USING THE DANCE TO INVESTIGATE THE PRAGMATIC/SEMANTIC BOUNDARY BETWEEN ARTIFICIAL AND NATURAL LANGUAGES

Laura Silver
University of Pittsburgh
Lawrence J. Mazlack
University of Cincinnati

0.0 ABSTRACT

This work addresses the pragmatic and semantic distinctions between natural and artificial languages by the development of a context-free generative grammar to describe motions in modern dance. The dance is a particularly good vehicle as it conveys meaning, but is undescribed by a generative grammar. Whether or not a grammar describing dance motion can be considered to be for a natural or artificial language is unclear.

1.0 INTRODUCTION

There are two different kinds of languages: natural and artificial. Artificial languages have been developed to deal with formal systems of man-created knowledge. Natural languages enable naturally arising entities to deal with their environment. Generally, artificial languages deal only with truth or knowledge that is specific to their artificial environment.

Both the written and spoken forms of human speech are universally considered to be languages. Animals as well as humans appear to communicate with each other through body motions. Whether or not body motion should be considered a language is open to debate. Some workers believe that the term "language" should be narrowly defined to include only signaling systems which are capable of manipulating abstractions. Others, would consider any origized system of signaling to be a language.

It is agreed that whatever a language is, its construction and interpretation is constrained by a specification mechanism. In languages, the construction specification is called a grammar.

Precisely how humans come to know the grammar of a language is unknown. One group of workers holds that it is learned. The other group, believes that the capability is innate. Irregardless of how men come to know the structures of spoken language, they certainly are capable of learning the grammars of artificial languages, for example, automata.

Both artificial and natural language can carry meaning; i.e., have semanticity. However, the semantic information represented by artificial languages appears to be of a different type than that of the information carried by a natural language.

In order to develop an understanding of the pragmatic and semantic differences between natural and artificial languages, a generative grammer is being developed to represent dance generation. The developed grammar is artificial, that described appears to communicate naturally. Whether or not the dance is a language is open to question as the tokens of the dance are never abstractions.

The problem is to understand the nature of language: of how humans perceive, understand and represent their world in their semiotic system. It is of further interest to develop an understanding of the relationship between the natural language system and the artificial language systems also developed by humans. These systems are intentionally created in order to represent systematically in systems of signs other perceptions in the symbolic manner or representation. Systems of signs can be represented as systems of signification where perceptions exist within the plane of content and are represented by the symbolic plane of expression. Artificial language systems such as mathematics and logic, are usually referred to as symbol systems, however they too are language systems and function as semiotic systems, as they are formal systems of signification.

In order to extend the analysis of information distinctions between the semantics of natural and artificial systems have to be clarified as do the distinctions between information and pragmatics.

2.0 BACKGROUND

Language, communication and information are three tightly interwoven concepts. The problem of information representation and communication is the focus of this work.

2.1 Language

Language is a process or symbolization that enables signification of some thing by representing it by something else. The "thing" represented has an existential space-time reality; the representation is an abstraction of the reality. Meaning is derived from the relationship between the physical and the symbolic

2.1.1 Meaning

Language provides the capability to functionally relate symbolized meanings. However, language is more than individual relationships among the meanings. Words, which are symbols, recursively become things themselves as they are utilized. As things themselves, they can be used symbolically to express or represent concepts as the next order or abstraction. Signs, symbols, words, tokens, pictographs are the tangible products of the interrelationship between the thought and the referent. This interface provides an operational definition for the nature of the concept of meaning, the property of language defined as "semanticity." That is, "the property of being able to convey meaning" [LYON 79].

2.1.2 Semiology: An Analytic Tool

De Saussure defined language as the Semiotic system; i.e., the science of signs. The sign is a subsystem, or a component of the system of language. The principle of "signification" indicates the relationship between the thing signified (the signified) and the things signifying it (the

signifier). Signifiers exist within the "plane of expression" and signifieds exist with the "plane of content." This relationship expressed by the sign as:

sign = (signifier, signified)

which is a specific relation between the plane of physical reality or "content" and symbolic reality or "expression." More generally, the sign is defined as:

sign = (plane of expression, plane of content)

A language is considered to be comprised of a set or system of signs.

