

On Explanatory Definition of English Clause

E. M. Pogossian

Institute for Informatics and Automation Problems
of NAS of RA and YSU

Abstract

An attempt of explicit definition of the scheme of basic affirmative indicative clauses and related syntax is described. Central idea is based on an interpretation of a pragmatic role of some basic constructions of English in the communication process and the fact of partition of all verbs into three classes: DO (active), BE and HAVE (stative). Appropriate modifications for some concepts are suggested, which, seems, make more transparent the syntax presentation. That explanatory approach may improve the efficiency of English teaching in a deductive way, the parsing of phrases for different computer applications, as well as the pattern formation and recognition of a linguistic approach.

1 Introduction

1.1. Children catch the spirit of their native language very quickly listening to samples of discourse. Even if they make mistakes in the vocabulary, forms of the verbs, etc., they reproduce the main laws of language in the right way. Such efficiency in an inductive way of learning may be explained, particularly, by the highly impressionable abilities of children, as well as by the fact that these basic laws are not very complicated.

For foreigners, educated but deprived of both a natural language studying environment and an ability to memorize intensively the use of the deductive ways of learning, particularly understanding and using the grammar of the language might become effective and helpful also. That process could become more efficient, if the grammar description is organized with explanation of emergency of its rules and shell by shell, to present the information about the structure of the language in the order of priority for communication efficiency. This could be done by portions similar to the portions children acquire in their step by step development. In that process primary and very essential information gives the syntactical structure of the clause. It allows to identify a subject and an object of some activity in the World, the time, aspect, voice, mode, subordination of components, etc., of that activity. Although the concept of a basic clause must be central for the syntax in the way that the whole variety of clauses could be derived from it in a natural way there is no its explicit and comprehensive definition. E.g. its very difficult and unexplainable for foreigners to catch the idea of appearance of a "did" in interrogative and negative English clauses.

To try to extract the hierarchy of grammar knowledge, we need to have guidelines for our choices or interpretations. For that we first remind ourselves of some pragmatic functions of a language in the process of a human survival as a Living Reality in the World. A more detailed analysis of these functions may be found, for example, in [1].

Then we give an explicit definition of a scheme of English clause underlining the primacy of its nomination, subject intentions and relations: *do*, *be* and *have*, identification functions. Finally, a

modification in the presentation of some basic constructions of English syntax is suggested, followed by a brief illustration of their influence.¹

2 Language as a Survival Instrument

Living Realities (LR s) are those that can preserve themselves, i.e. their main functions, in a extremely changeable and fearfully uncertain World. In the same sense, LR s are realizations or implementations of survival methods, i.e. knowledge of how to survive. Thus, "living" as a process is a cognition. In an individual (not community) level it means continuous reconstruction, renovation of an LR to react adequately to actual or possible changes in the World.

Knowledge of an LR, in fact, includes a persistent part in a form of the actual construction of an LR itself - a result of evolution, and more flexible adaptive part in a form of memorized descriptions and instructions acquired mainly during the individually period of life as an LR (compare with hardware and software of computers).

That persistent, genetic part of knowledge is so large and complete that in an ideally stable and unchanging environment, an LR might have no need to renew it. Only breakdowns in the expectations of an LR in anticipated and real events in the World cause the need to memorize the descriptions of situations and their resolution by actions.

The coordinated behavior of LR s is also an instrument for survival. Communities of LR s make possible results unachievable for individual ones. To coordinate their efforts, LR s need in exchanging of relevant information about themselves and the World. For that they prescribe identifiers or names, to the important realities of their life and organize their collective behavior by communication, i.e. transmissions from one to another of relevant samples of names to identify the situations and their own relations to them. Due to the concrete conditions, each community creates its specific communication symbolic system or language, with its own set of useful words (vocabulary) and rules of combination of them - the syntax of the language.

Before discussing these categories of language in more detail, let's remind ourselves that LR s symbolize only socially important parts of their knowledge, which are necessary for communication primarily for production or exchange activities.

