Bassey E. Antia
Department of Languages & Linguistics
University of Maiduguri
Nigeria

A THEORETICAL NOTE ON APPLIED CONCEPTOLOGY: CONCEPTUAL CATEGORIES AS AN ANCILLARY TO METHODS OF DATA ANALYSIS AND INTERPRETATION IN SOCIAL SCIENCE-ORIENTED RESEARCH

0. ABSTRACT

Concepts have again been on the front-burner of theoretical discussion within terminology, as witnessed by conference workshops dedicated to this topic at the 1997 European Symposium on Language for Special Purposes (Copenhagen) and at the 1999 Conference on Terminology and Knowledge Engineering (Innsbruck), among other fora. Without prejudice to whatever new perspectives are brought to bear, preoccupation with the same old issues (what a concept is, whether it can be used for anything practical, etc.) has the potential of preventing research from going in new directions that have applications and implications outside the narrow circle of 'buffs' of recondite theory. This article suggests one such direction. Taking its point of departure from *conceptual primitives* or *categories* such as underpin the organization of language thesauri, this article demonstrates how work on concepts may be applied to text interpretation in domains of socially inflected discourse, particularly in those contexts where robust interpretation requires the elevation of specific facts to the level of symbols (thus allowing for generalization) and a probing of the relationship between what is stated in a piece of discourse and its opposite in some conceptual scheme. Because both of these activities make it possible for a piece of discourse to transcend its immediate context, they may be said to constitute disarticulations from the specific facts that are articulated. Articulation is synonymous with discursification and textualization. This contribution is in the tradition of Khurshid Ahmad's research programme which employs heuristics derived from terminology to account for the evolving discourse of scientific communities.

1. INTRODUCTION

Philosophers of science, sociologists, political scientists, public policy makers, historians, literary critics, mass media analysts, among other social scientists, are all interested in the construction and/or evolution of 'consciousness' – whether this is seen in terms of scientific facts, collective representation (in an identity defining and contrasting sense), world view, or general awareness of products, policy, etc.: what is the accepted position on x in our field and how did we get there? As an ethnicity (recognised or previously marginalised): who are we? How do we distinguish ourselves from our neighbours? How do we institutionalise our uniqueness and convey same to our offspring? As a country or an alliance, why should we

attack x now when we could have done that y years ago? Why is x the cure-all wonder drug? Etc. These are all questions dealing with the construction and/or evolution of consciousness.

Somewhat more specifically, the interests of the aforementioned groups of social scientists may be said to cluster around two poles. One constellation of interests is oriented towards the relationship between a set of non-linguistic facts (e.g. acts of an enemy that are believed to justify an attack) and the mobilisation of public opinion through the discursification/textualisation of those facts. Discursification is the means by which consciousness/ideology of a particular community is created, recorded, transformed or evaluated. Calhoun (quoted by Chang 1997) notes that '[p]ublic discourse (and what Habermas later and more generally calls communicative action) is a possible mode of coordination of human life, as are state power and market economies...'.

It is central to the second constellation of interests that discursification simultaneously possess a projective/creative function (in the sense of actually creating a set of facts, creating/inventing reasons for an attack, offering a product to the market) and invite a perlocutionary analysis of the discourse (in the sense of examining intended/unintended effects). In speech act theory an analysis of the perlocutionary effect of an illocutionary act involves determining the act's effects relative to an original intention.

These constellations of interests are of course a matter of degrees. The first would clearly be associated with philosophers of science interested in scientific (r)evolutions, besides being identified with other social scientists concerned with how specific social experiences (e.g. current corporate scandals in America) are articulated in discourse and yet disarticulate from, or transcend, these specific experiences (corruption everywhere else). The second thrust would be associated with legislators, other public policy decision makers and, say, product manufacturers and their market analysts who are keen to obtain feedback on their offerings (product or policy) to the public.

What is common to both thrusts is the appeal to communities of discourse, or the invocation of discourse in the spheres corresponding to their respective interests. Books and other kinds of literature are written to create, document, but also to transform a given set of facts. Newspapers are founded to record and influence public opinion. Manufactured products are accompanied by some documentation giving information on usage, benefits, warranty, etc. – all crafted in a manner to elicit positive response from the target audience. Questionnaires are administered to ascertain the possible effect of these attempts to influence public opinion. It is one thing to obtain or to create all of these evidence sources, and another to obtain consensus on what they say.

