs of cutting and breaking).

TVJ is also difficult to apply when deictic expressions are involved, as they often require reference to the speaker, who is typically not part of the picture. For example, in English the sentence *The ball is in front of the tree* means that the ball is in between the speaker that faces the tree and the tree; the superficially corresponding sentence in Hausa means that the ball is behind the tree (cf. Hill 1982). In English, the tree is seen as facing the speaker, whereas in Hausa the speaker aligns with the tree (cf. article 110 *Expression of Space*). Such differences are not normally represented in pictures, but it can be done. One could either represent the picture from a particular angle, or represent a speaker with a particular position and orientation in the picture itself and ask the subject to identify with that figure.

The TVJ technique is systematically limited for sentences that do not have truth values, such as questions, commands, or exclamatives. But we can generalize it to a judgment of appropriateness of sentences given a situation, which sometimes is done to investigate politeness phenomena and the like. There are also subtypes of declarative sentences that are difficult to investigate with TVJ, namely modal statements, e.g. *Mary must be at home*, or habituals and generics that allow for exceptions, like *Delmer walks to school*, or *Birds fly* (cf. article 49 *Genericity*). This is arguably so because those sentences require to consider different possible worlds, which cannot be easily represented graphically.

### *2.7 TVJ with Linguistic Presentation of Situation*

The TVJ technique can be applied for modal or generic statements if we present the situation linguistically, by describing it. For example, we could ask whether *Delmer walks to school* is true if Delmer walks every day except Fridays, when his father gives him a ride. Of course, this kind of linguistic elicitation technique can be used in nearly all the cases de-

scribed so far. It has clear advantages: Linguistic descriptions are easy and cheap to produce and can focus the attention of the subject to aspects that are of particular relevance for the task. For this reason it is very popular for quick elicitations whether a sentence can mean such-and-such.

Matthewson (2004) argues that elicitation is virtually the only way to get to more subtle semantic phenomena. She also argues that it can be combined with other techniques, like TVJ and grammaticality judgments. For example, in investigating aspect marking in St'át'imcets Salish (Salishan, Suthwestern Canada) the sentence *Have you been to Seattle?* is translated using an adverb *lán* that otherwise occurs with the meaning 'already'; a follow-up question could be whether it is possible to drop *lán* in this context, retaining roughly the same meaning.

The linguistic presentation of scenes comes with its own limitations. There is the foundational problem that we get at the meaning of an expression α by way of the meaning of an expression β. It cannot be applied in case of insufficient linguistic competence, as with young children or language-impaired persons.

### *2.8 Acceptability Tests*

In this type of test, speakers are given an expression and a linguistic context and/or an description of an extralinguistic situation, and are asked whether the expression is acceptable with respect to this context or the situation. With it, we can explore the felicity conditions of an expression, which often are closely related to certain aspects of its meaning.

Acceptability tests are the natural way to investigate presuppositions and conventional implicatures of expressions. For example, additive focus particles like *also* presuppose that the predication holds for an alternative to the focus item. Hence in a context like *John went to Paris*, the sentence *John also went to PRAGUE* is felicitous, but the sentence *Mary also went*

*to PRAGUE* is not. Acceptability tests can also be used to investigate information-structural distinctions. For example, in English, different accent patterns indicate different focus structures; this can be seen when judging sentences like *JOHN went to Paris* vs. *John went to PARIS* in the context of questions like *Who went to Paris?* and *John went where?* (cf. article 74 *Questions*). As another example, Portner & Yabushita (1998) discussed the acceptability of sentences with a topic-comment structure in Japanese where the topic was identified by a noun phrase with a restrictive relative clause and found that such structures are better if the relative clause corresponds to a comment on the topic in the preceding discourse. Acceptability tests can also be used to test the appropriateness of terms with honorific meaning, or various shades of expressive meaning, which have been analyzed as conventional implicatures by Potts (2005).

When applying acceptability judgments, it is natural to present the context first, to preclude that the subject first comes up with other contexts which may influence the interpretation. Another issue is whether the contexts should be specified in the object language, or can also be given in the meta-language that is used to carry out the investigation. Matthewson (2004) discusses the various advantages and disadvantages – especially if the investigator has a less-than-perfect command over the object language – and argues that using a meta-language is acceptable, as language informants generally can resist the possible influence of the metalanguage on their responses.

### *2.9 Elicited Production*

We can turn the TVJ test on its head and ask subjects to describe given situations with their own words. In language acquisition research, this technique is known as "elicited production", and encompasses all linguistic reactions to planned stimuli (cf. Thornton 1998). In this technique the presumed meaning is fixed, and controls the linguistic production; we can

hypothesize about how this meaning can be represented in language. The best known example of probably is the retelling of a little movie called the Pear Story, which has unearthed interesting differences in the use of tense and aspect distinctions in different languages (cf. Chafe 1980 for the original publication). Another example, which allow to study the use of meanings in interaction, is the "map task", where one person explains the configuration of objects or a route on a map to another without visual contact.

The main problem of elicited production is that the number of possible reactions by speakers is, in principle, unlimited. It might well be that the type of utterances one expects do not occur at all. For example, we could set up a situation in which person A thinks that person B thinks that person C thinks that it is raining, to test the recursivity of propositional attitude expressions, but we will have to wait long till such utterances are actually produced. So it is crucial to select cues that constrain the linguistic production in a way that ensures that the expected utterances will indeed occur.

