Meaning Potentials in Words and Gestures

Jens Allwood [1], *Elisabeth Ahlsén*[1]

[1] SCCIIL Interdisciplinary Center, University of Gothenburg
jens@ling.gu.se, eliza@ling.gu.se

Abstract

This paper addresses the question of what and how gestures and speech, respectively, contribute to the construction of meaning. A point of departure of is the notion of "meaning potential" which we apply to both unimodal gestures and unimodal vocal-verbal units, as well as to multimodal vocal-gestural units, [1]. The purpose of this paper is to explore the notion of "meaning potential", not only for speech, but also for gesture. Specifically, we want to discuss the possibilities of extending the notion of a meaning potential for a symbolic sign (e.g. a word) to iconic and indexical signs.

Index Terms: speech recognition, human-computer interaction, computational paralinguistics

1. Purpose

The purpose of this paper is to explore the notion of "meaning potential", not only for speech, but also for gesture. Specifically, we want to discuss the possibilities of extending the notion of a meaning potential for a symbolic sign (e.g. a word) to iconic and indexical signs. The reason for this is that the non-verbal gestures accompanying speech (co-verbal gestures) are iconic and indexical. So if the notion of meaning potential can be used also in relation to such gestures, a significant step in providing an account of meaning in multimodal communication will have been taken.

The theory we will present will thus be part of a sketch of some of the steps towards a cognitive semiotic theory of the semantics/pragmatics of multimodal communication

2. Background and Points of departure

Below we will briefly introduce some notions our analysis is based on, namely: the notions of communication and multimodal communication and the Peircean three modes of activating information (index, icon ad symbol) [2].

2.1. Communication and Multimodal communication

The notion of communication we will be presupposing is a notion where communication is seen as productive activation and receptive co-activation of shared content (information/understanding), while drawing on contextual resources.

In multimodal face-to-face communication, this means that a speaker produces speech and gestures to be shared with a listener in a process involving co-activation of the produced content. It is important to note that there is a mutual flow of information between speaker and listeners, so that a speaker not only speaks and gestures, but also perceives and understands his own communication as well as simultaneous words and gestures from co-communicators. Similarly, a listener not only perceives and understands, but also behaviorally reacts, for example, by verbal and gestural feedback.

Both speakers and listeners make use of the context in which the communication is embedded in order to produce and interpret the content that is being shared.

In face-to-face interaction, communication is multimodal, in the sense that it involves activation through more than one of the sensory modalities (hearing, vision, touch, smell and taste). When it comes to hearing and speech, both segmental and suprasegmental (prosodic) features of speech are involved. Gestures are involved both in touch and in visually shared information.

2.2. Modes of activation and representation

Besides multimodal communication, another important presupposition for our discussion, are the three basic semiotic means of information activation (representation) suggested by Charles Sanders Peirce; symbol, icon and index.

Index: Indexes involve activation or representation by making use of contiguity in space and time.

Icon: Icons involve activation (or representation) by making use of similarity.

Symbol: Symbols involve activation (or representation) by making use of conventional associations.

All these three modes of activation and representation are used in simultaneous and consecutive combination with each other, in both cognition and face-to-face communication. The symbols used are mostly vocal verbal expressions, while the icons and indexes are mostly gestures.

2.3. Theories of semantics for symbols, icons and indices

Theories of semantics have almost exclusively been concerned with written or vocal verbal symbols (words and combinations of words). Other types of symbols, icons and indexes have rarely been considered. Some of the most common semantic theories for written and vocal words (morphemes, phrases, sentence) are:

1. Truth conditional semantics (applies primarily to sentences)

2. Common meanings, in the form of necessary and sufficient conditions (primarily applies to words and morphemes)

3. Basic meanings, in the form of basic exemplars or prototypes (primarily applies to words and morphemes) [3]

4. Meaning Potentials, in the sense of the potentially activizable information connected with verbal symbols – here the point of departure is the collection of all of a word's uses in individual and collective memory [4], [1] For collective memory, see for example the linguistic division of labor discussed by Putnam in [5]).