Reflective abilities of any language to simulate parts of the World - realities, at each moment of its development are restricted mainly by the composition of its vocabulary. That relates only to a part of knowledge of the World involved in each LR. What LR s symbolize in communications is only the top of the iceberg of their real knowledge.

LR s mainly transmit from one to another not their knowledge, but only the names (the signs) to trigger in the other communication partners a desirable state, which could allow them to recognize and operate in a problematic situation in an expected way. Here is, also, a possible explanation of the fact that actual results in simulation by computers of the human's intellectual abilities are mainly at the level of a "high school" education. They are impeded just as we try to simulate "primitive" abilities of LR s, which are common for all of them. Understanding of this fact implies also some regulation in our future attempts to simulate intelligence by computers [1].

To be more convincing let us discuss the question only for English.

3 A View of English Phrases

¹ Initial version of this concept was published in [7].

3.1. Thus, the primary function of English discourse is a triggering of the participants of the discourse into a desired state of acceptance of the problem situation. That is attained by transmitting to them some phrases of the language, descriptions, i.e. systems of names of the World's realities.

It is important to note that the structure of the basic descriptions is standard: two ordered nominators of realities and some of a few typical basic relations between them (modals and verbs: *do, be, have*). If the situation is described from the human's point of view, *modals* have a sense of human intentions, desires, duties, abilities, etc., to achieve something, to inform others, etc., in the problem situation. In other cases they are the result of the prescription to the described realities of imagined relations similar to the human ones.

The rest of those basic relations include the information about a type of activity of realities, i.e. indicating the fact of presence of their activity (*do*) or existence (*be*), possession, belonging (*have*) of some realities to others..

These basic relations, we suppose, historically played the role of the primary importance in the development of human communication and are reflected in the construction of the basic structure of English syntax - scheme of a clause. Other discourse information related to a time or period of events, a manner by which a description is focused, an identification of an action's doer and receiver, etc., is coded by concretizations of that scheme of a clause using forms of times, aspects, moods, voices, etc. It may be considered as a secondary level of information relative to the main one - nomination and the basic relations information.

3.2. Let's illustrate our view by examples of basic Affirmative Indicative clauses at a Progressive Aspect of the Active voice.

We say: *The car is stopping near the policeman. Driving car was stopping near the policeman at the time. The car will be stopping near the policeman. We should be stopping near the policeman.*

These clauses have a simple structure: Nominator1 Form(*be*) Nominator2 or Nominator1 Form(*Modal*) *be* Nominator2.

Here: *the car, driving car, we, stopping near the policeman*, are Nominators. They name some realities and their positions in the phrase inform us about Subject /Object subordination between them.

Time and Person information is coded by forms of *be* (*is, was*) or forms of modals (*will, should*). The type of Aspect (here Progressive) is coded by Present participial of the verb *stop: stopping*, placed at the first position of Nominator2.

We insist that a slight generalization of the above structures represents the scheme of Indicative clauses at the Active voice. Let's substantiate this assertion.

4 The Scheme of Basic Indicative Clauses in the Active Voice

4.1.1. The basic function of a word - string of letters in an alphabet, is nominative. This nominative function we differentiate in accordance with a static or dynamic nature of realities that they describe and a role they play in descriptions: principal or auxiliary. These functions depend from position of the word in a description or its context but may be an inherit too. We differentiate, particularly, the following main types or classes of words and their functions: Nouns (Nn), Adjectives (Adj), Determiners (Dt), Adverbs (Adv), Prepositions (Pre), Pronouns (Pro), Conjunctions (Cnj), Verbs (V or v), Modals (M).

Thus, we identify words as representatives of some classes but at the same time may prescribe them functions or role of another classes depending from their positions in the phrases.

For example, we identify Prepositions on, in, at, etc. and prescribe them genetic, primary function in expressing relationships of things in the space: movement, position, etc., in their literal senses. Some of those words along with those functions may play other roles in phrases depending from a position there, e.g. be adverb particles (come in, go out) or a coherent part of an idiom (phrasal verbs) with a meaning different from Preposition one..