2. PROBLEMS OF INTERPRETATION IN SOCIAL SCIENCE RESEARCH

There has been some concern about the methods used in analysing socially-inflected discourse, and consequently about the interpretations they allow for. To illustrate this, four dif-

ferent kinds of supporting evidence are cited from, respectively, the intersection of political science, sociology and mass communication; literary criticism; music; and finally the philosophy of science.

Our first example comes from the intersection of political science, sociology and mass communication. After World War II, the ideological component of U.S. foreign policy in its zone of occupied Germany called for a thorough programme of re-education of Germans in American values, particularly those values associated with democracy. This was the context for the Joint Chief of Staff's 4D agenda on denazification, democratisation, demilitarisation and decentralisation (cf. Gienow-Hecht 1997). On the basis of its conviction that a local news medium was an ancillary to the success of this programme, the Information Control Division (a unit of the U.S. Office of Military Government in Germany) founded the newspaper, Die Neue Zeitung. The paper would be run by Jewish émigrés who, after their flight from Germany, had become citizens of America - where they served in the U.S. Army's Psychological Warfare Division (cf. Gienow-Hecht 1997). According to Gienow-Hecht, the U.S. Office of Military Government (OMGUS) were quite negative in their evaluations of the work of the editors of Die Neue Zeitung. This was on account of the paper's emphasis on art, and scant reference to democracy or to America. In titling her discussion Art is Democracy and Democracy is Art, Gienow-Hecht catches and makes the point so sorely missed by the OMGUS and other U.S. political scientists: that is, that the conceptual categories of tolerance, diversity, respect, consensus, etc. developed by the editors in an artistic context were precisely the same that underpin democratic culture.

Our second example comes from African literature of French expression, from the era of anti-colonial protest literature. When the eminent Guinean novelist, Camara Laye, published his classic L'Enfant noir (most commercially successful English translation: The African Child) as a student in 1950s France, he was attacked by another celebrated novelist, Mongo Beti. The Marxist oriented Beti, putting the hat of a critic, wondered if Laye had not in the least been affected by the evils of colonialism. He wondered why Laye would indulge in the luxury of a nostalgic recollection of an idyllic, pre-European contact, Africa. This at a time when all creative and non-creative hands were on deck in the attempt to dismount colonialism. Now, only later did a general critical perception of Laye come to accept the theory of causality, which saw in Laye's apparent disarticulation from colonialism a most creative articulation of the negative effects of this system.

Our third example comes from music. Prior to the dramatic changes in South Africa culminating in the inauguration of the African National Congress (ANC) administration, Yvonne Chaka Chaka was one of the better known musical artists from that country in Nigeria. This was of course besides the legendary Miriam Makeba. A plan for her to visit Nigeria for a concert did not materialise because of what was reported as the uncooperative attitude of staff at what was then the ANC Office in Lagos, Nigeria. The official who was to serve as facilitator reportedly explained the attitude of his office by their commitment to only ambassador-artists who would articulate the sufferings and aspirations of the oppressed black majority in South Africa, and not some artist singing about male-female relationships and locally brewed beers. I cannot ascertain if at the time Yvonne had recorded the song Freedom, but the point is that her line of musical expression was perceived as far removed from

the front-burner existential concerns of the majority of her race. As in the Camara Laye case, it did not really occur to this self-appointed regulator of creative expression and of its interpretation that Yvonne's articulation of seemingly irrelevant issues could have been her own way of painting her dream society – one in which revolutionary rhetoric, killings and psychological torture would give way to family and socialisation as issues of the day. In other words, in Chaka Chaka's alleged disarticulation from the pressing issues of the day we find an articulation of the conceptual category of peace – also the destination of the fire-spitting activists.

Our final example, which is from the philosophy of science, serves to underscore the importance of concepts as a method of interpretation. Thagard (1992) contrasts belief revision approaches to knowledge change with conceptual approaches. The former, in their use of discursified community facts, operate with sentence-like propositions, while the latter operate with concepts. Thagard cites evidence in support of the claim that contemporaneously philosophy is oriented towards sentences. Sentences are contemporaneously the object of epistemological investigation. The consequence of slighting conceptual approaches is that it has not been possible to come up with a 'finer-grained theory' of knowledge change that employs tools that differ from the 'vague historical ones' (Thagard's description) used by Thomas Kuhn. More specifically in terms of results, Thagard notes that propositional approaches to knowledge change are ill-equipped to 'account for why some revisions are harder to make than others and why some revisions have more global effects.' It is hypothesised that these are issues that are best understood by 'noticing how beliefs are organised by concepts.'