The primary question we want to address in this paper is the question of whether meaning potentials can be extended from symbols to icons and indexes. Can we, drawing on memory and perception, in analogy with symbolic meaning potentials, also assume that there are iconic meaning potentials and indexical meaning potentials?

A Meaning Potential (in the sense we take it here, which is different from the sense it is used in for example, [6], [7],), is a structured collection of uses of a symbol (word), that is relevant both for understanding and production in communication.

Meaning potentials, thus, provide an analysis of linguistic meaning in line with the suggestion made by Wittgenstein [8] of seeing the meaning of a word as the set of uses of the word. But it is also in harmony with Vygotsky's suggestion that children learn language by learning linguistic labels (pseudo concepts) that are filled with content, as they successively learn to use these labels in different contexts [9]. The collection of uses forms the meaning potential of a word. Part of our linguistic competence is learning to activate (or actualize) this potential as triggered by different contextual features, such as the collocations (other words and morphemes), that a particular word is (often) combined with or the social activities in which the word are used.

The collection of uses (meaning potential) as stored in memory can then become a basis for a polysemy structure (analogous to what one might find in a dictionary) that is upheld by association with relevant contextual features like collocations and social activities. The meaning potential can also become the basis for cognitive processing, which can produce prototypes (typical examples), where they are relevant or necessary and sufficient conditions, where they are relevant, or both of these, when that is relevant. The cognitive processing is guided by cognitive operations supporting discrimination (analysis) and combination (synthesis), compression and abstraction of content, including such processes as contiguity abstraction and similarity abstraction and refinements of these. When the cognitive operations become associated with linguistic markers, we will refer to them as semantic-epistemic operations. Se also [4] and [1].

Let us now turn from the meaning potential of words and other symbols to a consideration of the role of meaning potentials for iconic and indexical gestures accompanying speech. As we shall see such gestures are often used to specify, highlight or illustrate features of the activated verbal (symbolic) meaning potential based content. Briefly, what happens is that in addition to the content activated by the words, the gestures activate additional content dependent on similarity (icons) and/or continguity relations (indices).

For example, iconic gestures might add illustrative pantomimes or metaphorical content and indexical gestures might add pointing to specific concrete or abstract locations or metonymic content.

In production, the gestures indicate, display or signal relevant information by making use of similarity and contiguity, often related to the verbal content. Similarly, in understanding, we interpret relevant information by similarity or contiguity, often in relation to the vocal verbal symbolic content.

2.4. Communication in context

Both of these processes – production and understanding – involve use of context as a resource for activation and contextual adaptation, accommodation, actualization and determination of content.

In fact, communication is always dependent on context for content, behavior action and type of interaction. The status and functions of new contributions are continuously being shaped by dimensions of context , such as:

- the physical environment
- the culture, the language, the current organizational setting
- the current social activity/activities
- the activity roles of the communicators
- the various traits of the communicators; gender, age and other psychological, social and biological properties
- the current contribution (compositionality)
- the currently preceding and/or simultaneous contributions (co-construction)
- other informative actions and behavior by the communicators
- the currently activated but also the potential shared background of the communicators (their "common ground")

The dimensions and features of context mentioned in this list form a background for pointing to two basic types of contextual determination of content

(i) Compositionality (combinability) in a wide sense

What we have in mind here is the contextual determination of the content of a multimodal contribution by drawing on the combined activation of several or all communicative features of the units (combining words and gestures) occurring in the same contribution. This is the issue we are discussing in this paper.

(ii) Co-construction

Here we move our contextual window from the content of a contribution of a single communicator to contextual determination of content, by drawing on the combined activation of several communicative contributions (mostly consecutive) from different communicators. This issue we will return to in future work.

2.5. Levels of awareness and intentionality

Our analysis also takes into account the fact that communication takes place on several simultaneous levels of awareness and intentionality. To facilitate analysis, we distinguish the following three levels on what basically is a continuous scale (cf. [10]).

Indicate (being informative)
Display (showing)
Signal (showing that you are showing)

These three levels can be combined with the three Peircean types of representation (index, icon, symbol) in the following manner, where all possibilities can occur but we have only indicated the most frequent cases in face-to–face communication.