Words may have person, degree, time and aspect forms. All Forms (Fr) of each type of word may be listed or enumerated by generation rules.

The totality of all the words of these types and their forms is the vocabulary of the language.

Nominators, in general, are combinations of arbitrary words of the vocabulary, except modals.

4.1.2. Modals indicate some deontic relation, modality - intention, duty, permission, obligation, etc., between Nominators. Logically, they are the following two-placed predicates: *will (shall), can, may, must, have to, ought to, need, dare*. Some modals have Past forms: *would, should, could, might, had to, needed*.

4.1.3. Each verb V has the following forms (FrV): infinitive (toV), person forms at Present (fPrV), Past (fPtV) and Future (fFtV), present (ingV) and past (edV) participles.

The forms fPr, fPt and fFt we name also Time/Person forms (tp).

The forms of all verbs are given by simple generation rules, except those of about 200 irregular ones, for which forms are listed.

All verbs may be split into three groups:

DO verbs (doV), BE verbs (beV) and HAVE verbs (hV).

DO or active verbs name some of variety of *do* related to activity, process, change, movement, etc., in the World.

BE verbs and HAVE verbs are named also stative verbs.

BE verbs include *be* and the verbs having a sense "being"

HAVE verbs include *have* itself, 19 verbs (see [2]) with a sense of "having", the verbs with involuntary sense and complete mental action that indicate their presence, existence or possession by the Sbj at the moment of communication.

4.2. The scheme (formula, type, code) of affirmative Indicative Basic Clauses of the Active voice, or aIBC, is defined as the following *positional* writing:

sN M dbh vN

or

sN dbh vN

where sN, M, dbh, vN are variables, sN- for a Nominator in the first or in the Subject position, M- for modals or their forms in the second position, dbh- for one of verbs: *do, be, have* or their forms in the third position, and vN - for verb Nominators in the forth position, correspondingly. A Time/Person information of the event is coded by using an actually present in the scheme the first component of the triple M dbh V.

dbh component of the aIBC is missed if aIBC expresses a Simple (General) Aspect of the event.

Verb Nominator vN always contains at its first position forms of verbs which provide nominative, Aspect and Voice information but also may code Time/Person one, e.g. for Simple Aspect.

4.3.1. The focus of a message in any clause is the Subject. sN provides its identification by its position and other externally recognizable characteristics.

M codes information of the internal state of the Sbj.

dbh and forms of verbs in the vN specify the relation of the Sbj to the World (W). Namely, dbh identifies the type of that relation in terms of the basic or archaic ones: *be-* for underlying Sbj's being or existence, *do* - to indicate presence of some activity of the Sbj in the W, and *have-* for saying about Sbj's possession of something of the W.

Note, that these archaic relations *do*, *be* and *have* also correspond to the three fundamental instruments of human survival: human's production (*do*), human's own evolution at genetic or knowledge-based ways (*be*), and exchange activities for needed items (*have*).

Forms of verbs at the first position of the vN provide nominative and Aspect /Voice information for a concrete activity of the Sbj. They also may code Person /Time information if both M and dbh are missed.

4.3.2. Let's note that the difference of the aIBC from known scheme: (Noun phrase) (Verb phrase) [2], is in extraction of the M and dbh from Verb phrase and positioning them at the scheme in a explicit way.

As we can see also the aIBC concept will coincide in fact with the concept of clause in [2], if we accept that active verbs may be always interpreted as a transitive. (The notion of a transitive stative verb seems very unnatural. I.e. we consider active verbs transitive and the stative ones intransitive.) Even if this assertion turns out to be false for some verbs, following it in the presentation of the syntax will clarify some arguments for its use as the first level of approximation.

4.4. Before completing the definition of aIBC for compound and complex clauses, we illustrate at first aIBC for a Simple Aspect of Active voice, then demonstrate advantages of the aIBC presentation for some syntactic constructions of the Indicative mood.

