Recognition and Action

Roland Hausser
Universität Erlangen-Nürnberg (em.)
email: rolandhausser662@gmail.com

DOI: 10.13140/RG2.2.12716.95366

Abstract

In nonlanguage cognition, recognition precedes action, but in language cognition, action (speak mode) is ahead of recognition (hear mode) – minimally in spoken face-to-face communication, but possibly thousands of years in the medium of writing. For content transfer, the speaker converts language-dependent surface *types* into raw data and the hearer matches raw data with language-dependent surface types, resulting in *tokens*. The type-token relation (Peirce CP 5.171) goes back to Aristotle's distinction between the necessary and the accidental (Metaphysics, Books ζ and η), and is a central notion in DBS.[1]

1 Principles of Recognition and Action

The hear mode in language recognition is a special case of perception. In a cognitive agent, perception begins with raw input data. Provided by the agent's interface component, raw data are matched by a suitable type, resulting in a token as output. Consider the perception of a square as a DBS schema:

1.1 Perception: from raw data to type to token

edge 1: 2 cm		edge 1: α cm		edge 1: 2cm
angle 1/2: 90^{o}		angle 1/2: 90^{o}		angle 1/2: 90^{o}
edge 2: 2 cm		edge 2: α cm		edge 2: 2cm
angle 2/3: 90^{o}		angle 2/3: 90^{o}		angle 2/3: 90^{o}
edge 3: 2 cm	$\Rightarrow$	edge 3: α cm	$\Rightarrow$	edge 3: 2cm
angle 3/4: 90^{o}		angle 3/4: 90^{o}		angle 3/4: 90^{o}
edge 4: 2 cm		edge 4: α cm		edge 4: 2cm
angle 4/1: 90^{o}		angle 4/1: 90^{o}		angle 4/1: 90^{o}
raw data		*type*		*token*

The raw input data are in the agent's external reality, while the type and the token are in the agent's cognition. The edge length of the raw data is specified as 2cm. The corresponding value in the type is the variable α. The variable α in the type is replaced by the edge length of the raw data, resulting in a token.

The speak mode counterpart to language recognition is language production as a special case of action. In a cognitive agent, elementary action begins with a concept type which is adapted to a purpose by turning it into a token which is realized as agent-external raw data. This may be illustrated as follows:

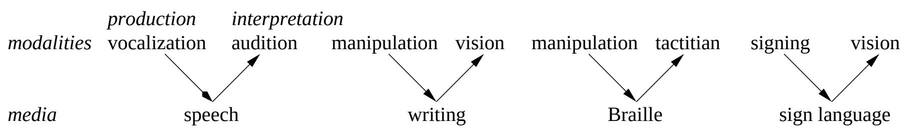
[1]Database Semantics, Hausser (2001).

1.2 Production: from type to token to raw data

type		token		raw data
edge 1: α cm		edge 1: 2cm		edge 1: 2 cm
angle 1/2: 90°		angle 1/2: 90°		angle 1/2: 90°
edge 2: α cm		edge 2: 2cm		edge 2: 2 cm
angle 2/3: 90°	⇒	angle 2/3: 90°	⇒	angle 2/3: 90°
edge 3: α cm		edge 3: 2cm		edge 3: 2 cm
angle 3/4: 90°		angle 3/4: 90°		angle 3/4: 90°
edge 4: α cm		edge 4: 2cm		edge 4: 2 cm
angle 4/1: 90°		angle 4/1: 90°		angle 4/1: 90°

What has been shown here abstractly for an uncomplicated example from two-dimensional geometry works analogously for language surface production and language surface recognition in the media of speech, writing, Braille, and signing, and serves as the declarative specification of computational implementations.

1.3 Sensory media and their modalities in communication



Each input medium has a complementary output medium (Hausser 2021, p. 32).

2 Vertical type-token matching in reference

In addition to horizontal type-token matching in language recognition (1.1) and production (1.2) there is vertical type-token matching in reference (Hausser 2017) as a matching between the concept and the context levels in the agent's cognition.