Table 1. *Levels of awareness and intentionality and types of representation.*

	index	**icon**	**symbol**
indicate	Vocal segmental, Gesture, Prosody		
display	Vocal segmental, Gesture, Prosody	Vocal segmental, Gesture, Prosody	Vocal segmental, Gesture, Prosody
signal	Vocal segmental, Prosody, Gesture, Prosody		Vocal segmental, and prosodic verbal, Gestural verbal

The table shows how the three means of expressions (words, gesture and prosody) are typically related both to the three levels of awareness and intentionality and the 3 types of representation.

3. Meaning potentials, multi-representational and multimodal contributions

Using the background introduced above, we now want to discuss in what sense there can be meaning potentials not only for verbal symbols, but also for accompanying gestural indeces and gestural icons.

We want to do this by discussing what could be meant by these three types of meaning potentials, and then discussing what could be meant by combining them

3.1. The meaning potentials of symbols (words)

A meaning potential of a word can be organized into a polysemy compatible structure sustained by collocations related systematically to encyclopedic (including iconic and indexical) information. As an example, we present a sketch of the meaning potential of the word *tree* below

Tree: **Meaning potential: Polysemy + collocations:**
Source: http://oxforddictionaries.com/definition/english/tree [11]

noun

1a woody perennial plant, typically having a single stem or a trunk growing to a considerable height and bearing lateral branches at some distance from the ground. (in general use) any bush, shrub, or herbaceous plant with a tall erect stem, e.g. a banana plant.

2 a wooden structure or part of a structure.
archaic or literary the cross on which Christ was crucified.
archaic a gibbet.

3 a thing that has a branching structure resembling that of a tree.
(also tree diagram) a diagram with a structure of branching connecting lines, representing different processes and relationships.

verb (trees, treeing, treed) [with object]

1 North American force (a hunted animal) to take refuge in a tree.
informal, chiefly US force (someone) into a difficult situation.

2 as **adjective** treed (of an area) planted with trees sparsely treed grasslands

Collocations

- Decision tree
- Solution tree
- Tree diagram
- Elm tree
- Fruit tree
- Christmas tree

The meaning potential also includes and integrates encyclopedic meaning so no systematic distinction is made between lexical and encyclopedic meaning.

- Source Wikipedia – Encyclopedia [12]:
- In botany, a **tree** is a plant with an elongated stem, or trunk, supporting leaves or branches.
- In some usages, the definition of a tree may be narrower, including only woody plants, only plants that are usable as lumber, only plants above a specified height or only perennial species. At its broadest, trees include the taller palms, the tree ferns, bananas and bamboo.
- In its broadest sense, a tree is any plant with the general form of an elongated stem, or trunk, which supports the photosynthetic leaves or branches at some distance above the ground.[6][7] Trees are also typically defined by height,[8][9][10] with smaller plants being classified as shrubs,[11] however the minimum height which defines a tree varies widely, from 10 m to 0.5 m.[10] By these broadest definitions, large herbaceous plants such as papaya and bananas are trees, despite not being considered as trees under more rigorous definitions.[3][5][12][13][14][15]
- Another criterion often added to the definition of a tree is that it has a woody trunk.[10][16][17] Such a definition excludes herbaceous trees such as bananas and papayas. Monocots such as bamboo and palms may be considered trees under such a definition.[18] Despite being herbaceous[19][20] and not undergoing secondary growth and never producing wood,[21][22][22][23] palms and bamboo may produce "pseudo-wood" by lignifying cells produced through primary growth.
- Aside from structural definitions, trees are commonly defined by use. Trees may be defined as plants from which lumber can be produced.

Finally the meaning potential of a symbol can also include iconic and indexical information and contextual information, over and above that given by collocations.