Semiology aims to take in any system of signs, whatever their substance, and limits; images, gestures, musical sounds, objects, and the complex associations of all these...constitute, if not languages, at least systems of signification" [Bart 9]. Seminology will be used as tool in the analysis conducted by their work.

2.2 Natural and Artificial Language

Whether or not artifically constructed languages, or language schemas, can be considered as language "proper" is not central to this work.

Semiology, although initially concerned with natural signalling or communication systems set the stage for the analysis of any system of signs, whether they be natural languages or artificial language systems.

In discussing languages, Carnap states

> "so long as we are concerned with building this language, and not with its application and interpretation respecting a given theory, the signs of our language remain uninterpreted. Strictly speaking, what we construct is not a language but a schema or skeleton of a language: out of this schema we can produce at need a proper language (conceived as an instrument of communication) by interpretation of certain signs." [CARN].

Cherry discusses the difference between the natural and artificial kinds of languages.

> By 'language' we shall mean those organically developed systems, whether spoken or scribed, by which humans transmit messages; but the work 'cipher,' or 'code,' will be used to mean any invented, self-consistent system, whereby one set of symbols may be transformed into another for certain special stated purposes" [CHER 93, 94].

This difference between "language" and "code" can be understood not as a difference in structure, but as a difference in development. The concept of language generally implies an organic or natural development, and consequently referred to as "natural" language. The concept of code implies an intentional development, and consequently if referred to as "artificial" language systems.

2.3 Analyzing Language

The language being observed is usually called the object-language. The language used to discuss the object-language is called the metalanguage. The semiotic of the object language is formulated in the metalanguage system. Carnap identified the semiotic analysis of the object language into the three components of syntax, semantics, and pragmatics.

The terms syntax, semantics, and pragmatics are somewhat ambiguously applied. Part of the ambiguity of these terms is a function of whether the analysis of the object language is either the natural or artificial form of language.

According to Carnap, syntax "attends strictly to the expressions and their forms." However, "syntax may include rules which determine certain logical relations between sentences, e.g., the relation of derivability" [CARN 79]. The inclusion of the property of derivability in the syntactic component blurs the boundary between syntax and semantics.

2.4 Differing Semantics: Descriptive and Logical

Both natural and artificial languages contain the components of syntax, semantics and pragmatics. In the artificial language system, semantics refers only to the expressions and their designations without reference to any particular external system. In the natural language system, semantics includes the analysis of meaning by pointing to referents in the extentional world. In essence, there are two kinds of semantics, which can be understood as depending on either context-sensitive or context-free grammar. In artificially or "logically" constructed language systems, the grammar is context-free. In natural language, the grammer is context-sensitive. This latter form of semantics could be referred to as descriptive semantics, following the terminological distinction that "descriptive linguists" do the analysis. Their analyses are context-sensitive, in that they include the pragmatic component of meaning. In contrast, the form of semantics pertaining to the artificially or logically constructed language system can be labeled logical-semantics.

2.5 Communication and Information

Languages can be both naturally developed and artificially created. At the semiotic metalevel form of analysis both forms of language are treated as object level languages as both fulfill the need to signify; i.e., to represent perceptions and abstractions. This process of signification is more generally known as communication, where the language serves as an instrument of communication.

In the analysis of the problems inherent in communication, workers such as Shannon and Weaver have identified three levels of difficulties which complicate the problem of identifying information, particularly at the semantic level: (a) accuracy of symbol transmission, (b) communication of meaning, and (c) effectiveness (how conduct is affected). These three levels are all concerned with the concept that is labeled information, yet which "information" applies is not consistent for all three levels. Level A uses "information" as the amount of signal transmission, where Levels B and C the "information" is the semantic and pragmatic sense.

