

1969 International Conference on Computational Linguistics

Semantics and the Syntactic Classification of Words^(*)

Ernst von Glasersfeld
(Georgia Institute for Research,
and University of Georgia)

In presenting some aspects of our somewhat unusual grammar to a gathering of computational linguists, I feel justified in taking two things for granted: first, that most people who have taken an active interest both in computers and in natural language have come to realise that computers, although impressively fast and reliable in many tasks, are not very brilliant when it comes to making inductive decisions on the basis of insufficient or not thoroughly defined data; and, second, that what we linguists know about the workings of natural language is by no means enough to supply computers, as they are, with a solid basis for the fully automatic handling of natural-language data.

This is not meant to be a disparaging comment on previous efforts in linguistics, for, after all, the contingency that has opened our eyes to these shortcomings did not exist until a few years ago. Traditional linguists, and especially grammarians, could carry on their business quite happily on the more or less explicit assumption that language - like so many things in the still wide-spread Platonic view of the world - could be separated into two levels: the ideal, uncontaminated one of pure structure

* The research reported in this paper was sponsored by the U.S. Air Force Office of Scientific Research (DAR), Information Sciences Directorate (Grant AFOSR 1319-67) and actively supported by the University of Georgia, Athens, USA.

or form that was inherently generalisable, and the slightly messy, unsystemic, and therefore far less interesting one of individual content. This dichotomy was possible and workable as long as the use of language was restricted to organisms who, by the time they embarked on linguistic activities, had necessarily absorbed a vast body of experiential and conceptual knowledge, on which, more or less consciously, they could draw whenever the formulation or comprehension of a linguistic expression required something beyond the rules of the ideal grammar.

With the advent of computers the situation was radically changed. Suddenly linguists could and would find themselves committed, for one reason or another, to transmit their know-how to a potential language user who did not possess any a priori experiential or conceptual knowledge whatsoever. Thus there arose innumerable questions which, hitherto, no one had ever been compelled to answer, and it became painfully obvious that the application of much of the linguists' cherished theoretical knowledge to actual language material presupposed a considerable amount of as yet unexplored preprocessing of that very material.

I shall not try to catalogue the types of question and the kinds of problem which have been thrust upon the linguist by the appearance of computers - every one of you is familiar with some if not all of them; instead I should like to present a few examples and show the direction in which, we believe, some solutions can be found.

By 'linguistic activities' or 'handling language data' we mean, for instance, formulating a given message in a specific natural language; or recognising, in a given piece of language, the message intentionally formulated

in that way by an author; or translating a formulation given in one natural language into a 'corresponding',^(*) formulation in another natural language, etc. Since the 'pre-processing' in all these cases involves what, summarily, is called semantics, it is necessary to stress that in our view of language there is no unbridgeable abyss or opposition between semantics and syntax. In fact, we speak of 'relational' semantics (a term already used by Ullmann⁽¹⁾) when we try to define that part of an expression's meaning that is determined by specific syntactic functions (or Correlators); and we call 'lexical' semantics the attempt - for instance the lexicographer's - to define the meaning of words as separate individual items.

In traditional grammars the lexical items of a natural language are classified as 'parts of speech' according to their generic syntactic functions and/or their morphological characteristics; in correlational grammar^(**) they are classified exclusively according to the actual roles they can play in correlational structures; moreover, while traditional grammars operate with about a dozen different ge-

* Note that since we are interested, not in developing a rigid logically formalised theory of grammar, but in developing a flexible operational system for the automatic interpretation and handling of natural-language sentences and text, we do not require formal instruments for the determination of 'meaning', 'interlinguistic correspondence', 'synonymy', etc.; for our empirical purposes the consensus of proficient language speakers is the relevant and sufficient criterion.

** Although both theory and practice of Correlational Grammar have been drastically modified in our empirical applications, much of the terminology and the basic concepts derive from the pioneering work of Silvio Ceccato, whom I was fortunate enough to have as friend and teacher.

neric syntactic functions, our correlational English grammar distinguishes several hundred correlators (in the present version of our automatic parser⁽³⁾ there are 350).