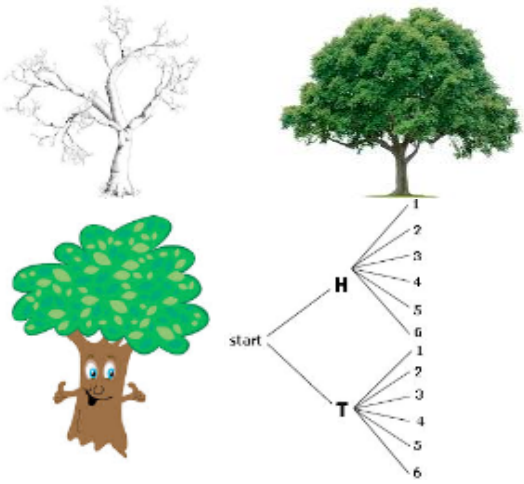

Figure 1: *Iconic information in meaning potential of tree.*

3.2. The meaning potentials of iconic gestures

The meaning potential of an icon relies on similarity, i.e. homomorphic-isomorphic relations that can be used both for production and understanding; activated by cognitive processes and semantic-epistemic operations triggering cognitive specification from memory or perception.

Let us first consider what might be the meaning potential of an icon without regard for context [13]. We explored this topic by asking a panel of judges to give interpretations of an iconic tree gesture, which consisted of: Both hands lifted in front of face, palms about 15 cm apart, turned towards each other, then hands coming apart and then together forming a circle, then both hands moving down in parallel.

Below are the interpretations of these iconic gestures as given by the panel of judges.

A. Showing a shape – possibly woman
B. A man or a person
C. Round a the top getting thinner – showing form
D. Tree
E. Showing the shape of something
F. Showing the form
G. Female earth mother (showing hip rouding)
H. "this shape"
I. quite
J. "symbolizing a woman/female body"
K. a tree
L. narrow it down

If we take these responses as indications of the meaning potential of the exhibited gesture, we can see that what seems to be going on is an activation of shapes from memory that are similar to the gesture.

As we can see, the meaning potential of a decontextualized iconic gesture, in general, seems more open and less structured than the meaning potential of a decontextualized verbal symbol. A circular movement of a hand or a finger can be similar to many things and we need either to add conventionalization or context, or both, to arrive at a more limited specific type of content. Deaf sign language has many examples of how iconic signs become conventionalized, that is, they combine iconic with symbolic representation and in this way can be used to activate more specific content.

For iconic signs that are less conventionalized than deaf sign language, context is needed to guide the users in what features, of the content being shared, are the relevant ones to focus on for the similarity based abstraction and activation of information. If the shared contextual information is not sufficient for this, there is a risk that the content activated will not be shared. This can clearly be seen in the variety of the responses presented above, where no shared context was provided.

In face-to-face multimodal communication, the most important context for iconic gestures is usually the content of the vocal verbal messages they co-occur with.

3.3. The meaning potentials of indexical gestures

To get an idea of the meaning potential of decontextualized indeces, let us consider two examples of indexical gestures: (i) a pointing index finger or (ii) a smile. As with iconic gestures, the meaning potential of these gestures, without conventionalization or context, will allow for a too large number of information activations.

What is being pointed to by an index finger is in no way easily restricted, since it could be pointing to both concrete and abstract entities. What is being expressed by a smile is more restricted, but in itself allows for many interpretations, like friendliness, shame, fear, ingratiation, happiness, contentment, malevolence etc.

As with iconic gestures, context is needed to determine what the contiguity relation activated by the gesture should apply to.

3.4. Multimodal combinations of symbols, icons and indices in face-to-face communication

(i) Symbol with symbol

The first combination to consider is perhaps the multimodal combination of a vocal verbal symbol with a gestural verbal symbol. Such combinations are common in giving communicative feedback in English, where, for example, the vocal verbal symbol *yeah* is often accompanied by a gestural verbal symbol, affirmative *head nod,* providing a multimodal combination of a vocal and a gestural symbol, both expressing assent and affirmation, the function of which is a reinforcement of the affirmation. The same would happen, in English, if the vocal verbal *no* is combined with a gestural verbal *head shake.*

(ii) Symbol with icon

Let us now see what happens if the meaning potential of an iconic gesture is combined with the meaning potential of a

word. Let us consider an example from a discussion about Nature.