3. DISCOURSE INTERPRETATION WITHIN TERMINOLOGY

Terminology focuses on specialised or specific subject areas within which it studies knowledge (units [e.g. concepts], structure, representation, evolution, acquisition, use, etc.) in its relation to expression (cf. Antia 2000:1). Conceptology, or the study of concepts, is well-developed within terminology because terminology approaches knowledge from the standpoint of conceptual logic, as opposed to propositional logic. Concepts (to which labels – linguistic or non-linguistic would normally be assigned) are the building blocks of knowledge. The difference between a sample of textualised specialised knowledge and a sample of general knowledge text lies in the type-token ratio of concepts, more commonly referred to as lexical type-token ratio. The specialised knowledge text would have a higher concept type-token ratio than the general knowledge text. Seen in terms of concepts, the lexical closure hypothesis simply means that, in a statistical analysis, a specialised text reaches closure or enumerates itself conceptually faster than a general knowledge text. The foregoing explains why terminology sees knowledge in terms of concepts, and also why the terminology framework is chosen as context for this study, involving conceptual categories.

Now, within the terminology community, Khurshid Ahmad's research programme is easily the closest to the issue at hand in this study. Ahmad has been keen to interpret scientific or knowledge discourses from the standpoint of conceptual change. This change is studied in

texts, in the representative works of the knowledge universe of interest. The heuristic for the analysis is the terminology dynamic in evidence, in other words, the empirically attested patterns of term use or the competing term preferences in the carefully constituted text corpus. These terms are believed to index stages, researchers, etc. within the knowledge space studied. Illustrative data have come from nuclear physics, philosophy of science, mathematics and linguistics. We shall use the latter (cf. Ahmad 2002).

In support of the movement in Chomskyan thought, from the perception of grammar as description to grammar as rules, the following are observed in Ahmad's Chomsky sub-corpus: from a combined frequency of occurrence of 1.68% in (and relative to the words in) Syntactic Structures (1957), the term Grammar and its variants (e.g. grammars, grammatical, grammatically, etc.) drop to 1.58% in Aspects of the Theory of Syntax (1965), then to 0.64% in Government and Binding (1981). These observations are further buttressed when it is known that in the corpus under analysis terms like grammar, sentence and language which were leitmotifs in Syntactic Structures – with a very high frequency comparable to that typically associated with the determiner the (5%) – lose about 2% frequency in Aspects of the Theory of Syntax, and have only 1% left in Government and Binding. From a frequency analysis of compound terms, Chomsky is seen as heading conceptually towards core grammar, case filter, structural case, case theory, etc. – apparently leaving behind generative grammar.

Ahmad's studies of specialised discourses, for which a semasiological or term-driven approach is most appropriate, even supports the kinds of onomasiological or concept-driven interests (conceptual relations, concept branch jumping, etc.) which Thagard has in his analysis of conceptual revolutions in Geology, etc. The lexicogrammatical dimension in Ahmad's analysis of terms makes this feasible.

The data Ahmad uses are from specialised fields. That is of course what terminology studies. It however seems that in the application of insights on concepts in terminology to the interpretation of discourses in less structured (i.e. the one-off type or more general knowledge-oriented) contexts, as is the case particularly with the first three of the four examples cited earlier, a device or framework complementary to Ahmad's is required.

Below we suggest a complementary conceptology-inspired framework, but the presentation is preceded by some theoretical scaffolding derived from discourse research in the social sciences.

4. A THEORETICAL FRAMEWORK

Articulation is a construct that has been used (by Wuthnow 1989) to describe the relationship between discourse and the social environment that produces it. Simply put, articulation is a socially conditioned illocutionary act. For our broader purposes we might replace social environment with community. Articulation, if it is not to be adjudged too parochial to be self-commending, or too abstract to be relevant, must necessarily involve a balancing act; in other words, it should imply a measure of disarticulation. According to Wuthnow, '[s]ome

features of an ideology resonate closely with the social context in which they appear; others point towards context-free concepts and generalizations.' The latter would be disarticulations of experiences of community facts. The idea of disarticulation being ideally embodied in articulation is taken by Wuthnow to suggest that the 'search for features of ideology [read 'consciousness'] that resemble features of the social milieu [read 'community'] must also include an account of ways in which an ideology becomes at least partially free of contextual determination.' Among other distinctions which Wuthnow makes in connection with articulation (context, manner), there is the one that discusses what is actually articulated. Three 'whats' are identified.