### *2.10 From Sentence Meanings to Word Meanings*

The TVJ technique and its variants test the meaning of sentences, not of words or subsentential expressions. Also, with elicitation techniques, often we will get sentence-like reactions. With elicited translations, it is also advisable to use whole sentences instead of single words or simpler expressions, as Matthewson (2004) argues. It is possible to elicit the basic meaning of nouns or certain verbs directly, but this is impossible for many other words. The first ten most frequent words in English are often cited as being *the, of, and, a, to, in, is, you, that*; it would be impossible to ask a naïve speaker of English what they mean or discover there meanings in other more direct ways, with the possible exception of *you*.

We can derive hypotheses about the meaning of such words by using them in sentences and judging the truth value of the sentences with respect to certain situations, and their ac-

ceptability in certain contexts. For example, we can unearth the basic uses of the definite article by presenting pictures containing one or two barking dogs, and ask to pick out the best picture for *the dog is barking*. The underlying idea is that the assignment of meanings to expressions is compositional, that is, that the meaning of the complex expression is a result of the meaning of its parts and the way they are combined.

## 3 Communicative Behavior

Perhaps the most important function of language is to communicate, that is, to transfer meanings from one mind to another. So we should be able to find evidence for meaning by investigating communicative acts. This is obvious in a trivial sense: If A tells B something, B will often act in certain ways that betray that B understood what A meant. More specifically, we can investigate particular aspects of communication and relate them to particular aspects of meaning. We will look at three examples here: Presuppositions, conversational implicatures and focus-induced alternatives.

Presuppositions (cf. article 102 *Presupposition*) are meaning components that are taken for granted, and hence appear to be downtoned. This shows up in possible communicative reactions. For example, consider the following dialogues:

A: *Unfortunately, it is raining.*

B: *No, it isn't.*

Here, B denies that it is raining; the meaning component of *unfortunate* expressing regret by the speaker is presupposed or conventionally implicated.

A: *It is unfortunate that it is raining.*

B: *No, it isn't.*

Here, B presupposes that it is raining, and states that this is unfortunate. In order to deny the presupposed part, other conversational reactions are necessary, like *But that's not unfortunate*, or *But it doesn't rain.* Simple and more elaborate denials are a fairly consistent test to distinguish between presupposed and proffered content (cf. van der Sandt 1988).

For conversational implicatures (cf. article 103 *Implicature*) the most distinctive property is that they are cancelable without leading to contradiction. For example, *John has three children* triggers the scalar implicature that John has exactly three children. But this meaning component can be explicitly suspended: *John has three children, if not more*. It can be explicitly cancelled: *John has three children, in fact he has four.* And it does not arise in particular contexts, e.g. in the context of *People get a tax reduction if they have three children.* This distinguishes conversational implicatures from presuppositions and semantic entailments: *John has three children,* {*if not two / in fact, two*} is judged contradictory.

Our last example concerns the introduction of alternatives that are indicated by focus, which in turn can be marked in various ways, e.g. by sentence accent. A typical procedure to investigate the role of focus is the question-answer test (cf. article 74 *Questions*). In the following four potential question-answer pairs (A1-B1) and (A2-B2) are well-formed, but (A1-B2) and (A2-B1) are odd.

A1: *Who ate the cake?*

A2: *What did Mary eat?*

B1: *MARY ate the cake.*

B2: *Mary ate the CAKE.*

This has been interpreted as saying that the alternatives of the answer have to correspond to the alternatives of the question.

To sum up, using communicative behavior as evidence for meaning consists in evaluating the appropriateness of certain conversational interactions. Competent speakers generally

agree on such judgments. The technique has been used in particular to identify, and differentiate, between different meaning components having to do with the presentation of meanings, in particular with information structure.

## 4 Behavioral Effects of Semantic Processing

When discussing evidence for the meaning of expressions we have focused so far on the meaning themselves. We can also investigate how semantic information is processed, and get a handle on how the human mind computes meanings. To get information on semantic processing, judgment tasks are often not helpful, and might even be deceiving. We need other types of evidence that arguably stand in a more direct relation to semantic processing. It is customary to distinguish between behavioral data on the one hand, and neurophysiologic data that directly investigates brain phenomena on the other. In this section we will focus on behavioral approaches (cf. also article 15 *Experimental Methods*).

### *4.1 Reaction times*

The judgment tasks for meanings described so far can also tap into the processing of semantic information if the timing of judgments is considered. The basic assumption is that longer reaction times, everything else being equal, are a sign for semantic processing load.

For example, Clark & Lucy (1975) have shown that indirect speech acts take longer for processing than direct ones, and attribute this to the additional inferences that they require. Noveck (2004) has shown that the computation of scalar implicature takes time; people that reacted to sentences like *Some elephants are mammals* with a denial (because all elephants and not just some do) took considerably longer. Kim (2008) has investigated the processing

of *only*-sentences, showing that the affirmative content is evaluated first, and the presupposition is taken into account only after.

Reaction times are relevant for many other psycholinguistic paradigms, beyond tasks like TVJ, and can provide hints for semantic processing. One notable example is the semantic phenomenon of coercion, changes of meanings that are triggered by the particular context in which meaning-bearing expressions occur (cf. article 25 *Mismatches and coercion*). One well-known example is aspectual coercion: Temporal adverbials of the type *until dawn* select for atelic verbal predicates, hence *The horse slept until dawn* is fine. But *The horse jumped until dawn* is acceptable as well, under an iterative interpretation of *jump* that is not reflected overtly. This adaptation of the basic meaning to fit the requirements of the context should be cognitively costly, and there is indeed evidence for the additional semantic processing involved. Piñango e.a. (2006) report on various studies and their own experiments that made use of the dual task interference paradigm: Subjects listen to sentences and, at particular points, deal with an unrelated written lexical decision task. They were significantly slower in deciding this task just after an expression that triggered coercion (e.g. *until* in the second example, as compared to the first). This can be taken as evidence for the cognitive effort involved in coercion; notice that there is not syntactic difference between the sentences to which such reaction time difference could be attributed.