Example 1:

B: he was there // with his senses and open to it just then // maybe sitting on his tractor //and
D: yes
B: he probably didn't then // but normally [it is probably (…)]
C: [(…) forerunners] with with modern tractors // with // air condition // radio // and headphones // machine panel
D: but // but surely // e // e // surely // there is someone who has a //quick // association // e ö with a // **fruit tree (+ iconic gesture)** blossoming // and who sits driving a tractor // and turns around

Note: (// = pause, [ ] = overlap, (…) = inaudible speech)

In this example, the iconic gesture triggers a cognitive similarity specification, operating on the actualized content of the word *tree* and other perceptual memories related to this word.
The gesture highlights the shape of a blossoming fruit tree and in doing this also emphasizes and specifies the tree and the shape of the tree.

(iii) Symbol with index

As with iconic gestures, the context of an indexical gesture will often be given by simultaneously produced vocal verbal content. For example

1. The house is over there, accompanied by *pointing* gesture

The gesture specifies direction to the location of the house by contiguity and the verbal element tells us what is to be located.

2. *I am happy to see you*, accompanied by a *smile*

Here the smile indicates an inner state of happiness expressed by the word *happy*.

In both cases, the gestures (the pointing finger and the smile), that could potentially have many other meanings, trigger an epistemic contiguity operation which further specifies the content of the vocal verbal symbol.

(iv) Symbol with both index and icon

Often, vocal verbal symbols are combined with both iconic and indexical gestures, which can occur either separately or as simultaneous features of one gesture. Let us consider some examples.

Example 2.

A: *I have tried to start to study also English, so that I wont* ***forget*** **(the word *forget* is combined with an indexical/iconic gesture of a circling index finger pointing to the temple of the head)**

The activated meaning potential of the word *forget* here provides the contextual content basis for the gesture, which in itself combines indexical and iconic features.

The indexical features (contiguity in space and time) of the gesture locates "the forgetting" in the head. Here perception of this gesture and memory interact in giving further associations to cognitive processes. The iconic features of the gesture, "a circling motion", simultaneously with the indication of the location, highlights a memory problem (circling – not finding).

To some extent, this particular association between the gesture used (the circling finger) and a cognitive problem has been conventionalized, which can be seen when we asked a panel of judges to give interpretations of circling index finger pointing to head and the majority all indicate a cognitive problem of some sort.[13].

Description of "forget" gesture:
Preparation: lifts hand up towards head
Stroke: index finger points to head, circular movement
Retraction: hand goes back to lap

Suggested meanings by panel of judges:

A. crazy (in the head) or confused /about self
B. "I am confused"
C. Don't understand – crazy/nuts
D. ?
E. I'm crazy /confused
F. You have a hole in your head, you know = are stupid /don't understand
G. I am confused!
(pointing to ear and circling to show confusion)
H. "this person is crazy"
I. thinking all the time
J. mind-bogging
K. "my head is going round" = "cocco
L. cannot remember, or cannot think sth up
M. Hopeless to talk to

If we consider the contextual influence on the meaning activation of the multimodal contribution, we see that we have in this case is a combination of

(i) activation of the symbolic meaning potential of the word *forget*, which is contextually determined by the fact that it occurs in the activity context of a discussion on learning and is a collocation of *not forget*.

(ii) a gesture providing an indexical highlighting of the locus of forgetting and an iconic highlighting of a dynamic circle, which can display some type of cognitive problem.

The combined multimodal effect of the gesture will be to highlight and emphasize the locus of forgetting in the head.

4. Summary and concluding discussion

We have seen that multimodal face-to-face communication not only involves a combination of information in many modalities, but also a combination of several modes of representation on several level of awareness and intentionality.

What we frequently have is a combination of vocal verbal symbols with gestural icons and indices. However, vocal verbal symbols can also be combined with gestural verbal symbols, as is the case in communicative feedback, where, in English, words like *yes* and *no* are combined with head nods and head shakes. The most common effect of a combination of a vocal verbal symbol with an iconic or indexical gesture is that features of the activated symbolic content are specified, highlighted or illustrated by making use of cognitive semantic-

epistemic operations relying on similarity (homomorphism) and/or contiguity relations. When vocal verbal symbols are combined with gestural verbal symbols, the effect is rather one of reinforcement and emphasis.