The list of these correlators is the result of intuitive analysis of English texts, continual refinement by means of insights gained in the translation of English sentence structures into other languages^(*), and by experiments with an automatic parsing procedure whose output, or lack of output, inevitably demonstrates the shortcomings or, as we prefer to put it, the degree of completeness reached by the system's grammar. In our terminology a 'correlator' is a connective function which links two pieces - either words or word combinations - and thus forms a unit we call 'correlation'. Correlators are divided into several types, the main distinction being made between implicit correlators, which are indicated in the sentence merely by the juxtaposition of the items they link, and explicit correlators, which are indicated by specific words (mostly prepositions and conjunctions).

To serve as a valid tool in the parsing or interpretation of sentences, a correlator must not only have an individual code number ('Ic', or correlation index), but the particular relation it establishes, between the two items it correlates, must be characterised or explicated. This explication may be an ad hoc description, an illustrative paraphrase, a suitable transform à la Chomsky⁽⁴⁾ or Fillmore⁽⁵⁾, or a symbolic expression devised along the lines of function symbols in logical calculus⁽⁶⁾.

* Translation frequently helps to pinpoint relational ambiguities which, although relevant in the disambiguation of other structures, are not immediately perceived by the monolingual speaker or reader.

As samples of correlator explication, here are two (one concerning an implicit correlator, the other an explicit one) taken from the correlator list which has been implemented in our operational system:

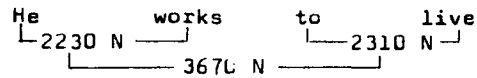
Corr.No.	Description
3670 N	generic type: verb phrase complemented by infinitive; <u>Explication</u> : the gram.subject is the actor of the infinitive activity; the infinitive specifies the <u>purpose</u> of the subject's primary activity. e.g. "He works to live", "I braked to avoid the child"; but <u>not</u> : "He began to live" (3430 N), "I had to avoid the child" (3410 N); <u>nor</u> : "He was eager to live" (3350 N), "I was clever to avoid the child" (7012 N).
0243 E	generic type: BY, temporal limitation; <u>Explication</u> : the item on the right of "by" specifies the point in time <u>after which</u> what is asserted by the item on the left of "by" is to be considered an accomplished fact. e.g. "You will be paid by Christmas", "He had left by 1865"; but <u>not</u> : "You will be paid by cheque" (0257 E), "He had left by the back door" (0251 E); <u>nor</u> : "You will be paid by the treasurer" (0247 E), "He had left by sheer determination" (0255 E).

It should be clear that the 'Explication' is intended to describe as univocally as possible the relation qua relation without regard for the particular items (*) that are

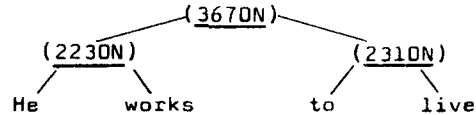
* In practice it is, of course, not always easy to formulate the explication without reference to particular items; but as long as the analyst remains scrupulously aware of the intention, this is not a serious impediment and, in our experience, a sufficiently general formulation can always be found sooner or later.

eligible to be linked by that relation. Correlator analysis and the explications this analytical investigations aims at are consequently what we may call 'the semantics of relations'.

The term 'correlation' refers to the word-combination a correlator produces, i.e. the ternary unit consisting of one specific correlator and the two items (words or themselves word-combinations) that are linked by it; and we represent correlations - for instance the ones operative in the first phrase sample) in this manner:



which corresponds to a conventional tree-structure with labelled nodes:



In order to recognise this structure in the linear input string of the sentence, first each of the four word items must be characterised as a possible candidate for linkage by correlators 2230 N and 2310 N respectively (*); and both these correlations have then to be characterised as possible candidates for linkage by correlator 3670 N. This characterisation is achieved by correlation indices (or Ic's) assigned to the individual items and specifying (a) the code numbers of the correlators by means of which

* The correlator numbers we use were originally significant as to the kinds of relation; continual corrections and additions, however, have thoroughly obliterated this significance and the numbers are now neither consecutive nor characteristic.

the item can be linked to other items; and (b) whether the item can occupy the right-hand (RH) or the left-hand (LH) place in the specified correlation.