### *4.2 Reading process: Self-paced reading and eye tracking*

Another window into semantic processing is the observation of the reading process. There are two techniques that have been used: (i) Self-paced reading, where subjects are presented with a text in a word-by-word or phrase-by-phrase fashion; the subject has control over the speed of presentation, which is recorded. (ii) Eye tracking, where the reading movements of the subject are recorded by cameras and matched with the text being read. While self-paced

reading is easier to handle as a research paradigm, it has the disadvantage that it might not give fine-grained data, as subjects tend to get into a rhythmical tapping habit.

Investigations of reading have been provided many insights into semantic processing; however, it should be kept in mind that by their nature they only help to investigate one particular aspect of language use that lacks many features of spoken language.

For example, reading speed has been used to determine how speakers deal with semantic ambiguity: Do they try to resolve it early on, which would mean that they slow down when reading triggers of ambiguity, or do they entertain an underspecified interpretation? Frazier & Rayner (1990) have shown that reading slows down after ambiguous words, as e.g. in *The records were carefully guarded* {*after they were scratched / after the political takeover*}, showing evidence for an early commitment for a particular reading. However, with polysemous words, no such slowing could be detected; an example is *Unfortunately the newspaper was destroyed,* {*lying in the rain / managing advertising so poorly*}.

The *newspaper* example is a case of coercion, which shows effects for semantic processing under the dual task paradigm (see discussion of Piñango e.a. 2006 above). Indeed, Pickering e.a. (2006) have shown that the aspectual coercion cases do not result in increased reading times; thus different kinds of tests seem to differ in their sensitivity.

Another area for which reading behavior has been investigated is the time course of pronoun resolution: Are pronouns resolved as early as possible, at the place where they occur, or is the semantic processor procrastinating this decision? According to Ehrlich & Rayner (1983), the latter is the case. They manipulated the distance between an antecedent and its pronoun and showed that distance had an effect on reading times, but only well after the pronoun itself was encountered.

### *4.3 Preferential Looking and the Visual World Paradigm*

Visual gaze and eye movement can be used in other ways as windows to meaning and semantic processing.

One technique to investigate language understanding is the preferential looking paradigm, a version of the picture selection task that can be administered to young infants. Preferential looking has been used for the investigation of stimulus discrimination, as infants look at new stimuli longer than at stimuli that they are already accustomed to. For the investigation of semantic abilities, so-called "Intermodal Preferential Looking" is used: Infants hear an expression and are presented at the same time with two pictures or movie scenes side by side; they preferentially look at the one that fits the description best. Hirsh-Pasek & Golinkoff (1996) have used this technique to investigate the understanding of sentences by young children that produce only single-word utterances.

A second procedure that uses eye gaze is known as "Visual World Paradigm". The general setup is as follows: Subjects are presented with a scene and a sentence or text, and have to judge whether the sentence is true with respect to the scene. In order to perform this verification, subjects have to glance at particular aspects of the scene, which yields clues about the way how the sentence is verified or falsified, that is, how it is semantically processed.

In an early study, Eberhard e.a. (1995) have shown that eye gaze tracks semantic interpretation quite closely. Listeners use information on a word-by-word basis to reduce the set of possible visual referents to the intended one. For example, when instructed to *Touch the starred yellow square*, subjects were quick to look at the target in the left-hand situation, slower in the middle situation, and slowest in the right-hand situation. Sedivy e.a. (1999) have shown that there are similar effects of incremental interpretation even with non-intersective adjectives, like *tall*.

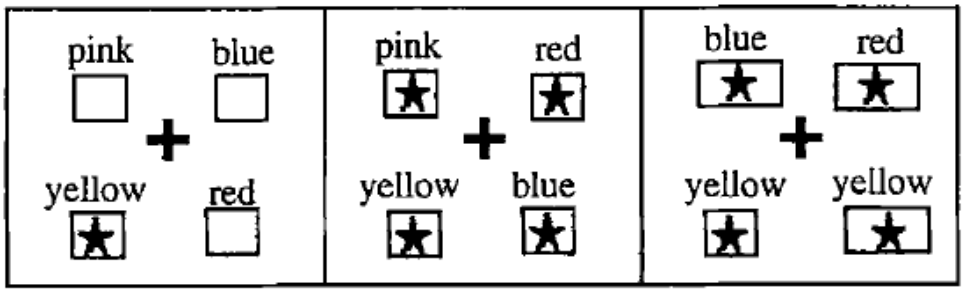


FIGURE 2: Stimulus of eye gaze test. From Eberhard e.a. (1995)

Altman & Kamide (1999) have shown that eye gaze is not just cotemporaneous with interpretation, but may jump ahead; subjects listening to *The boy will eat the…* looked preferentially at the picture of a cake than at the picture of something non-edible. In a number of studies, including Weber e.a. (2006), the effect of contrastive accent has been studied. When listeners had already fixated one object – say, the purple scissors – and now are asked to touch *the RED scissors* (where there is a competing red vase), they gaze at the red scissors more quickly, presumably because the square property is given. This effect is also present, though weaker, without contrastive accent, presumably because the use of modifying adjectives is inherently contrastive.

For another example of this technique, see article 15: *Experimental methods in semantics: An illustrative case of quantifier scope.*

## 5 Physiological effects of semantic processing

There is no clear-cut way to distinguishing physiological effects from behavioral effects. With the physiological phenomena discussed in this section it is evident that they are truly beyond conscious control, and thus may provide more immediate access to semantic processing.