The framework we have proposed, thus, provides some steps towards a cognitive, holistic semiotic theory of the semantics/pragmatics of multimodal contributions to interactive communication. We have suggested that communication should be seen as involving productive activation and receptive coactivation of shared content (information/understanding), drawing on contextual resources determining the meaning potentials of symbols (words), but also of icons and indices, making use of cognitive operations combining activation of conventional information with activation making use of similarity and contiguity relations, helping to determine the symbolic content.

We have also seen that the meaning potentials of symbols, icons and indices vary in how specific and structured the information is that they enable activation of. Conventionalization, in general, seems to make the activizable information more specific and structured, so that decontextualized symbols have more fine grained meaning potentials than decontextualized icons and indices. It seems likely that the same difference might also hold true for icons versus indices. The homomorphism of icons puts more restrictions on what information can be activated than the contiguity of indices.

In all cases, activation of meaning potentials requires activation of contextual resources to acquire a more determinate actualized meaning. Since the meaning potentials of icons and indices are more open ended than the meaning potentials of symbols, this need is stronger for icons and indices than for symbols. In this way, iconic and indexical gestures have a natural functional role to play as coverbal reinforcers and specifiers of features of content in activated symbolic verbal meaning potentials.

Thus, our analysis provides a basis for a rethinking not only of the "compositionality" of symbols (symbol + symbol) in terms of a combination of symbolic meaning potentials under contextual enablements and restrictions, but also for considering the combination of symbols (words) with icon and index (usually gestures) under contextual enablements and restrictions.

Finally, we have noted that meaning potentials with the aid of cognitive (semantic-epistemic) operations can be used not only as a basis for meaning determination and meaning actualization in context, but also to operate innovatively on shared information, creating new generalizations, prototypes and metaphors; sometimes reinforced by innovative gestures.

5. References

.

[1] J. Allwood, "Meaning Potential and Context: Some Consequences for the Analysis of Variation in Meaning." In H. Cuyckens, R. Dirven and J. R. Taylor, (eds.) *Cognitive Approaches to Lexical Semantics*. *Mouton de Gruyter*. pp. 29-65, 2003.

[2] C. S. Peirce, *Collected Papers of Charles Sanders Peirce*, vols. 1–6, S. Hartshorne and Paul Weiss, eds., vols. 7–8, 1958, A. W. Bukrs (ed), Harvard University Press, Cambridge, Massachusetts, 1931-35, 1958.

[3] E. Rosch, "Cognitive representations of Semantic Categories". *Journal of Experimental Psychology:* General 104 (3): 192–233, 1975.

[4] J.. Allwood, "Semantics as Meaning Determination with Semantic Epistemic Operations." In J. Allwood and P. Gärdernfors (eds.) *Cognitive Semantics*. Amsterdam: Benjamins. pp. 1-18, 1999.

[5] H. Putnam, "The meaning of 'meaning." In *Philosophical Papers*, Vol. 2. *Mind, Language and Reality*. Cambridge University Press, 1975/1985.

[6] M. A. K. Halliday, *Explorations in the Functions of Language* . London : Edward Arnold , 1973.

[7] M. A. K. Halliday,. *Language as Social Semiotic: The social interpretation of language and meaning*. London : Edward Arnold, 1978.

[8] L. Wittgenstein,, *Philosophical Investigations*. Oxford: Blackwell Publishing, 1953.

[9] L. Vygotsky, *Thought and Language*. In A. Kozulin (Ed. and Transl.). Cambridge, MA.: MIT Press, 1986.

[10] J. Allwood, "A multidimensional activity based approach to communication." In I. Wachsmuth,, J. P. de Ruiter, P. Jaecks and S. Kopp (eds) *Alignment in Communication.* Amsterdam: John Benjamins, pp. 33-55, 2013.

[11] http://oxforddictionaries.com/definition/english/tree

[12] Wikipedia – Encyclopedia. https://en.wikipedia.org/wiki/Tree

[13] E. Ahlsén, and J. Allwood, "What's in a gesture?", *Allwood, J,. Ahlsén, E., Paggio, P. (2013). Proceedings of the Fourth Nordic Symposium on Multimodal Communication, Nov 15-16, University of Gothenburg. NEALT Proceedings Series No. 21.*. Linköping: Linköping University Electronic Press, pp. 13-20, 2013.