

Ilya Geller
2442 E26th st
Brooklyn NY 11235
igeller@lexiclone.com

Key words:

predicative definition

living language

slang

In Search of a Living Language (re-writing, September 2004).

Abstract.

This article affirms that Living Language is a dialectical sum of slangs which cannot be formalised or objectively understood, since:

1. Living Language and slangs are in a state of constant change: formalization and objective understanding indicate a cessation of change, the fixation of Living Language in time;
2. Understanding Living Language and slangs is always subjective; Living Language carries only information about subjective Truth, never information about Reality.

It is supposed that change is inherent to the slangs of Living Language because of their striving to provide, by means of information, a full, well-rounded description of an unlimited quantity of subjects/objects and their changes within an unlimited interval of time. Such a description is arrived at:

1. By applying the sum of predicative definitions that are primordially inherent to each slang;
2. As well as by including in the slang an unlimited quantity of predicative definitions associatively borrowed by means of synonymic expansion from other slangs that are at least indirectly related to the given slang; where synonymic expansion by predicative definitions represents the way slangs evolve, by:
 - a. Replacing predicative definitions that were a part of the slang before new predicative definitions were included in it;
 - b. As well as by enlarging the general quantity of predicative definitions contained in the given slang.

A Slang.

The Merriam-Webster Online Dictionary defines a slang as follows:

Main Entry: slang

Pronunciation: 'sla[ng]

Function: noun

Etymology: origin unknown

1. language peculiar to a particular group;
2. an informal nonstandard vocabulary composed typically of coinages, arbitrarily changed words, and extravagant, forced, or facetious figures of speech.

I am completely in accord with the view expressed in Merriam-Webster: every area of human activity presupposes the existence of a specific means of communication – a single one for all the subjects of a

given activity – a slang. Any geographically linked (now or in the past) group of people, united by a specialization in certain types of activity that are necessary for the physical and spiritual existence and development of those people, is the carrier of particular slangs.

Slangs don't exist in an immutable form -- they appear and disappear, as an obligatory response to the necessity of making available to many people information about particular types of activity that are necessary for the physical and spiritual existence and development of those people.

Cynicism.

I belong to the Cynical school of thought, the foundations of which were laid by those contemporaries of Plato and Aristotle -- Ecclesiastes and Jeremiah. The basic principle of that school is: "For my people have committed two evils. They have forsaken me, the fountain of living waters, to hew for themselves cisterns, broken cisterns, that can hold no water¹". Translated into modern language and applied to linguistics, it means the following: Living Language and the slangs that make it up are in a constant state of change, and any formal rules imposed on Living Language and its slangs become obsolete even before they have been definitively formulated.

The formula "The dialectical sum of slangs is Living Language" is a philosophical one. Aristotle affirmed² that when one is looking for a true understanding of Nature, one must use Dialectics. But when it comes to scholastic exercises, to empty demagogy, there's no room for Dialectics. That is, Dialectics is the philosophy of progress, of uninterrupted change; while Formalism is the philosophy of what has been ossified into changelessness. For example, the rules and applications of so-called Natural Language, in its current form, are pure demagogy, Formalism.

Indeed, one finds the following definition of Natural Language: "everyday language; in database searching, a natural language search allows the user to type words in the same way that a person normally speaks them³." That is, Natural Language works according to Formal rules, with words that stand by themselves and aren't linked in any way. Or, in other words, Natural Language works with non-predicative definitions. But Living Language is the sum of the predicative definitions in slangs, not of individual words.

Living Language and Slangs.

For the concrete understanding of information, people call certain subjects and/or objects of certain events by Names that are inherent only to given slangs. The Name of an object is always a noun or a pronoun.

As the number of people engaged in certain types of activity that are necessary for their physical and spiritual existence and development grows, so do their slangs become Dialectically included in Living Language. This happens through the mutual enrichment of both Living Language and the given slang: the matchless predicative definitions of the given slang become a part of Living Language, while predicative definitions from Living Language's other slangs become parts of the given slang.

1 **Jeremiah.** Bible. Chapter 2 & 13

2 "...dialectic probes where philosophy seeks understanding, and sophistic is imagined to be science but is not really." **Aristotle**, A new Aristotle reader, Princeton University Press, Princeton, 1995, 1004^B,25

3 http://www.academic.rccd.cc.ca.us/~lib/glossary_lib_terms.htm

That is, the more there are people engaged in particular types of activity, the more a given slang becomes included in Living Language. For example, we've all witnessed the ever increasing inclusion of the specialized slang of Information Retrieval into Living Language, in relation to increased use of the Internet.