Physiological evidence can be gained in a number of ways: From lesions of the brain and how they affect linguistic performance, from excitations of brain areas during surgery, from the observable metabolic processes related to brain activities, and from the electro-magnetic brain potentials that accompany the firing of bundles of neurons. There are other techniques that have been used occasionally, such as pupillary dilation, which correlates with cognitive load. For example, Krüger e.a. (2001) show with this measure that representations of event sequences following their natural order are cognitively less demanding than when not following the time line.

### *5.1 Brain Lesions and Stimulations*

Since the early discoveries of Broca and Wernicke, it has been assumed that specific brain lesions affect the relation between expressions to meanings. The classical picture of Broca's area responsible for production and Wernicke's area responsible for comprehension is now known to be incomplete (cf. Damasio e.a. 2004), but it is still assumed that Broca's aphasia impedes the ability to use complex syntactic forms to encode and also to decode meanings. From lesion studies it became clear that areas outside the classical Broca/Wernicke area and the connecting Geschwind area are relevant for language production and understanding. Brain regions have been identified where lesions lead to semantic dementia (also known as anomic aphasia) that selectively affects the recognition of names of persons, nouns for manipulable objects such as tools, or nouns of natural objects such as animals. These regions are typically situated in the left temporal lobe, but the studies reported by Damasio e.a. also indicate that regions of the right hemisphere play an important role.

It remains unclear, however, whether these lesions affect particular linguistic abilities or more general problems with the pre-linguistic categorization of objects. A serious problem with the use of brain lesions as source of evidence is that they are often not sufficiently lo-

cally constrained as to allow for specific inferences.

Stimulation techniques allow for more directed manipulations, and hence for more specific testing of hypothesis. There are deep stimulation techniques that can be applied during brain surgery. There is also a new technique, Transcranial Magnetic Stimulation (TMS), which affects the functioning of particular brain regions by electromagnetic fields applied from outside of the skull.

### *5.2 Brain Imaging of Metabolic Effects*

The last decades have seen a lively development of methods that help to locate brain activity by identifying correlated metabolic effects. Neuronal activity in certain brain regions stimulate the flow of oxygen-rich blood, which in turn can be localized by various means. While early methods like PET (Positron-Electron Tomography) required the use of radioactive markers, the method of fMRI (functional Magnetic-Resonance Imaging) is less invasive; it is based on measuring the electromagnetic fields of water molecules excited by strong magnetic fields. A more recent method, NIRS (Near Infrared Spectroscopy), applies low-frequency laser light from outside the skull; it is currently the least invasive technique. All the procedures mentioned have a low temporal resolution, as metabolic changes are slow, within the range of a second or so. However, their spatial resolution is quite acute, especially for fMRI using strong magnetic fields.

Results of metabolic brain-image techniques often support and refine findings derived from brain lesions (cf. Damasio e.a. 2004). As an example of a recent study, Tyler e.a. (2008) challenge the view that nouns and verbs are represented in different brain regions; they rather argue that inflected nouns and verbs and minimal noun phrases and minimal verb phrases, that is, specific syntactic uses of nouns and verbs, are spatially differentiated. An ongoing discussion is how general the findings about localizations of brain activities are, given the

enormous plasticity of the brain.

### *5.3 Event-Related Potentials*

This family of procedures investigates the electromagnetic fields generated by the cortical activity. They are observed by sensors placed on the scalp that either track minute variations of the electric field (EEG) or the magnetic field (MEG). The limitations of this technique are that only fields generated by the neocortex directly under the cranium can be detected. As the neocortex is deeply folded, this applies only to a small part of it. Furthermore, the number of electrodes that can be applied on the scalp is limited (typically 16 to 64, sometimes up to 256), hence the spatial resolution is weak even for the accessible parts of the cortex. Spatial resolution is better for MEG, but the required techniques are considerably more complex and expensive. On the positive side, the temporal resolution of the technique is very high, as it does not measure slow metabolic effects of brain activity, but the electric fields generated by the neurons themselves (more specifically, the action potentials that cause neurotransmitter release at the synapses). EEG electrodes can record these fields if generated by a large number of neurons in the pyramidal bundles of neurons in which the cortex is organized, in the magnitude of at least 1000 neurons.

ERP has been used for thirty years in psycholinguistic research, and specifically for semantic processing since the discovery by Kutas & Hillyard (1980) of a specific brain potential, the N400. This is a frequently observed change in the potential leading to higher negativity roughly 400ms after the onset of a relevant stimulus. See Kutas e.a. (2006) for a review of the vast literature, and Lau e.a. (2008) for a partially critical view of standard interpretations.

The N400 effect is seen when subjects are presented in an incremental way with sentences like *I like my coffee with cream and* {*sugar / socks*}, and the EEG signals of the first and the second variant is compared. In the second variant, with a semantically incongruous word, a

negativity around 400ms after the onset of the anomalous word (here: *socks*) appears when the brain potential development is averaged over a number of trials.