The social horizon [read community] provides the facts from which a consciousness eventually emerges through the selective and transformational process of textualisation or discursification. The discursive field gives a structure to the consciousness, and specifies the conceptual categories to be employed in talking about this consciousness. Articulation here involves mapping onto this structure specific facts from the community, while disarticulation involves, in Wuthnow's words, 'identifying ways in which the discursive field provides contrasts with features of the social horizon itself, thereby evoking a conceptual space in which creative reflection can take place.' Figural actions or actors, in the slot and filler terms of frame descriptions (in the field of Artificial Intelligence), are the fillers of the discursive field slots. Articulation here involves identifying prototypical facts (behavioural modes or personages) from the social horizon (as it is structurally mediated by the discursive field). Disarticulation or the transcending of specific behavioural modes and personages means increasing the metonymic function of such modes or personages. We have here something of an algorithm for creative/poetic writing, which invites interesting comparisons with stages of discourse production in text/discourse linguistics (cf. discussion of Frederiksen's model in Antia 2000: 156).

The implications of the foregoing need to be stated. The creation of the discursive field, just like disarticulation at the level of this field and of figural actions, calls for operating with conceptual categories. Operating with conceptual categories may involve operating with a system of contrasts, e.g. affirmation and negation, where either can serve as articulation and the other as disarticulation. From the standpoint of the analysis of evidence sources, concerns obviously relate to reconstituting the social horizon, identifying the structure of the discursive field, and finally assigning to figural actions and actors dimensions which, while not necessarily being universal, have fewer spatio-temporal constraints.

5. TOWARDS COMPLEMENTARY DIRECTIONS OF INTERPRETATION

Let us attempt to first work out the methodological implications of the examples we provided in section 2 above. From these four examples, particularly the first three, we draw a number of implications concerning the use of conceptual categories. Without it amounting to seeing facts where they do not exist, the use of conceptual categories (within the theoretical framework sketched earlier) to interpret evidence sources has vertical and horizontal implications.

Vertically, conceptual categories or probes of various degrees of specificity or depth are called for. For instance, to retrieve the structure of a discursive field and to allow figural actions and actors assume their full metonymic functions, it should be possible to map concept tokens to concept types or primitives. American social scientists and the OMGUS were clearly unable to disarticulate, to go beyond tokens.

Horizontally, because articulation necessarily involves disarticulation, the presence of a conceptual category would perhaps be as significant as the absence of another with which it correlates (to form a pair) along a certain dimension (e.g. opposition) and in a certain logical, axiomatic, system. This is as evident in the literary example as in the OMGUS and Yvonne Chaka Chaka examples. What these three examples share with the philosophy of science example is the fact that in all four cases propositional approaches to evidence sources make it difficult to retrieve important subtleties in community facts.

If we saw the above as specifications for a conceptology-inspired or -associated framework of general knowledge text interpretation, then a thesaurus would seem to meet them. This would be particularly true of a computerised thesaurus.

6. THESAURUS MODEL OF CONCEPT CATEGORIES FOR INTERPRETATION

A language thesaurus is a resource that enters words, not according to the alphabet, but conceptually or according to ideas. For instance, an entry in a language thesaurus would list words that are more or less synonymous. The pre-modifier ‘language’ seeks to alert to the existence of another type of thesaurus meant for documents, but using the same basic principle of organisation.

One of the best known thesauri in the English language is Roget’s eponymous thesaurus, Roget’s Thesaurus of English Words and Phrases, which first appeared in 1852. Among the many motivations for this (and similar resources in other languages), there is the one on knowledge classification (cf. Antia 2000). It was Roget’s hope that his work would contribute to the search for a universal scheme for classifying general or pre-scientific knowledge. In setting for himself this goal, Roget was going back to a 17th century research concern of philosophers like Descartes, Leibniz, Dalgarno, Wilkins, and the like.

The point in our discussion being that the articulation of a conceptual category can simultaneously read as disarticulation, and that at the discursive field the relation between articulation and disarticulation can be one of affirmation and negation, it is easy to see how Roget’s thesaurus might support such a reading. Table 1 is the plan of classification of the thesaurus, but Table 2 (synopsis of a class of categories) is perhaps more revealing.