(The process of assigning Ic's, obviously, is of one type when the item to which the Ic's are to be assigned is a pre-established vocabulary word, and of quite another when the item is a correlation formed in the course of the analysis procedure; in the first case the word's string of Ic's is the result of a priori assignation; in the second case it is the result of a dynamic process called 'reclassification', implemented by means of an intricate system of rules which take into account the correlator responsible for the made correlation as well as the individual character of the pieces correlated in the particular instance; for a full discussion of the operational reclassification procedure see ref.No.7).

From the sample phrases given under correlator 3670 N it is evident that the two partial phrases constituting the right-hand piece of the listed correlations, i.e. "to live" and "to avoid the child", must bear RH-Ic's not only of correlator 3670 N but also of correlators 3430 N, 3410 N, 3350 N, and 7012 N, while the partial phrases constituting the left-hand piece of the two correlations 3670 N will bear only the LH-Ic of correlator 3670 N and not those referring to the correlators operative in the remaining four samples. (In the sample phrases given under correlator 0243 E, the situation is symmetrically inverted, i.e. the two partial phrases on the left will bear only the LH-Ic of correlator 0243 E, while the partial phrases on the right must bear RH-Ic's of the other four listed correlators as well.)

This assignation of Ic's automatically precludes the

from the one required to decide whether two given items are compatible as RH and LH pieces in one particular syntactic structure⁽⁸⁾.

The question of predicability - especially if elaborated in the way suggested in an essay by Fred Sommers⁽⁹⁾ - is a case in point. Whether phrases such as:

"blue grass" or "cerise ideas",
"the indigestion of angels" or "the wings of the morning",

are acceptable, acceptable only metaphorically, or not acceptable at all, is apparently not a relational question; the relation - a kind of 'appurtenance' - does not seem to change, nor could we say that any of the four phrases is unacceptable because one of its items is such that it can never be related in that way; since we have no objection when the same relation is asserted in "cerise paper" or "clear ideas", we can only conclude that there are certain items to which certain properties or things cannot be said to appertain and that, therefore, it is a question of lexical semantics.

If the semantic analyses initiated by Ceccato in the 1950's^(*) were pushed further, it seems likely that a relation such as 'appurtenance' could be demonstrated to incorporate (i.e. to have confused) a number of specifiable subrelations; and, once we had isolated these subrelations, much if not all of what we now, for the lack of demonstrable distinctions, have to call 'lexical ambiguity' could perhaps be resolved relationally in a satisfactory way.

* called 'operational' semantics, by which was meant the analysis of the mental operations that lead to the formation of the items (or concepts) designated by words (cf. ref. No. 10).

Thus it may indeed be so, that our need to resort to a static, non-relational semantic classification of items, in order to interpret phrases and sentences, is only the measure of our ignorance concerning the basic character and composition of relations; and that, eventually, it will become possible to derive a comprehensive and foolproof general semantics from the investigation of relational conditions as they manifest themselves in our actual use of language.

Let me try to make these conjectures a little less obscure. One area of English grammar that has given considerable trouble to analysts is that of a string of constituents which, in traditional terms, would have the specification:

nominal + to be + adjective + infinitive

A survey of contemporary text shows that this string, with different lexical items and in different contexts, gives rise to ten different relational interpretations (*). Explicating the relevant relations, we get the following listing:

A "John is easy to please"

Paraphrase: to please John is easy

Explication: the gram.subject is the object of the infinitive activity; the adj. specifies an aspect (adverbial) of the infinitive activity as enacted by the given subject.

B "John is eager to please"

Paraphrase: to please is what John wants to do

* I do not wish to claim that ten is all and that no others are possible; but these ten are the interpretations we came across in one year's conscious scanning of everything contemporary we happened to read.

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies the subject's disposition towards the infinitive activity, and this activity is merely envisaged.

C "John was slow to understand"

Paraphrase: John was slow about understanding

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies an aspect of the subject's performance.

D "John is likely to leave"

Paraphrase: that John leaves is likely

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies an assessment of the activity's incidence (i.e. whether or not it will take place).

E "John is clever to leave"

Paraphrase: to leave is clever of John

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies an assessment (regarding the subject) based on the subject's enacting the given activity.

F "John is young to go to school"

Paraphrase: John is young for going to school

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies an assessment of the subject's adequacy (or inadequacy) as actor of the given activity.