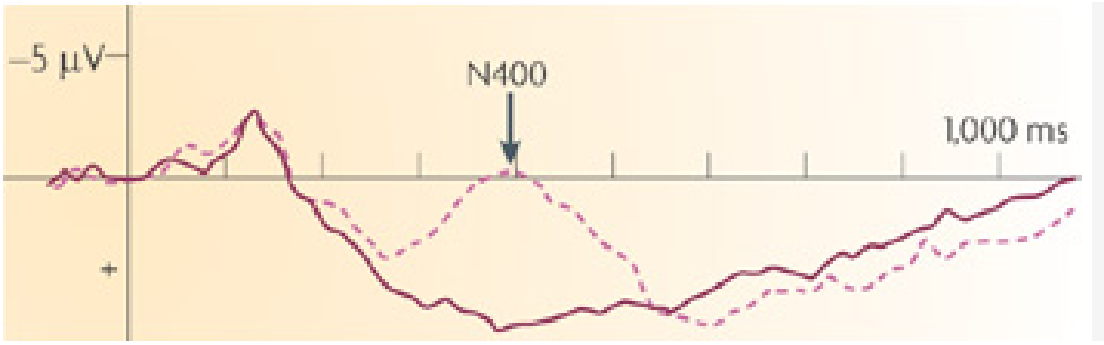


FIGURE 3: Averaged EEG over sentences with no semantic violation (solid line) and with semantic violation (dotted line); vertical axis at the onset of the anomalous word. (from Lau e.a. 2008).

There are at least two interpretations of the N400 effect: Most researchers see it as a reflex of the attempt to integrate the meaning of a subexpression into the meaning of the larger expression, as constructed so far. With incongruous words, this task is hard or even fails, which is reflected by a stronger N400. The alternative view is that the N400 reflects the effort of lexical access. This is facilitated when the word is predictable by the context, but also when the word is frequent in general. There is evidence that highly frequent words lead to a smaller N400 effect. Also, N400 can be triggered by simple word priming tasks; e.g. in *coffee* – {*tea* / *chair*}, the non-primed word *chair* leads to an N400. See Lau e.a. (2008) for consequences of the integration view and the lexical access view of the N400.

The spatial location of the N400 is also a matter of dispute. While Kutas e.a. (2006) claim that its origins are in the left temporal lobe and hence can be related to established language areas, the main electromagnetic field can be observed rather in the centroparietal region, and often on the right hemisphere. Lau e.a. (2008) discuss various possible interpretations of

these findings.

There are a number of other reproducible electrophysiological effects that point at additional aspects of language processing. In particular, Early Left Anterior Negativity (ELAN) has been implicated in phrase structure violations (150ms), Left Anterior Negativity (LAN) appears with morphosyntactic agreement violations (300-500ms), and P600, a positivity after 600ms, has been seen as evidence for difficulties of syntactic integration, perhaps as evidence for attempts at syntactic restructuring. It is being discussed how specific N400 is for semantics; while it is triggered by phenomena that are clearly related to the meaning aspects of language, it can be also found when subjects perform certain non-linguistic tasks, as in melody recognition. Interestingly, N400 can be masked by syntactic inappropriateness, as Hahne & Friederici (2002) have shown. This can be explained by the plausible assumption that structures first have to make syntactic sense before semantic integration can even start to take place.

There are a number of interesting specific finding around N400 or related brain potentials (cf. Kutas e.a. (2006) for an overview). Closed-class words generally trigger smaller N400 effects than open-class words, and the shape of their negativity is different as well – it is more drawn out up to about 700ms. As already mentioned, low-frequency words trigger greater N400 effects, which may be seen as a point in favor for the lexical access theory; however, we can also assume that low frequency is a general factor that impedes semantic integration. It has been observed that N400 is greater for inappropriate concrete nouns than for inappropriate abstract nouns. With auditory presentations of linguistic structures, it was surprising to learn that N400 effects can appear already before the end of the triggering word; this is evidence that word recognition and semantic integration sets in very early, after the first phonemes of a word.

The larger context of an expression can modulate the N400 effect, that is, the preceding

text of a sentence can determine whether a particular word fits and is easy to integrate, or does not fit and leads to integration problems. For example, in a context in which piercing was mentioned, *earring* triggers a smaller N400 than *necklace*. This has been seen as evidence that semantic integration does not differentiate between lexical access, the local syntactic fit and the more global semantic plausibility; rather, all factors play a role at roughly the same time.

N400 has been used as evidence for semantic features. For example, in the triple *The pizza was too hot to* {*eat / drink / kill*}, the item *drink* elicits a smaller N400 than *kill*, which can be interpreted as showing that the expected item *eat* and the test item *drink* have semantic features in common (ingestion), in contrast to *eat* and *kill*.

Brain potentials have also been used to investigate the semantic processing of negative polarity items (cf. article 71 *Polarity items*). Saddy e.a. (2004) and Drenhaus e.a. (2006) observe that negative polarity items in inappropriate contexts trigger an N400 effect (as in {*A / no*} *man was ever happy*}. With NPIs and with positive polarity items, a P600 could be observed as well, which is indicative for an attempt to achieve a syntactic structure in which there is a suitable licensing operator in the right syntactic configuration. Incidentally, these findings favor the semantic integration view of the N400 over the lexical access view.

There are text types that require special efforts for semantic integration – riddles and jokes. With jokes based on the reinterpretation of words, it has been found that better comprehenders of jokes show a slightly higher N400 effect on critical words, and a larger P600 effect for overall integration. Additional effort for semantic integration has also been shown for metaphorical interpretations.

A negativity around 320ms has been identified by Fischler e.a. (1985) for statements known to the subjects to be false, even if they were not asked to judge the truth value. But semantic anomaly clearly overrides false statements; as Kounios & Holcomb (1992) have

showed, in the examples like *No dogs are* {*animals / fruits*}, the latter triggers an N400 effect.