2.1 Reference in the speak mode

	edge 1: α cm		edge 1: 2 cm	
	angle 1/2: 90°		angle 1/2: 90°	
	edge 2: α cm		edge 2: 2 cm	
concept	angle 2/3: 90°		angle 2/3: 90°	*concept*
type	edge 3: α cm	⇒	edge 3: 2 cm	*token*
	angle 3/4: 90°		angle 3/4: 90°	
	edge 4: α cm		edge 4: 2 cm	
	angle 4/1: 90°		angle 4/1: 90°	
	⇑ *matching*			
	edge 1: 2 cm			
	angle 1/2: 90°			
	edge 2: 2 cm			
raw data	angle 2/3: 90°			
in memory	edge 3: 2 cm			
	angle 3/4: 90°			
	edge 4: 2 cm			
	angle 4/1: 90°			

The raw data input is matched by a concept type, resulting in a concept token.
Reference in the hear mode may be illustrated as follows:

2.2 Reference in the hear mode

raw data		*concept type*
edge 1: 2 cm		edge 1: α cm
angle 1/2: 90^o		angle 1/2: 90^o
edge 2: 2 cm		edge 2: α cm
angle 2/3: 90^o	$\Rightarrow$	angle 2/3: 90^o
edge 3: 2 cm		edge 3: α cm
angle 3/4: 90^o		angle 3/4: 90^o
edge 4: 2 cm		edge 4: α cm
angle 4/1: 90^o		angle 4/1: 90^o

$\Downarrow$ *matching*

edge 1: 2 cm
angle 1/2: 90^o
edge 2: 2 cm
angle 2/3: 90^o
edge 3: 2 cm
angle 3/4: 90^o
edge 4: 2 cm
angle 4/1: 90^o

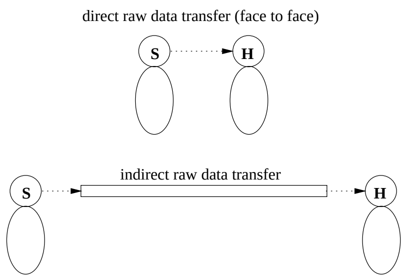
concept token

Here, the result of the hearer's raw data recognition is a concept token. The concept token serves as the referred-to item (*referent*) in both speak mode and hear mode.

3 Direct and Indirect Transfer

Transfer of content from the speaker to the hearer may be direct or indirect. An example of direct content transfer is face-to-face communication while examples of indirect transfer are talking on the phone or sending a letter. The distinction between direct and indirect transfer in natural language communication may be shown schematically as follows:

3.1 COMPARING DIRECT AND INDIRECT TRANSFER

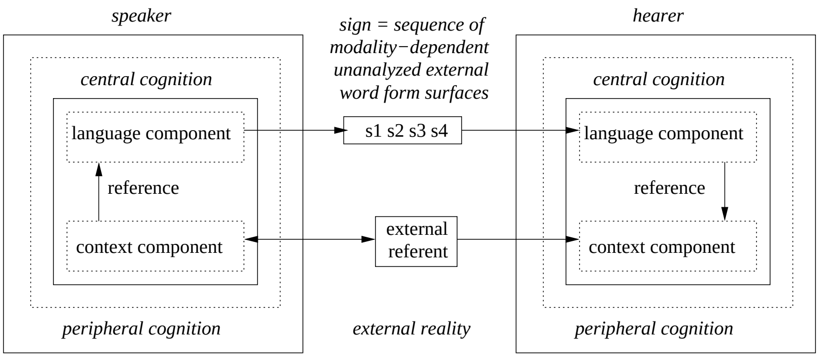


All that is required of an artificial or natural transfer channel is the transmission of data without distortion (Shannon and Weaver 1949). Though the transfer channel is not the place for reconstructing cognition (pace Eco 1975), it poses a crucial structural requirement for language communication: the signs must be in a linear order (canonized by de Saussure ([1916]1972) as his *seconde principe*). This is because humans can neither produce nor interpret two or more words, phrases, sentences, or texts simultaneously.[2]

4 Immediate and Mediated Reference

In communication, DBS distinguishes [+surface, +external], [+surface, -external], [-surface, +external], and [-surface, -external], with '±surface' indicating the presence or absence of a language sign and '±external' the location of the referent.

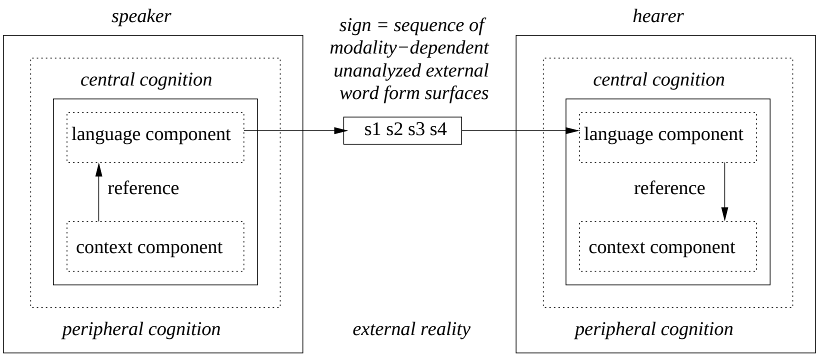
4.1 [+surface, +external]: immediate reference in communication



Agent-externally, language surfaces (shown here as s1 s2 s3 s4) are modality-specific unanalyzed external signs (raw data) which are passed from the speaker to the hearer and have neither meaning nor any grammatical properties whatsoever at all (no reification in DBS), but may be measured by the natural sciences.