G "John is heavy to lift"

Paraphrase: John is heavy with regard to being lifted

Explication: the gram.subject is the object of the infinitive activity; the adj. specifies an aspect (adjectival) of the subject as object of the given activity. (Note: the paraphrase given for type A is impossible here; and if an ambiguous adjective occurs in G, the construction determines its meaning; e.g. "mushrooms are good to eat" requires the interpretation "good"='pleasing', since it does not mean "to eat mushrooms is good for you", where

"good"='beneficial').

H "John is sad to go away"

Paraphrase: to go away causes John to be sad,

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies the subject's state which is a reaction to the given activity.

I "John was critical to upset the speaker"

Paraphrase: John was critical in order to upset the speaker

Explication: the gram.subject is the actor of the infinitive activity; the adj. specifies a deliberate attitude of the subject's, and the infinitive specifies the purpose of the subject's attitude.

(Note: there often is an irresolvable ambiguity between type I and type E; e.g. "the dog was clever to get the biscuit" may be interpreted as I, 'the dog was being clever in order to get the biscuit', or as E, 'it was clever of the dog to get the biscuit')

J "It is sad to go away"

Paraphrase: to go away is sad

Explication: the nominalised infinitive is the subject of the sentence; the "it" functions as subject marker; the adj. specifies an evaluation of the given activity as event.

As far as relational analysis goes, this discrimination of types is fairly satisfactory (although, to be really solid, it would require the detailed definition and coherent application of the terms used in the explications, many of which, e.g. 'aspect', 'assessment', 'attitude', etc., are still rather vague). when we come to a sentence-interpretive procedure, however, such a listing of relational possibilities does not get us anywhere, unless we are able to provide each type with some criterion by means of which we can recognise it in the input text.

In an attempt to discover some such criterion, we as-

sembled a corpus of about 100 relatively frequent adjectives from a recent compendium of English word frequency⁽¹¹⁾ and examined their individual possibilities to function as acceptable constituents in the ten types of construction (for a complete report on our findings, see ref.No.12). The results of this investigation were compiled in the form of a matrix, showing for each adjective the types of construction in which it can occur.

Summarising some of the observations that could be made regarding that matrix, we can say:

a) the adjectives that fit construction type D (viz. certain, expected, known, likely, said, sure, unknown, unlikely) do not occur in any of the other constructions.

b) the adjectives that fit construction type H (viz. content, glad, happy, proud, sad, satisfied, sorry) do not occur in any of the other constructions - with the exception of sad, which can occur also in type J.

c) the adjectives that fit construction type B (viz. able, afraid, anxious, careful 2, desirous, eager, fit, mad 2, prepared, ready, reluctant, wild 2, willing, unable) do not occur in any of the other constructions - with the exception of fit and ready, which can occur also in type G.

(Of the other adjectives of the corpus, approximately one quarter fits only one construction, one half fits two constructions, and one quarter fits three; but since many of these adjectives have more than one meaning - e.g. good 1 = 'pleasing', good 2 = 'beneficial', good 3 = 'moral' - the relational listing of the individual meanings is not an immediate help in the disambiguation of a given string. Nevertheless, it is a step forward from the position where every adjective has to be considered a potential candidate for all

ten constructions.)

The groups of adjectives given under (a), (b), and (c), on page 13, constitute extensional definitions, within the selected corpus, of adjective classes, and we can now examine each of these groups to see whether an intensional definition can be derived from it.

Group (a) obviously has a common semantic element which could be described as 'assessment of probability and/or actuality' (of the item to which the adjective is applied);

Group (b) has a common semantic element which could be described as 'a temporary state of mind usually associated with a specific cause';

Group (c) has a common semantic element which could be described as 'attitude or disposition towards an event'.

I should like to stress that we are at the beginning of this kind of investigation and are presumably still rather clumsy in formulating valid definitions of semantic elements; what is relevant in this context, however, is not the efficiency or reliability of the definitions we tentatively formulate, but the fact that semantic definitions can be derived at all from word groups compiled on the strength of relational considerations. What we are, in fact, trying to show, is that the analysis of the relations found to obtain between the items of phrases or sentences leads, first, to an extensional, and eventually, to an intensional semantic classification of the lexical items constituting these phrases or sentences.