More recent experiments using MEG have discovered a brain potential called AMF (Anterior Midline Field) situated in the frontal lobe, an area that is not normally implied in language understanding. The effect shows up with coercion phenomena (cf. article 25 *Mismatches and coercion*). Coercion does not lead to an N400 effect; there is no anomaly with *John began the book* (which has to be coerced to *read* or *write* the book). But Pylkkänen & McElree (2007) found an AMF effect about 350ms after onset. This effect is absent with semantically incongruous words, as well with words that do not require coercion. Interestingly, the same brain area has been implied for the understanding of sarcastic and ironic language in lesion studies (Shamay-Tsoory e.a. 2005).

## 6 Corpus-linguistic Methods

Linguistic corpora, the record of past linguistic production, is a valuable source of evidence for linguistic phenomena in general, and in case of extinct languages, the only kind of source (cf. article 124 *Corpus linguistics*). This includes the study of semantic phenomena. For the case of extinct languages we would like to mention, in particular, the task of deciphering, which consists in finding a mapping between expressions and meanings.

Linguistic corpora can provide for evidence of meaning in many different ways. An important philosophical research tradition is hermeneutics, originally the art of understanding of sacred texts. Perhaps the most important concept in the modern hermeneutic tradition is the explication of the so-called hermeneutic circle (cf. Gadamer 1960): The interpreter necessarily approaches the text with a certain kind of knowledge that is necessary for an initial understanding, but the understanding of the text in a first reading will influence and deepen the understanding in subsequent readings.

With large corpora that are available electronically, new statistical techniques have been developed that can tap into aspects of meaning that might otherwise be difficult to recognize. In linguistic corpora, the analysis of word co-occurrences and in particular collocations can yield evidence for meaning relations between words.

For example, large corpora have been investigated for verb-NP collocations using the so-called Expectation Maximation (EM) algorithm (Rooth e.a. 1999). This algorithms leads to the classification of verbs and nouns into clusters such that verbs of class X frequently occur with nouns of class Y. The initial part of one such cluster, developed from the British National Corpus, looks as in the following table. The verbs can be characterized as verbs that involve scalar changes, and the nouns as denoting entities that can move along such scales.

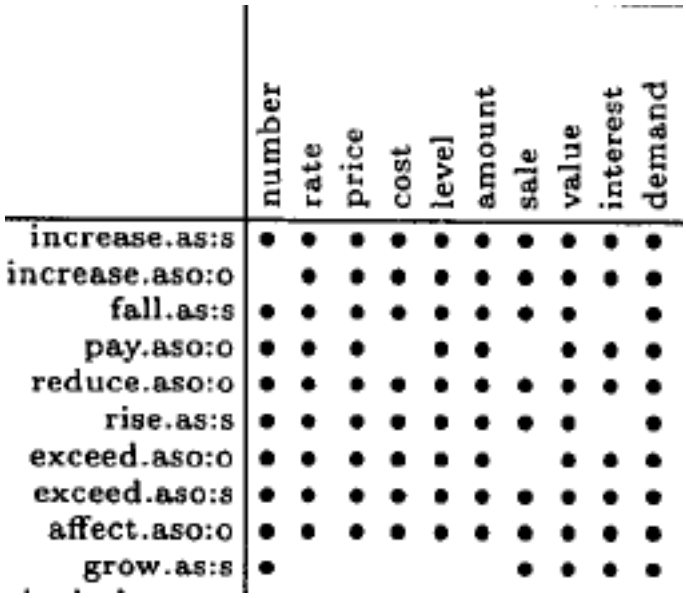


FIGURE 4. Clustering analysis of nouns and verbs; dots represent pairs that occur in the corpus. “as:s” stands for subjects of intransitive verbs, “aso:s” and “aso:o” for subjects and objects of transitive verbs, respectively. From Rooth e.a. (1999)

We can also look at the frequency of particular collocations within this cluster, as illustrated in the following table for the verb *increase*.

| *increase* 17 | 0.923698 | | |
|---|---|---|---|
| number | 134.147 | proportion | 23.8699 |
| demand | 30.7322 | size | 22.8108 |
| pressure | 30.5844 | rate | 20.9593 |
| temperature | 25.9691 | level | 20.7651 |
| cost | 23.9431 | price | 17.9996 |

FIGURE 5: Frequency of nouns occurring with *increase*. From Rooth e.a. (1999).

While pure statistical approaches as Rooth e.a. (1999) are of considerable interest, most applications of large-scale corpus-based research are based on a mix between hand-coding and automated procedures. The best-known project that has turned into an important application is WordNet (Fellbaum (ed.) 1998). A good example for the mixed procedure is Gildea & Juravsky (2002), a project that attempted semi-automatic assignment of thematic roles. In a first step, thematic roles were hand-coded for a large number of verbs, where a large corpus provided for a wide variety of examples. These initial examples, together with the coding, were used to train an automatic syntactic parser, which then was able to assign thematic roles to new instances of known predicates and even to new, unseen predicates with reasonable accuracy.

Yet another application of corpus-linguistic methods involves parallel corpora, collections of texts and their translations into one or more other languages. It is presupposed that the meanings of the texts are reasonably similar (but recall the problems with translations mentioned above). Refined statistical methods can be used to train automatic translation devices on a certain corpus, which then can be extended to new texts that then are translated automatically, a method known as example based machine translation.

For linguistic research, parallel corpora have been used in other ways as well. If a language $\alpha$ marks a certain distinction overtly and regularly, whereas language $\beta$ marks that distinction only rarely and in irregular ways, good translations pairs of texts from $\alpha$ into $\beta$

can be used to investigate the ways and frequency in which the distinction in β is marked. This method is used, for example, in von Heusinger (2002) for specificity, using Umberto Eco's *Il nome della rosa*, and Behrens (2005) for genericity, using Sait-Exupéry's *Le petit prince*. The articles in Cysouw & Wälchli (2007) discuss the potential of the technique, and its problems, for typological research.