5 The Scheme of Basic Affirmative Clauses for a Simple Aspect

Basic clauses in a Simple Aspect of the Active voice may be naturally interpreted in frame of the aIBC.

Indeed, the interpretation of phrases like: *Driving at high speed may be dangerous. We should have a great benefit. He will do the work.*, where *be*, *have* or *do* are used explicitly, is the following: corresponding dbh component in the scheme is missed due its actual redundance - repetition in the vN.

Phrases such as: *A boy runs. The car moves. I do it.*, etc., which underline an activity and use active verbs, may be interpreted as equal to phrases: *A boy does run. The car does move. I do do it.*, with a scheme: sN fPr(*do*) (do)N. In actually used phrases redundant *do* is missed and fPr functions are driven to the verb in the vN.

Phrases: *We exist many years. It reminds me of the school.*, with stative BE verbs, different from *be*, may be interpreted as equal to phrases: *We are exist many years. It is remind me the school.*, with archaic scheme: sN fPr(*be*) vN, and an actual one: sN fPr(v)N, with missed the redundant *be* in the dbh position.

At last, the phrases: *He belongs to the group. I hear you. We believe him.*, with stative HAVE verbs, different from *have*, have an archaic scheme: sN fPr(hv) vN, and an actual one: sN fPr(v)N, with missed the redundant *have* in the dbh position.

6 An Outline of Syntax for the Basic Indicative Clauses

6.1 We are going to demonstrate that English syntax of Indicative clauses may be presented in compact and enough complete form if operate with aIBC.

Let's start with consideration of Aspect forms of the Active voice. To this end we will base ourselves on their transparent explanation in [3].

Information on activity or events in the World, coded in phrases, in their essential include the following:

- Time period T of an activity: Present, Past or Future,
 - an Aspect of the activity relatively to the focus point T [3].
- There are the following types of Aspects:
- Progressive: the activity is continuing now at the focus point,
 - Perfective: the activity started at some period before the focus point in the past and just ended,
 - Perfective +Progressive: the activity started at some period in the past before the focus point and is continuing now,
 - General, or Simple: the activity exists also after the focus point of consideration..

The Aspect information at time T is coded in the vN part of aIBC at the following simple way:

- for Progressive Aspect: vN = ingV Obj,
- for Perfective Aspect: vN = edV Obj,
- for Perfective +Progressive Aspect: vN = edBE ingV Obj,
- for Simple Aspect: vN = v Obj or vN = tpV Obj (if there is no modal in the clause)..

Correspondingly dbh is equal to the following :

- for Progressive Aspect dbh = be,
- for Perfective and Perfective/Progressive Aspect dbh = have,
- for Simple Aspect dbh is omitted, as a rule.

Time /Person information is coded by actually present first component of the triple M dbh V in the clause .

Let's remind that only DO verbs usually have progressive Aspect.

6.4. For clauses: *A boy can make a shelf. A boy makes a shelf. A boy has made a shelf.*, clauses in Passive will be: *A shelf can be made by a boy. A shelf is made by a boy. A shelf has been made by a boy.*

As we can see a Sbj is "passive" in the Passive voice (in opposite to the Active voice), i.e. Sbj is not a Doer of the action but only Receiver of it.

There's no normally Passive for clauses with stative verbs [2, p.226]. E.g. for clauses like: *George has sat down on my hat. He is my friend.*, etc. Haven't Passive, too, clauses with Perfective /Progressive Aspect, like : *A boy has been making a shelf.*, [2].

The Passive is formed by specific coding of an Aspect information for active verbs. Namely, it is using Aspect forms of the verb *be* in Active voice and edV participial of the current verb.

Thus, aIBC in Passive will be :

- for Progressive Aspect: sN [M] be ingBE edV [by] Obj ,
- for Perfective Aspect: sN [M] have edBE edV [by] Obj,
- for Simple Aspect: sN [M] be edV[by]Obj.