This particular sector of adjective construction is, of course, not the only area of English grammar which makes it seem plausible that semantic classifications of lexemes can be derived from empirical grouping according to their rela-

tional behaviour. The range of relations expressed in English by prepositions is extremely fertile in this regard, but since it is also extremely wide, we have not yet brought our survey of it to a definitive conclusion. Partial results⁽¹³⁾, however, indicate that, here too, relational semantics successfully absorbs a great deal of what, hitherto, was considered lexical or unsystemic.

A comprehensive study of the verb-object relations (a still poorly defined area in our operational system) promises to yield the perhaps most convincing confirmation of our thesis.

Even a very superficial examination of a transitive verb and the grammatical objects that occur with it, shows that the way in which the two are related may vary widely. If, for instance, we take the verb "to pay", we find that:

- 1) in "He paid the driver" the subject gives up something of economic value (e.g. a sum of money); and the object specifies the receiver;
- 2) in "He paid his bill" the subject gives up something of econ. value; the object implies a specific amount and that this amount is due to some not further specified (but specifiable) remote entity as the rightful receiver;
- 3) in "He paid fifty dollars" the subject gives up something of econ. value; the object specifies the amount.

Leaving aside the rarer and the more metaphorical uses of the verb (such as: to pay attention, homage, a visit, etc.) we can now take a corpus of nouns, try each one of them as object of "to pay", and determine which of the three described relations it fits. This will extensionally define three noun classes (not necessarily mutually exclusive, because some of our nouns may fit into more than one of the relations) for which we can then tentatively formulate intensive semantic definitions:

- 1') items that can act as receiver of items having an economic value; e.g. boy, butcher, college, court, driver, girl, tax collector, tailor, etc.;
- 2') items implying a definite econ. value and the fact that the implied amount is due to someone; e.g. dues, fare, fee, fine, postage, tax, etc.;
- 3') items indicating a specific econ. value; e.g. any numeral followed by an indication of currency, and pronominal expressions such as "a lot", "little", "much", etc.

The verb "to pay" is also doubly transitive, i.e. it can be constructed with a dative and an accusative object in one phrase, e.g.:

- 4) "He paid the butcher ten dollars" (in which "the butcher" is the receiver).

If we now test our corpus of nouns in this construction, we find, first, that only the items listed in group 1' can be used as dative object and that they cannot be used as accusative object; second, that if items of group 2' or 3' are used as direct object, the dative object always plays the part of 'receiver'; third, we find that we have to consider some new items (occurring only in conjunction with a dative object) which are not members of the three listed classes and which, moreover, change the role of the dative object, as for instance in:

- 5) "He paid his driver a holiday" (where "his driver" is the beneficiary of the subject's act).

We thus get a fourth group of possible objects; their intensional definition is less obvious, but we can tentatively put down:

- 5') items intended for personal consumption; e.g. a drink, a term at college, a meal, a trip round the world, a vacation, etc.

One point that is of special interest in these still very crude and incomplete results of relational analysis and semantic classification, is the fact that, although the classes of 'receiver' and 'beneficiary' are co-extensive in the corpus, they do not seem to create ambiguities, since the particular role of an item that is ambivalent in this respect, is, in any given example, determined by the classification of the direct object; in other words, if the direct object belongs to either group 2' or 3', the dative object plays the part of 'receiver'; if the direct object belongs to group 5', the dative object plays the part of 'beneficiary'.

This rule, incidentally, seems rigid enough to deal with at least some phrases which, on a purely experiential basis (i.e. 'knowledge of the world'), would have to be rated rather odd or unlikely. If, for instance, we came across the sentence "She paid her lover a week of clams", we might be uncertain how to interpret precisely "a week of clams", but we would have no doubt that it had to be something her lover could personally consume. - And there is considerable reassurance to be got from the implication that (at least in some cases) the logic of relational semantics is more powerful than the statistics of factual experience.

We could adduce many more examples of transitive verbs and relevant object classification, but at the present stage of the analysis this would add little: the results are invariably suggestive, even indicative, but they cannot be considered conclusive. Therefore, I shall merely summarise what we expect from these investigations.