## 7 Conclusion

This article, hopefully, has shown that the elusive concept of meaning has many reflexes that we can observe, and that semantics actually stands on as firm grounds as other disciplines of linguistics. The kinds of evidence for semantic phenomena are very diverse, and not always as convergent as semanticists might wish them to be. But they provide for a very rich and interconnected area of study that has shown considerable development since the first edition of the *Handbook Semantics* in 1991. In particular, a wide variety of experimental evidence has been adduced to argue for processing of meaning. It is to be hoped that the next edition will show an even richer and, hopefully, more convergent picture.

References

Altmann, Gerry & Yuki Kamide 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73, 247-264.

Behrens, Leila. 2005. Genericity From a Cross-Linguistic Perspective. *Linguistics* 43, 257-344.

Berlin, Brent, Dennis E. Breedlove & Peter H. Raven 1973. General principles of classifi-

cation and nomenclature in folk biology. *American Anthropologist* 75, 214-242.

Berlin, Brent & Paul Kay 1969. *Basic color terms. Their universality and evolution***.** University of California Press, Berkeley.

Bloomfield, Leonard. *Language.* New York, 1933.

Chafe, Wallace (ed.) 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, Ablex, Norwood, N.J.

Clark, Herbert H. & Peter Lucy 1975. Understanding what is meant from what is said: A study in conversationally conveyed requests, *Journal of Verbal Learning and Verbal Behavior* 14.

Crain, Stephen 1991. Language Acquisition in the Absence of Experience. *Behavioral and Brain Sciences* 14. 597-650.

Crain, Stephen, Rosalind Thornton, Laura Conway & Diane Lillo-Martin 1998, Quantification without quantification. *Language Acquisition* 5, 83-153.

Cysouw, Michael & Bernhard Wälchli 2007. Parallel texts: Using translational equivalents in linguistic typology (Theme issue) *Sprachtypologie und Universalienforschung* 60, 95-99.

Damasio, H., D. Tranel, T. Grabowski, R. Adolphs, A. Damasio. 2004. Neural systems behind word and concept retrieval. *Cognition* 92. 179-229.

Drenhaus, Heiner, Peter beim Graben, Doug Saddy & Stefan Frisch 2006. Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language* 96, 255-268.

Eberhard, Kathleen, Michael J. Spivey-Knowlton, Judy Sedivy & Michael Tanenhaus 1995. Eye movement as a window into real -time spoken language comprehension in a natural context, *Journal of Psycholinguistic Research* 24m 409-436.

Ehrlich, Kate & Keith Rayner 1983. Pronoun assignment and semantic integration during reading: Eye movement and immediacy of processing, *Journal of Verbal Learning and Ver-*

*bal Behavior* 22, 75-87.

Fellbaum, Christiane (ed.) 1998. *WordNet: An electronic lexical database***,** MIT Press, Cambridge, Mass.

Fischler, Ira, Donald G. Childers, Achariyapaopan, Teera & Nathan W. Perry 1985. Brain potentials during sentence verification: Automatic aspects of comprehension. *Biological Psychology* 21, 83-105.

Frazier, Lyn & Keith Rayner 1990. Taking on semantic commitments: Processing multiple meanings vs. multiple senses, *Journal of Memory and Language* 29, 181-200.

Frege, Gottlob 1892. Über Sinn und Bedeutung, *Zeitschrift für Philosophie und philosophische Kritik, NF* 100, 25-50.

Gadamer, Hans Georg 1960. *Wahrheit und Methode. Grundzüge einer philosophischen Hermeneutik***,** Tübingen.

Gärdenfors, Peter 2000. *Conceptual Spaces - The Geometry of Thought***,** The MIT Press, Cambridge, Mass., London.

Gerken, LouAnn & Michele E. Shady 1998. The picture selection task. In: Dana McDaniel, Cecile McKdee & Helen Smith Cairns (eds), *Methods for Assessing Children's Syntax*, Cambridge, Mass, MIT Press, 125-146.

Gildea, Daniel & Daniel Jurafsky 2002. Automatic labeling of semantic roles, *Computational Linguistics* 28, 245-288.

Goddard, Cliff 1998. *Semantic analysis. A practical introduction***,** Oxford University Press, Oxford.

Goodman, Nelson 1955. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard UP,

Gordon, Peter 1998. The truth-value judgment task. In: Cecil McKee & Helen Smith Carins, *Methods for Assessing Children's Syntax*, Cambridge, MIT Press, 211-235.

Gregory, Robert J. 2003. An early history of land on Tanna, Vanuatu. *The Anthropologist*

5, 67-74.

Grice, H. Paul 1975. *Logic and conversation*. In Peter Cole & Jerry L. Morgan (eds.) *Syntax and Semantics 3: Speech Acts*, New York, Academic Press, 41-58.

Grice, Paul 1957. Meaning. *The Philosophical Review* 66, 377-388.

Hahne, Anja & Angela D. Friederici 2002. Differential task effectson semantic and syntactic processes are revealed by ERPs, *Cognitive Brain Research.* 13, 339-356.

Haviland, John B. 1974. A last look at Cook's Guugu-Yimidhirr wordlist, *Oceania* 44, 216-232.

Hill, C. 1982. Up/down, front/back, left/right: A contrastive study of Hausa and English. In Jürgen Weissenborn & Wolfgang Klein, *Here and there: Cross-linguistic Studies on Deixis and Demonstration*, Amsterdam, Benjamins, 11-42.

Hirsh-Pasek, Kathryn & Roberta M. Golinkoff 1996. *The Origin of Grammar: Evidence from Early Language Comprehension***,** MIT Press, Cambridge, Mass.