3.0 PROBLEM DOMAIN

In order to develop a context-free information representation a domain other than human verbal communication had to be selected. Verbal communication is too context-sensitive. Rather than working with the ambiguities of human verbal communication where it is difficult not to be pragmatic, or with information system design where the objective is to be pragmatic, the information system of human movement communication was selected.

Just as linguists have attempted to develop the notation for natural language grammars, seeking to represent the logical-semantic component, a similar grammatical structure of human movement can be developed. Generally, the domain of human communication is categorized into the verbal and the non-verbal. The verbal includes both the verbal and written forms of natural language. The non-verbal includes everything that is not verbal communication. Within this large category of non-verbal, the domain of human movement communication has been selected in order to construct a formal grammar representing the logical-semantic component of human movement information.

3.1 Purpose

This research investigates the semantic component of the artificial language system. The concern addressed is the clarification between the descriptive-semantics with the context-sensitive grammar (CSG) representation, and the logical-semantics with the context-free grammar (CFG) representation. The purpose is to illustrate the separation of the logical-semantic from the pragmatic, in order to demonstrate that it is possible to separate the information structure from potential meaning. The CFG is a template, providing the structure for the set of possible constructions any eventual user could select in order to represent any intended meaning. Prior to the representation of meaning a structure has to be defined whereby meaning representation can be made possible. Just as the information system can be viewed as both a process and referred to as a thing, a grammar can also. It is a template and therefore a thing, but it is a dynamic processing structure.

3.2 Existing Systems Representing Human Movement

The representational systems for human movement are data systems in that they are bound to some pragmatic component and that they are context-sensitive. Each has a basic set of symbols representing units that the user needed to represent. Although each representation system identifies a variety of syntactical units, no logical-semantic or grammar has yet been developed. Thus there is as yet no representation for the process of human movement information.

3.2.1 Notations

Labanotation is one of the most widely used representational systems for notating human movement [DNOT] [HUTC]. The notations are syntactic representations specifying syntactic units: direction, level, timing, and areas of the body, which are represented by unique symbols. It is not possible to use the system for anything but the description of the movement in the units provided by the initial symbol set. There are no structures or rules indicating relations among the syntactical units to generate more complex units. Consequently there is no representation by a grammar expressing the information system that is communicated by the movement.

Other notation systems are similar [SILV]. The Eschol-Watchman, the Benesh, Kinesics, Choreometrics, to name only a few, only differ in the specific particularization of the syntactic representation.

Why is there no system for representing the human body and its movement apart from any context? Perhaps because in the development of the representational systems, distinctions between the syntax and the logical-semantic were never clearly understood.

3.2.2 Models and Simulations

The objectives of various designers of models and simulations of human movement have been to extend the representational facilities of the human by mechanizing the laborious task of describing and computing problems in human movement. The goal was to develop a computer graphic display of a human model [POTT] [BILL].

Another area of research was the development of an interactive graphic editor for Labanotation [BROW]. The objective was to use the computer to facilitate the laborious process of hand writing Labanotation. This work was extended as part of the development of a graphic simulation for human motion [BALD] [TRAC].

3.2.3 Recognizing What To Know

The visual aspects of the perception of movement are essential in the design of mobile robots and the context-sensitive forms of representations are usefull. However, a context-free form of representation is preferable prior to any context-sensitive (i.e. applied) form of representation. For example, the visual aspect can be specified as a context-sensitive situation, which subsequently can be defined using a context-free grammatical structure. Research on this problem is important not only for the solution to problems in movement understanding, representation and generation, but also to illustrate the context-sensitivity of systems.

If we wish to represent the dynamic process of information, research must be done on abstracting the information from the pragmatics of use of that information. The structure of the process of information must be represented prior to the pragmatic application of the information.