On the basis of work accomplished during the last 18 months, it appears that the semantic classes derived from relational analysis are recursive and that their num-

ber will remain usefully smaller than the number of classified items. There will, of course, be many more semantic classes than grammars have hitherto contemplated - but given the versatility of natural language, this should not really surprise us; nor is there any need to be particularly pessimistic about the possibility of implementing such a voluminous and intricate data base in a computer system. We all have seen how the early computational linguists' worries concerning storage capacity and processing speeds have ~~been~~ been made to appear anachronistic by technological progress, and this progress does not yet seem to be anywhere near its ceiling.

As to the theoretical implications of our kind of language analysis, I should like to put forward one suggestion. Assuming that we can derive (and the material presented here does imply precisely this) a satisfactory semantic classification of lexical items from their relational properties defined in terms of an adequately differentiated syntax, it may be more profitable (and, perhaps, also more correct) to view syntax and semantics, not as a pair of mutually exclusive opposites, but rather as the axes of a continuum of meaning; every semantic element, particle, feature (or however we want to call it) would, in such a frame of reference, have both a relational and a lexical coordinate - which would not only make it possible for us to discuss one and the same item from two points of view without contradiction, but, I believe, it would also be a useful advance towards an economical representation of linguistic data in computational procedures.

R e f e r e n c e s

- 1) Stephen Ullmann, *The Principles of Semantics*, Blackwell, Oxford, 1957.
- 2) Silvio Ceccato et al., *Linguistic Analysis and Programming for Mechanical Translation*, Feltrinelli, Milan, 1960, and Gordon & Breach, New York, 1962.
- 3) E.v.Glasersfeld and P.P.Pisani, *The Multistore System MP-2*, Georgia Inst.for Research, Athens (Georgia, U.S.A.), 1968, and *The Multistore Parser for Hierarchical Syntactic Structures*, to be published in the Communications of the ACM, 1969.
- 4) Noam Chomsky, *Aspects of the Theory of Syntax*, M.I.T. Press, Cambridge (Mass.), 1965.
- 5) Charles J. Fillmore, *Entailment Rules in a Semantic Theory*, Ohio State University, Columbus (Ohio), 1965.
- 6) Hans Reichenbach, *Elements of Symbolic Logic*, Free Press Paperback, New York, 1966.
- 7) E.v.Glasersfeld, P.P.Pisani, J.Burns, B.Notarmarco, B. Dutton, *Automatic English Sentence Analysis*, Reports ILRS T-11 and ILRS T-14, IDAMI Language Research Section, Milan (Italy), 1965/66.
- 8) Jehane Burns, *A Scheme for Semantic Controls in Automatic Sentence Analysis*, Literary and Linguistic Computing Centre, University of Cambridge (England), 1968.
- 9) Fred Sommers, *Predicability*, in Max Black (editor), *Philosophy in America*, Cornell University Press, 1965.
- 10) Silvio Ceccato, *Un Tecnico fra i Filosofi*, vol.I and II, Marsilio Editore, Padova (Italy), 1964/66.
- 11) Henry Kučera and W.N.Francis, *Computational Analysis of Present-Day American English*, Brown University, Providence, R.I., 1967.
- 12) E.v.Glasersfeld and B.Notarmarco, *Some Adjective Classes derived from Correlational Grammar*, to be published in American Speech, 1969.
- 13) E.v.Glasersfeld, *An Approach to the Semantics of Prepositions*, Proceedings of the Symposium on Computer-related Semantic Analysis (Las Vegas, 1965), Wayne State University, Detroit, Mich., 1967.

Semantics and the Syntactic Classification of Words

A B S T R A C T

Traditional grammars classify words according to generic syntactic functions or morphological characteristics. For teaching humans and for descriptive linguistics this seemed sufficient. The advent of computers has changed the situation. Since machines are devoid of experiential knowledge, they need a more explicit grammar to handle natural language. Correlational Grammar is an attempt in that direction. The paper describes parts of correlational syntax and shows how a highly differentiated syntax can be used to establish word classes for which an intensional semantic definition can then be found. It exemplifies this approach in two areas of grammar: predicative adjectives and transitive verbs. The classification serves to eliminate ambiguity and spurious computer interpretations of natural language sentences.

Author's address:

Ernst von Glasersfeld
Georgia Institute for Research
711, C & S Bank Bldg.
ATHENS, Ga., 30601, U.S.A.