Figure 1: Roget's Plan of classification

Class	Section	Heads
1 Abstract relations	1 Existence	1 - 8
	2 Relation	9 - 25
	3 Quantity	26 - 59
	4 Order	60 - 84
	5 Number	85 - 107
	6 Time	108 - 142
	7 Change	143 - 155
	8 Causation	156 -182
2 Space	1 Space in general	183 - 194
	2 Dimensions	195 - 242
	3 Form	243 - 264
	4 Motion	265 - 318
3 Matter	1 Matter in general	319 - 323
	2 Inorganic matter	324 - 357
	3 Organic matter	358 - 446
4 Intellect: the exercise of the mind		
Division one: Formation *of ideas*	1 General	447 - 452
	2 Precusory conditions and operations	453 - 465
	3 Materials for reasoning	466 - 474
	4 Reasoning processes	475 - 479
	5 Results of reasoning	480 - 504
	6 Extension of thought	505 - 511
	7 Creative thought	512 - 513
Division two: Communication *of ideas*	1 Nature of ideas communicated	514 - 521
	2 Modes of communication	522 - 546
	3 Means of communicating ideas	547 - 594

Class	Section	Heads
5 Volition: the exercise of the will		
Division one: Individual volition	1 Volition in general	95 - 616
	2 Prospective volition	617 - 616
	3 Voluntary action	676 - 699
	4 Antagonism	700 - 724
	5 Results of action	725 - 732
Division two:	1 General social volition	733 - 755
	2 Special social volition	756 - 763
	3 Conditional social volition	764 - 770
	4 Possessive relations	771 - 816
6 Emotion, religion and morality	1 General	817 - 823
	2 Personal emotion	824 - 879
	3 Interpersonal emotion	880 - 912
	4 Morality	913 - 964
	5 Religion	965 - 990

Table 2: Extract from Roget's Tabular synopsis of Categories

Class one: Abstract relations		
1 Existence		
Abstract:	1 Existence	2 Nonexistence
Concrete:	3 Substantiality	4 Insubstantiality
Formal: (internal / external)	5 Intrinsicality	6 Extrinsicality
Modal: (absolute / relative)	7 State	8 Circumstance

Table 2 amplifies Class 1 and section 1 (Existence) in Table 1. Table 2 shows the heads in a section as being presented in contrasting pairs (existence-nonexistence, etc.). This pattern is replicated in the text of the thesaurus. In his introduction, Roget writes:

For the purpose of exhibiting with greater distinctiveness the relations between words expressing opposite and correlative ideas, I have, whenever the subject admitted, placed them in two parallel columns in the same page, so that each group of expressions may be readily contrasted with those which occupy the adjacent column, and constitute their antithesis (p. xxix).

This layout has regrettably been changed in several revised editions (e.g. in the 1982 edited by Lyold used here).

How might this work in light of the vertical and horizontal methodological implications of the examples studied as stated in section 5 above? Recall that, vertically, conceptual categories of various degrees of specificity were called for, so as to make possible the mapping of concept tokens to concept types; this was said to be important in retrieving the structure of a discursive field while allowing figural actions assume their metonymic function. Horizontally, it was a case of correlating conceptual categories (to form a pair) along a certain dimension (e.g. opposition) and in a certain logical system.

In a computerized thesaurus environment, there could be developed for a conceptual category (e.g. a Roget 'head' or simplification/refinement thereof) an archive of lexical types. These lexical types would be possible verbal manifestations of the category. A lexicalised concept token occurring in text gets matched with a corresponding entry in the lexical archive, from where mapping onto the parent conceptual category (i.e. a Roget 'head') takes place. The result of this matching would be what is articulated. Now because this 'head' or conceptual category is related to another category, the next step in processing would be to connect to the opposite or correlative category. The outcome would provide the basis for a disarticulated reading.

This author welcomes views on the use of general language ontologies (e.g. WordNet) for man and machine text interpretation within the framework described here.

REFERENCES

AHMAD, K. (2002). Como nascono i termini: zeri, comportamenti e nuclei. In: M. Magris, M.T. Musacchio, L. Rega & F. Scarpa (eds) Manuale di terminologia. Aspetti teorici, metodologici e applicativi. Milan: Hoepli, 63-71.

ANTIA, B.E. (2000). Terminology and Language Planning: an alternative framework of discourse and practice. Amsterdam/Philadelphia: John Benjamins, xxiv + 264.

CHANG, Han-pi (1997). Taiwan: Community of Fate and Cultural Globalization. Münster: Lit (Market, Culture and Society 3).

GIENOW-HECHT, C.E. (1997). Art is Democracy and Democracy is Art: Culture, Propaganda, and the *Neue Zeitung* in Germany, 1945 – 1947. Manuscript.

LYOLD, S.M. (ed.). Roget's Thesaurus of English Words and Phrases. Essex: Longman.

THAGARD, P. (1992). Conceptual Revolutions. Princeton, New Jersey: Princeton University Press.

WUTHNOW, R. (1989). Communities of Discourse. Cambridge, Massachussets: Harvard University Press.