"Kernel Sentences" as Predicative Definitions.

It has been suggested earlier that we refer to minimal constructions carrying a basic load of meaning in texts as "kernel sentences". It has also been supposed that kernel sentences must provide an unambiguous, formal description arising from the subjects/objects of the given Names.

But I'm a philosopher. In this case, it seems more expedient to use a philosophical terminology -- the terminology used by Aristotle, Moore, Russell and Wittgenstein, as the most significant philosophers to have dealt closely with the study of language. Those philosophers spoke of predicative and non-predicative definitions for the description of events and changes occurring in subjects or objects of the Universe. Moreover, predicative definitions and kernel sentences are one and the same thing.

Following philosophical tradition, it has been shown that information in textual form is transmitted by means of minimal lexical constructions that are predicative definitions: the articulation of three words, relating to three parts of speech -- substantives, verbs and adjectives. All other parts of speech, with the exception of prepositions and interjections, can be, in some way or other, taken to be substantives, verbs or adjectives, where:

1. A substantive has the meaning of the abstract Name of certain objects and subjects;
2. A verb defines the abstract Name of an action;
3. An adjective is the abstract Name describing objects and subjects in the process of change.

In formalizing the understanding of what a "predicative definition" is, one must note that it almost always consists of the sequence "noun/pronoun-verb-adjective" obtained by means of the technology described in US patent 1.199.067. A verb-adjective pair (or equivalent parts of speech) must also necessarily be present in a predicative definition.

Non-predicative definitions, however, are individual words, words in themselves: as is well known, Bertrand Russell introduced the notion of a "non-predicative" definition, in which what is to be defined is brought in through its relation to a class of which it is an element. For example, "the set of all sets that are not elements of themselves". It is said that the use of "non-predicative" definitions leads to paradoxes, so they should be dealt with carefully.

For example, take the non-predicative definition "blue". What kind of information can be communicated by means of the non-predicative definition "blue"? What does this non-predicative definition mean?

Context and Subtext.

It's obvious that a non-predicative definition is meaningful – that is, it can be understood more or less monosemantically – only as part of predicative definitions which manifest the contexts and subtexts of texts: the sum of predicative definitions, taking into account the relative frequency with which they occur in a given text, are the context and the subtext of that text:

1. The context consists of those predicative definitions where a substantive is used as the abstract Name of objects and/or subjects;
2. The subtext consists of those predicative definitions where pronouns and interjections are used as the abstract Name of objects and/or subjects.

The Expansion of a Living Language's Store of Predicative Definitions by Way of Synonyms as the Mechanism for the Dialectical Linking of Slangs Within Living Language.

A Name is what a subject and/or object is called in a given slang. Other slangs may use different Names to refer to the same subjects and/or objects. In those cases, the Names are synonyms.

A synonymous description of a change in the form of a predicative definition is a description of the change by means of another slang, not the original one.

It is precisely such synonymous predicative definitions that link different slangs together within the structure of a Living Language -- for example, within the structure of Living (as opposed to Formal) Russian, English, French or any other language.

As has been shown by a program using the technology of synonymic expansion of predicative definition⁴, whenever new subjects and/or objects or new, not previously described events appear, predicative definitions are borrowed by association (through the mechanism of synonymic expansion) from another slang or other slangs.

Synonymic expansion of predicative definitions (at a purely technical level) means that all possible synonyms are found for every word in every predicative definition, and every possible combination of these words is structured according to the substantive-verb-adjective pattern. The term "by association" indicates that the subjects and/or objects of an event resemble (or the event itself resembles) –in some cases, very distantly – the subjects and/or objects of events described in other slangs.

For example, the phrase "How do you do?" has become a standard greeting, and no longer a question about how someone does something: "How do you do?" has been borrowed from one slang into another.

On the Semantics of Living Language.

Neither Living Language nor slangs can carry any objective information: I'm proceeding from the assumption that subjective opinions are Truths, and that subjective Truths are always opinions⁵. Consequently, there's no point in looking for objective Truth in slangs that were created subjectively and are used subjectively, and one has to discover a person's thought/opinion and accept that thought as a subjective Truth. In other words, communication between people is always a search for objective information – as Reality – through the recognition of many subjective Truths.

Looking for a meaning in the semantic sense, as the Merriam-Webster Online Dictionary understands it⁶, may not lead one to the desired result.

Lexical Cloning.