The corresponding [+surface, −external] constellation between the speaker and the hearer may be shown as follows:

4.2 [+surface, −external]: mediated reference in communication



[2] The work of Shannon and Weaver is not exactly popular in Generative Grammar because substitution-based Phrase Structure Trees do not fit the time-linear transfer channel.

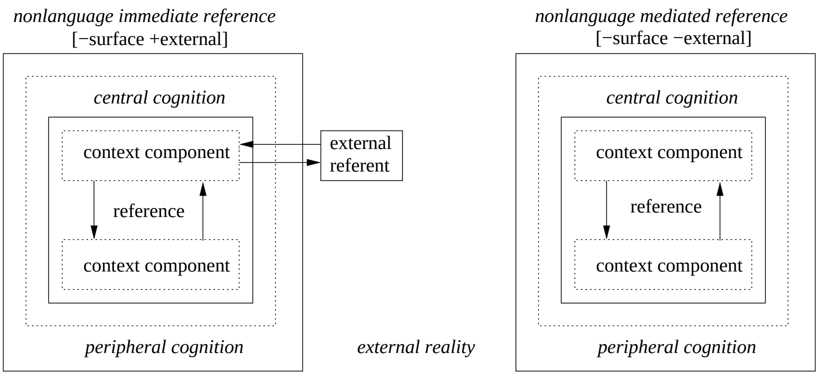
The reference relation begins with content in the memory of the speaker and ends as content in the memory of the hearer. The conventions of assigning surfaces to content and content to surfaces (Saussure 1916 [1972], première principe) are the same in immediate and mediated language reference.

The graphs 4.1 and 4.2 show the speaker on the left, the sign in left-to-right writing order in the middle, and the hearer on the right. This is a possible constellation which is in concord with the naive assumption that time passes with the sun from left to right (→) on the Northern Hemisphere. Yet it may appear that the first surface s1 leaves the speaker last and the last surface s4 arrives at the hearer first, which would be functionally incorrect.

It is a pseudo-problem, however, which vanishes if each surface is transmitted individually and placed to the right of its predecessor, i.e., (((s1 s2) s3) s4). This *left-associative*[3] departure and arrival structure allows incremental surface by surface processing, provided the derivation order is based on computing possible continuations, as in Left-Associative Grammar (Hausser 1992).

Finally consider the constellations of nonlanguage reference:

4.3 Nonlanguage immediate vs. mediated reference



The referring content in the [−surface +external] constellation is a current nonlanguage recognition, as when recognizing a person on the street. In the [−surface −external] constellation of nonlanguage mediated reference, in contrast, the referring content is activated without an external trigger, for example, by reasoning. In both, the referred-to content is resonating in memory (Hausser 2019, Sects. 3.2, 3.3).

5 Conclusion

The precomputational foundations of theoretical computer science in the 1930ties, 40ties, and 50ties inherited the sign-based substitution-driven ontology from mathematics and symbolic logic. Continuing from Frege, Hilbert&Ackermann, and

[3] Aho and Ullman 1977, p. 47.

Russell, the work of Church, Gödel, Kleene, Post, Tarski, and Turing resulted in a rich harvest of undecidable or undecided problems, such as the Entscheidungsproblem, the $P =? NP$ problem, Gödel's incompleteness proofs, the halting problem of Turing's virtual machines, and Post's correspondence problem.

To these venerable achievements, Chomsky added the claim that natural language is undecidable as well. The proof[4] relies on the complexity hierarchy of PSG (Chomsky hierarchy) and the theory-specific assumption, that the 'innate human language ability' combines a context-free base with a transformation component. This system, called *Generative Grammar* (GG), is recursively enumerable.[5]

Instead of distinguishing between the speak and the hear mode, Chomsky's GG derivations are all initiated by the same input, namely the nonterminal S node, working like a start button. The intended output is a random generation of well-formed expressions of a natural language, using the recursive substitution of nonterminals which are finally substituted by word form surfaces. Just as a vehicle requires a skilled human driver with vision and manipulation to keep it on the road, the stand-alone generation algorithm of GG requires a competent speaker of the natural language to distinguish between grammatical and ungrammatical output.