In Passive clauses a part: "by Obj", often is omitted and *by* may be replaced by prepositions : *with, in, at*, etc.[2].

6.3. For further consideration we'll use the following notification.

Let denote by Y an actually present first component of the triple $M\ dbh\ V$ in a $aIBC$ by Y_n the negative of Y .

We present a clause in a form $C = sN\ Y\ C'$, too, where C' complement $sN\ Y$ to the complete $aIBC$.

Time/Person information in clauses is coded by Y . Negative forms Y_n for each Y are generated by known easy rules.

Let's remark that for all Active and Passive voice clauses Y is equal either a Modal, or BE, or HAVE. The only exclusion are clauses in a Simple Aspect of Active voice when used verbs are different from BE and HAVE. At those clauses dbh component are omitted for reasons interpreted above. In further analysis and definitions for such clauses below we will suppose that Y denotes omitted verb DO (see also the Note at the end of this paragraph).

6.4. Negative clauses, or $nIBC$, are defined as clauses formed by replacing Y by Y_n in clauses $sN\ Y\ C'$.

6.5. Interrogative, or question, clauses ($qIBC$) for the first degree of their approximation $qIBC$ are the following:

- $C?$ and $nC?$ with rising intonation, if we are pronouncing them.
- $Y\ Sbj\ C'?$ and $Y_n\ Sbj\ C'?$
- Question tags: $Sbj\ Y\ C'$, $Y_n\ ProSbj?$ and $Sbj\ Y_n\ C'$, $Y\ ProSbj?$

Here $ProSbj$ means Pronoun corresponding to Sbj .

- $Wh\ Y\ C'?$ or $Wh\ Y_n\ C'?$ if the question is about the Sbj .

Here Wh corresponds to one of Sbj question words: Who, What, Which, Whose.

If $Wh = Whose$ and $N(Sbj)$ is the Noun of the Sbj , then $q\ aIBC$ will be: $Wh\ N(Sbj)\ TFrV\ C'?$.

- $Wh\ Y\ Sbj\ TFrV\ C'/P?$, if the question is about any part P of the $TFrV\ C'$. Here Wh is one of the corresponding P question words: Who, What, Which, Whose, Where, When, Why, How, How many, How long, etc., C'/P is the rest of C' without P . [2]. For Y_n a definition is similar.

6.6. Note. To construct interrogative or negative for the above exception clauses in a Simple Aspect of Active voice standard syntax recommends a use of DO for the verbs of all types: active and stative. In our logic BE and HAVE verbs had to use omitted $tpBE$ and $tpHAVE$, correspondingly, instead of $tpDO$. For example, the interrogatives for: *We exist many years.*, might become: *Are we exist many years?*, or *We exist many years, aren't we?*, instead of: *Do we exist many years?*, or: *We exist many years, don't we?*.

At the same logic: *He belongs to the group.*, might become: *Has he belong to the group?* or: *He belongs to the group, hasn't he?*, instead of: *Does he belong to the group?* and: *He belongs to the group, doesn't he?*

As we can see in these examples the use of DO, in opposite to BE and HAVE, adds some new sense of activity absent in the initial $aIBC$. This fact may be an argument also for splitting the verbs into DO, BE and HAVE classes.

7 Indicative Clauses

7.1. We have defined above all types of **basic** clauses for the Indicative mood. Let's make an inductive step to complete that definition for complex and compound Indicative clauses. Here we outline it based on the examples from [4] only. The full definition will demand more details.

Given affirmative or negative IBC clause C (aIBC or nIBC) we name:

phrases *that C, what C, how C, etc.*, as a **Noun phrase C** (NnPh C),

phrases *that C, which C, who C, whom C, etc.*, as an **Adj phrase C** (AdjPh C),

phrases *when C, where C, because C, if C, so that C, etc.*, as an **Adv phrase C** (AdvPh C).