3.4 Separating Logical and Semantic Descriptive Structures

Human movement and human verbalization both have the association of meaning with the sensorial transmittable component of movement and of speech that is transferred as a product of the information system. Where natural language has the symbolic representational facility of the written form, providing another channel for the transfer of the information, the movement notations do not. Just as the information process of human verbalization has been grammatically coded, the aspect of the problem that first needs to be addressed is the definition of the logical-semantic, i.e. the formal representation of the grammatical structure to code the information transferred via human

movement.

Before the grammar can be context-sensitive, it needs a context-free form. The problem is to develop a way to write movement information such that context-sensitive meanings can then be communicated in a written symbolic form.

3.4 Communication Systems: Human Movement Compared to Natural Language

The semiotic system of human movement communication was selected as the domain in which to investigate representation of a natural activity in an artificial language.

Adequate representation of the generative structure of the semiotic system seems to be the necessary and sufficient conditions which linguists, anthropologists and philosophers require for "language" identification. Where natural language encompasses both verbal and written forms of the sounds and their meanings, movement language exists only with what can be equated with the verbal level of natural languages. A comparison of this difference would be equating the notations for movement with the phonological orthographic representations of the sounds of natural languages. Each natural language has particular orthographic symbols necessary to represent the sounds of that language, just as each form of movement has developed notational symbols to represent the visual perception of movements particular to that form. However, where verbal language has not only the particular phonologogical representation, it further has a representational form which is called the written from where the meaning in the experiential form can symbolically be represented.

4. INVESTIGATIVE STRUCTURE

The specific problem addressed is the development of a prototype for information representation, by experiment with representions of the logical-semantic structure of human movement information using the BNF form of the context-free grammar as the analytic tool.

It is posited that the situation in natural language representation is analogous to problems in information system design. Before the grammatical structure is constructed, comprised of the vocabulary elements of the system and the set of relations among them, particular referents to the units are assigned, building the context-sensivitity into the initial design of the system. The idea of the context-free form of representation preceeding any context-sensitive representation is the direction of this research.

4.1 Role or The Grammar

The grammar itself is a representational template, in that it does not contain the meaning, but rather provides a structure. The grammar is a process in that it is used as a template. It is a commodity in that it is a tool constructed for analytical and representational purposes. It is tied to a particular form of representation, but it is relatively context-free. Any language system is a particular form of a semiotic system useful to communicate a range of meanings, and sensitive to that range. (Whorf defined this concept as "linguistic relativity" [WHOR].) Yet, the same semiotic system is context-free, in that it has the feature of productivity, and can generate valid expressions to represent new meanings even in the extensional world to which it is bound.

Looked at in this way, a grammar exists at the meta-level, providing a form of analysis for an information structure for any possible object-level expression that is generated in that language system.

BNF was chosen to represent the grammar. It provides a method of notation with the capability to code information that is dense and non-linear. BNF also lends itself to consistency verification.

4.2 Scope: Context-Free Representation of An Information System

In the analysis of the communication or semiotic system of human movement that is to be represented, only the logical-semantic form of the semantic component will be considered in order to illustrate that context-free representation is possible when the coding is only of the information system rather than including the pragmatics of the communication system.

This reception of data as information by the receiver is the pragmatic component, which is added to the input data from the sender. Meaning is the result of the contextual processing of data given some information input.

In order to develop a context-free grammar for the logical-semantic of human movement information, a non-purposeful context needs to be examined, i.e., where the movement is not intended to communicate any meaning but where the units of movement are learned for the production of movement itself, which subsequently can be used in various contexts to communicate a variety of meanings.

4.3 Dance Units

Dance instructors teach the units of the movement language without any intended transfer of information other than how to produce the units of movement. The vocabulary of movement that is used for dance is a complex series of units, which are deriveable in terms of initial units, plus rules for connecting the various units. These more complex units are referred to as "combinations." The units and the combinations are the information communicated in dance instruction.