Based on the derivation principle of computing possible substitutions, GG is "not intended"[6] as a model of the speaker-hearer. Instead, the goal is a "universal" characterization of the human language "ability." This may have been misinterpreted as the implicit promise that understanding the universal human language ability would fundamentally facilitate computational language processing, despite GG's undecidability. Today, however, after more than half a century and a tremendous international effort, the assumed promise remains unfulfilled.

DBS and GG differ in their input and their output. GG takes an S node as input and is intended to derive all and only the well-formed expressions of a natural language. DBS, in contrast, distinguishes between the speak and the hear mode. The speak mode takes a content as input and produces a language-dependent surface as output. The hear mode takes a language-dependent surface as input and produces a content as output. Furthermore, the speak mode derivations may have a choice between paraphrases which are semantically equivalent. The hear mode derivations, in contrast, may be ambiguous, whereby the readings represent different contents.

The difference in the respective input and the output of GG and DBS requires different derivations. The substitution-based derivations of GG are vertical top-down; there is no distinction between the speak and the hear mode, and no upper limit on the number of substitution operations for the length of an *output*.[7] DBS,

[4]Peters, S., and R. Ritchie (1973) "On the Generative Power of Transformational Grammar," *Information and Control*, Vol. 18:483–501

[5]A generative grammar is recursively enumerable if, implemented as a Turing Machine, it will halt on grammatical but not on ungrammatical input (undecidable).

[6]Chomsky (1965) p. 9.

[7]In GG, a derivation of output intended to be of finite length may take infinitely many derivation steps because there is no limit on recursive cycles of alternating deletion and insertion transformations.

in contrast, distinguishes between the speak and the hear mode from the outset. The modes share the horizontal (left-associative) direction of derivation, but take opposite inputs and outputs. In either mode, the number of operations is a function of the input length.

Resulting from these structural differences, GG and DBS have orthogonal hierarchies of computational complexity (Hausser 1992). The classes of DBS are the C1, C2, C3, B, and A languages. The classes of PSG are the regular, context-free, context-sensitive, and unrestricted languages. The C-languages of DBS are so-called because they are C*onstant* in that each navigation (speak mode) or concatenation (hear mode) operation may take only a finite number of primitive operations, i.e. below a grammar-dependent upper bound; the only way to raise the complexity of a C-language above linear would be *recursive ambiguity*, which is empirically absent in natural language. The B languages are so-called because the number of steps in an operation is Bounded by the length of the input. The A language are so-called because they comprise only and All recursive languages.

In summary, the claim that natural language is undecidable holds only for a theory, namely sign-based substitution-driven GG, which fails to distinguish between the speak and the hear mode and is defined as a transformation component on top of a context-free base. Another theory, namely agent-based data-driven DBS, in contrast, would require recursive ambiguity for any complexity degree above linear. Recursive ambiguity, however, is absent in natural language (Hausser 2022). Consequently, the computational processing of natural language in DBS is of linear complexity, which is essential for real time performance.

Bibliography

Aristotle: J. Barnes (ed.), *The Complete Works of Aristotle*, Princeton Univ. Press 1974

Eco, U. (1975) *A Theory of Semiotics*, Bloomington, Indiana: Indiana Univ. Press

Hausser, R. (1992) "Complexity in Left-Associative Grammar," *Theoretical Computer Science*, Vol. 106.2:283-308, Elsevier

Hausser, R. (2001) "Database Semantics for Natural Language," *Artificial Intelligence*, 130.1:27–74, Elsevier

Hausser, R. (2017) "A Computational Treatment of Generalized Reference," *Complex Adaptive Systems Modeling*, Vol. 5.1:1–26, Springer

Hausser, R. (2019) *Computational Cognition*, `lagrammar.net`

Hausser, R. (2021) *The Ontology of Communication*, lagrammar.net

Hausser, R. (2022) “Grammatical Disambiguation, The Linear Complexity Hypothesis for Natural Language,” *Language and Information*, Korean Society of Linguistics, in print

Peirce, C.S.: C. Hartshorne and P. Weiss (eds.), *Collected Papers*, Cambridge, MA: Harvard Univ. Press. 1931-1935

Peters, S., and R. Ritchie (1973) “On the Generative Power of Transformational Grammar,” *Information and Control*, Vol. 18:483–501, Elsevier

Saussure, F. de [1916](1972) *Cours de linguistique générale*, Édition critique préparée par Tullio de Mauro, Paris: Éditions Payot

Shannon, C.E., and W. Weaver (1949) *The Mathematical Theory of Communication*, Urbana, Illinois: University of Illinois Press