Complex clauses are defined as the following:

if B and C are affirmative or negative IBC clauses in a form: sN [M] dbh vN, then the substitutions:

- of the Noun (Nn) in the Subject sN of the B by Noun phrases: *what C, or how C, etc.*, or Noun in vN of B by Noun phrase: *that C*,
- Noun, Adj Noun or Noun Adj in B by Adj phrase: Noun AdjPhC,
- Adv in some couples with V, toV and ingV in B by Adv phrase AdvPhC

are **clauses**, or **complex clauses**, of affirmative or negative types in a correspondence with the type of B.

B is named a principal clause and C - a subordinate one.

Correspondence of Times and Aspects of the component clauses of a compound clause is subject of special consideration.

This definition is expanded for other basic forms of B and C in an evident way.

7.2. Given affirmative or negative clauses B, C, a text B Cnj C is a **compound clause** where Cnj-s are Conjunctions: *and, but, etc.*

It will be affirmative or negative clause depending upon the combination of the types of B and C.

7.3. *Thus, Indicative clauses are affirmative, negative and question basic, complex and compound clauses.*

8 Subjunctive and Conditional Clauses Coordination of Times and Aspects

Given affirmative or negative clauses B and C, conditional clause is a compound one with Adverbial clause: "if C", in Present Simple and principal one in Future Simple.

Example.

Thus, in conditional clauses component ones have the same Aspect but Times are shifted for one.

Subjunctive clauses preserve the same structure but relates it to the Past of Simple or Perfect Aspects. Correspondingly, instead of *will* they use *would*, i.e. instead of Future Time they use Future in the Past.

Example.

9 Some Recommendations for English Deductive Learning

The above view of syntax, if would be correct, could imply a natural order in English presentation. Namely, first presenting the scheme of a clause, then step by step its forms, followed by indication of the parts of the language that ought to be regularly memorized: vocabulary, exclusions, prepositional verbs, etc. In Appendix we give some examples of that.

10 Conclusion

We suggest to replace accents and modify some definitions in English basic syntax to try to make more explainable its constructions. Namely, we expand nominative role of verbal phrases, more concentrate attention on the role of modals and verbs DO, BE and HAVE, slightly reclassify stative verbs and equalize them with intransitive ones. This modifications allow to set forth the core of syntax in compressed and logically clear form as well as to give complete definition for clauses in all their variety.

As a result the essential of the whole syntax in the form, compatible for its father systematic enrichment, may be easy cached by experienced in a symbolic analysis students from the first steps of their learning and used as a powerful deductive instrument for selfcorrections.

Author has not illusions to cover by some formulas an alive language in whole. But he is convinced in importance of attempts to extract the logical nature of realities. He risked to try it for English based on his experience at language learning essentially in a deductive way and many years efforts to explain knowledge formation mechanisms [5]. Author only tried to demonstrate an advantages of that approach for some syntax constructions in an attempt to attract language professionals to make the final verdict

Nevertheless, he never start until study Terry Winograd ideas in a symbolic systems at the seminars with Hrachia Bajadian and continuous support of his teachers and friends from language department of the American University of Armenia. Specially he thanks Patricia Boyle, Maria Bobrova and Virgil Strohmeyer for their exclusive efforts and patience to advance such terrible student as me.

References

1. Winograd T., Flores F.. Understanding Computers and Cognition. A New Foundation for Design, Stanford, 1989.
2. Close R. A.. A Reference Grammar for Students of English. Longman Group, 1975, 352pp.
3. Asher A..Think About Editing. Heile&Heile Pub.,1993, 334pp.
4. Eckercley C.E.. Essential English for Foreign Students.,Sofia, 1967.
5. Pogossian E.. Adaptation of Combinatorial Algorithms. Pub. by Academy of Sciences of Armenia, 1983, 293pp.(in Russian).
- 6.Reid W. Verbs and Nouns Number in English. A functional explanation. Langman Publ.,1991.
- 7.Pogossian E. On a Transparent Presentation of Written English Syntax. Procieedings of the Intern. Workshop in Computational Linguistics and its Applications (Dialogue'96), Puschino, Russia, May 4-5, 1996, 209-213.