4.4 The Goal: A Movement Semantic

This methodology formally can be represented as an operation of the logical structure of the BNF grammar, operating upon the selected scope of the verbal channel of the domain of the information system of human movement yielding as a product a grammatical representation of the logical-semantic component of the information system.

The movement semantic will be the grammar derived from the operation of the template processing the logical-semantic structure of the information into a representational form. This product will represent the results of research of the representation of the dynamic structure of the information process in a grammatical context-free form. The form of representation is that of a formal logical system.

4.5 Verification

A form of logical verification can be accomplished

by using LEX, the lexical analyzer, and YACC, the compiler compiler of the UNIX operating system. One of the advantages of using the BNF notation is that the movement semantic being developed and the code that LEX recognizes are both in the context-free form which is based upon the BNF notation.

4.6 Project Summary

The project will: 1) represent a portion of the logical-semantic information structure of a selected domain of human movement information, 2) represent a prototype for a written code of a representational rather than an experimental human movement language where 3) the symbolic representation of human movement information is accomplished using a dramatical rather than a descriptive template. The aim is to define a subset of human movement information code that meets these criteria of the logical or artificial system, such that it can be used without the problems of contradictory and ambigious expressions that are inherent, for example, in natural language systems.

5.0 REFERENCES

BADL Badler, N. J., Smoliar, S. W., "Digital Representation of Human Movement," Computing Surveys, Vol. 11, No. 1, March 1979.

BART Barthes, Roland, Elements of Semiology, Layers, Smith (trans.), Hill and Wang, New York, 1968, c1964 Elements de Semiologie.

BILL Billings, M.P., Yucker, W. R., "The Computerized Anatomical Man (CAM) Model," NASA-CR-134043, MDC-G4655, CNT: NAS9-13228, Issue 23, 1970-71.

BROW Brown, M. D., Smoliar, S. W., "A Graphic Editor for Labanotation," Computer Graphics, Vol. 10, No. 2, Summer 1976.

CARN Carnap, Rudolph, Introduction to Symbolic Logic and Its Applications, Meyer, W. H., Wilkinson, J. (trans.), Dover Publications, Inc., New York, 1958, c1954 Einfubrung in die symbolische logik, Springer.

CHER Cherry, Colin, On Human Communication, MIT Press, Cambridge, Massachusetts, 1980, c1957.

DNOT The Dance Notation Bureau, Courses and Programs, New York, 1980.

HUTC Hutchinson, A., Labanotation, Theather Arts Books, New York, 1977.

LABN Laban, R., The Language of Movement: A Guide to Choreutics, Plays, Inc., Boston, 1974, c1941.

LYON Lyons, J., Semantics, Cambridge University Press, Cambridge, 1977.

OTTE Otten, K., "Basis for a Science of Information," Information Science: Search for Identity, Debons, A. (ed.), Marcel Decker, New York, 1974.

POTT Potter, T. E., Willmert, K. D., "Three-Dimensional Display Model," Office of Naval Research, July 1975.

SAVA Savage, G. J., Officer, J. M., "CHOREO: An Interactive Computer Model for Dance," International Journal of Man Machine Studies, Vol. 10, 1978.

SHAN Shannon, C., Weaver, W., The Mathematical Theory of Communication, University of Illinois Press, Urbana, Chicago, 1980, c1949.

SILV Silver, L. D., "Towards a Movement Lanugage: On the Representation of Movement Knowledge," manuscript, Interdisciplinary Department of Information Science, University of Pittsburgh, Pittsburgh, Pennsylvania, 1981.

TRAC Tracton, W. P., "GEL: A Graphic Editor for Labanotation with an Associated Data Structure," Movement Project Report No. 15, The Moore School of Electrical Engineering, The University of Pennsylvania, Philadelphia, August 1979.

WHOR Whorf, B. L., Language, Thought and Reality, The MIT Press, Cambridge, Massachusetts, 1979, c1956.