Kay, Paul, Brent Berlin, Luisa Maffi & Wiliam R. Merrifield 2008. *The World Color Survey*. Center for the Study of Language and Information, Stanford.

Kempton, William. 1981. *The Folk Classification of Ceramics.* New York: Academic Press.

Kim, Christina 2008. Processing presupposition: Verifying sentences with only. *Proceedings of the 31 st Penn Linguistics Colloquium*, University of Pennsylvania.

Kounios, John, and Phillip J. Holcomb 1992. Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General* 121, 459-79.

Kranstedt, Alfred, Andy Lücking, Thies Pfeiffer, Hannes Rieser 2006. Deixis: How to determine demonstrated objects using a pointing cone. *Lecture Notes in Computer Science* 3881, Berlin-Heidelberg: Springer, 300-311.

Krüger, Frank, Antje Nuthmann & Elke Van Der Meer 2001. Pupillometric indices of the temporal order representation in semantic memory. *Zeitschrift für Psychologie* 209, 402-415.

Kutas, Marta & Steven A. Hillyard 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207, 203-205.

Kutas, Marta, Cyma K. Van Petten & Robert Kluender 2006. Psycholinguistics Electrified II 1994-2005, in M.A. Gernsbacher & M. Traxler (eds), *Handbook of Psycholinguistics*, New York, Elsevier, 83-143.

Kutas, Marta, Cyma K. Van Petten & Robert Kluender 2006. Psycholinguistics Electrified II 1994-2005), in M.A. Gernsbacher & M. Traxler, *Handbook of psycholinguistics*, New York, Eslevier.

Lakoff, George, Johnson, Mark 1980. *Metaphors we live by***,** University of Chicago Press, Chicago.

Lau, Ellen F., Colin Phillips & David Poeppel 2008. A cortical network for semantics: (de)constructing the N400. *Nature reviews neuroscience* 9, 920-933.

McDaniel, Dena, Cecile McKee & Helen Smith Cairns (eds) 1996. *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press.

Majid, A. et al. 2006. The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18, 133-152.

Matthewson, Lisa 2004. On the methodology of semantic fieldwork. *IJAL* 70, 369-415.

Noveck, Ira 2001. When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition* 78, 165-188.

Noveck, Ira 2004. Pragmatic inferences linked to logical terms. In: Ira Noveck & Dan Sperber (eds.) *Experimental pragmatics*, Basingstoke, Palgrave Macmillan, 301-321.

Pickering, Martin J., Brian Mcelree & Steven Frisson 2006. Underspecification and aspectual coercion. *Discourse Processes* 42, 131-155.

Piñango, Maria Mercedes, Aaron Winnick, Rashad Ullah & Edgar Zurif. 2006. Time course of semantic composition: The case of aspectual coercion. *Journal of Psycholinguistic Research* 35, 233-244.

Portner, Paul & Katsuhiko Yabushita 1998. The semantics and pragmatics of topic phrases. *Linguistics and Philosophy* 21, 117-157.

Potts, Christopher 2005. *The logic of conventional implicature***.** Oxford University Press, Oxford.

Pylkkänen, Liina & Brian McElree 2007. An MEG study of silent meaning, *Journal of Cognitive Neuroscience* 19, 1905-1921.

Quine, W. V. O. 1960. Translation and meaning., in *Word and object*, Cambridge, Mass, MIT Press, 26-79.

Rayner, Keith 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124, 372-422.

Rooth, Mats, Stefan Riezler, Detlef Prescher & Glenn Carroll 1999. Inducing a semantically annotated lexicon via EM-based clustering, *37th Annual Meeting of the ACL,* 104-111.

Saddy, Douglas, Heiner Drenhaus & Stefan Frisch 2004. Processing polarity items: Contrastive licensing costs. *Brain and Language* 90, 495-502.

Sedivy, Julie C., Michael K. Tanenhaus, Craig C. Chambers & Gregory N. Carlson 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition* 71, 109-147.

Shamay-Tsoory, S.G., R. Tomer & J. Aharon-Peretz 2005. The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology* 19, 288-300.

Smith, Carlotta L. 1980. Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, 30, 191-205.

Stalnaker, Robert 1974. Pragmatic presuppositions. In: Milton K. Munitz & Peter K.

Unger, *Semantics and Philosophy*. New York, New York University Press, 197-214.

Thornton, Rosalind 1998. Elicited production. In: Dana McDaniel, Cecil McKee & Helen Smith Carins, *Methods for assessing children's syntax*, Cambridge,Mass, MIT Press, 77-95.

Tomasello, Michael 2008. Why don't apes point?, in Regine Eckardt, Gerhard Jäger & Tonjes Veenstra, *Variation, selection, development. Probing the evolutionary model of language change*, Berlin, Mouton de Gruyter, 375-394.

Tyler, L. K., B. Randall & E.A. Stamatakis 2008. Cortical differentiation for nouns and verbs depends on grammatical markers, *Journal of Cognitive Neuroscience* 20, 1381-1389.

van der Sandt, Rob A. 1988. *Context and presupposition***,** Croom Helm, London.

Von Heusinger, Klaus 2002. Specificity and definiteness in sentence and discourse structure, *Journal of Semantics* 19, 245-274.

Weber, Andrea, Bettina Braun & Matthew W. Crocker 2006. Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech* 49, 367-392.

Weinreich, Uriel 1964. Webster's Third. A Critique of Its Semantics. *International Journal of American Linguistics* 30: 405-09.

*Manfred Krifka, Berlin Germany)*

Keywords

meaning, psycholinguistics, communication