4 www.lexiclone.com

5 "For we know in part and we prophesy in part. But when that which is perfect has come, then that which is in part will be done away." Paul, Bible. I to Corinthians, 9&10

6 **1** : the study of meanings: **a** : the historical and psychological study and the classification of changes in the signification of words or forms viewed as factors in linguistic development **b** : a branch of semiotic dealing with the relations between signs and what they refer to and including theories of denotation, extension, naming, and truth

Lexical cloning is the computerized indexing of the predicative definitions in the texts of a particular author, calculating the frequency with which given predicative definitions occur in the text. It means that when one looks for the answer to a given question in a text being used for lexical cloning, one seeks and finds the thought/opinion of that text's author, as a Truthful answer to the question.

In the beginning lexical cloning was thought to be a simple demonstration of the possibilities of a technology searching for information through predicative definitions with the use of synonymic expansion. But it later came to be understood that looking for information in an enormous range of texts is a meaningless task. Such a search may not lead one to the looked-for information. It was understood, then, that – according to the context of the given texts – one had to go to specialists, and seek the needed information from them. "According to context" means that a professional uses a slang in which certain predicative definitions are repeated with a necessary frequency. Indeed, don't we go to an expert to get worthwhile advice?

For example, does it make sense to look in the dialogues of Plato for Truth about "stem cells" and their significance in the treatment of immune disorders? Wouldn't it be better to turn to the microbiologists of the 21st century for that Truth? With Plato, meanwhile, we can discuss eternal problems, such as: who is good?

The Summary of a Conversation With a Lexical Clone.

A program has been created which allows one to "converse" with lexical clones. Moreover, the lexical clone is taken to be a source of opinion on the problem being discussed with him: the program finds, within the texts used as the basis for the lexical cloning, those opinions of the lexical clones' prototypes that are most closely related to the problem under discussion.

In so far as the task was to demonstrate the program's ability to find opinions relevant to the widest possible audience of readers, famous writers, politicians and thinkers were chosen, whose utterances are intelligible to all, as the slangs they used are, by definition, a part of Living Language (if their slangs weren't intelligible to all, they wouldn't be famous).

Sometimes the themes proposed for discussion with the lexical clones are completely different from the themes dealt with in the texts that were used for the lexical cloning. This is done in order to demonstrate how the lexical clones are able to find those opinions that are most appropriate to the new themes. For example, Dostoevsky was challenged to a duel, and George Bernard Shaw was told that he had been cloned after his death.

In order to bring about a common slang between the user and the clone, a summary⁷ is created of the predicative definitions in the conversation between them, constituting a summary of the history of that conversation. On the basis of that summary computer can conclude which texts or excerpts of texts present the greatest interest for the user: what is searched for is what the user needs. Moreover, the clone's

2 : a doctrine and educational discipline intended to improve habits of response of human beings to their environment and one another especially by training in the more critical use of words and other symbols

3 a : the meaning or relationship of meanings of a sign or set of signs; *especially* : connotative meaning **b** : the language used (as in advertising or political propaganda) to achieve a desired effect on an audience especially through the use of words with novel or dual meanings

⁷ <http://lexiclone.com/SummarySample.htm>

interests are ignored -- so far, a lexical clone cannot show initiative, but only reply to the questions put by his interlocutor.

That is, I affirm that lexical clones "understand" (by means of the technology of synonymic expansion) their interlocutors, participating in the construction of a new slang: the slangs used by the parties in the conversation evolve, becoming maximally intelligible to both parties.

To wit – I began: "You are a fool! You are an idiot! I challenge you!"

Dostoevsky replied:

[50.0% Fyodor Dostoevsk_353]

I thought you were.

[40.8% Fyodor Dostoevsk_175]

Where are you going

[40.8% Fyodor Dostoevsk_353]

Yes, you are, indeed

[33.3% Fyodor Dostoevsk_175]

Do you know that you are killing me

I continued: "How dare you! I duplicated you and gave you a new life after your death and you are thinking all these dirty things about me! Duel, only duel!"

Dostoevsky reacted: [8.1% Fyodor Dostoevsk_695]

You are lying, your aim is to convince me you exist apart and are not my nightmare, and now you are asserting that you are a dream.

Then I tried to explain to George Bernard Shaw's lexical clone that I cloned him⁸. I said: "You exist as your duplicate, a duplicate of your soul." George Bernard Shaw used retorts taken from the history of the conversation, showing that Shaw's clone had learned by borrowing my predicative definitions from the slang I had proposed for the dialogue:

[38.4% History]

I cloned, duplicated your soul after your death.

[20.4% History]

I am speaking with your duplicate, not with you.

This means that my predicative definitions have a higher Compatibility than any sentences in the texts by George Bernard Shaw used for the lexical cloning, and that the clone used them while constructing a new slang. Indeed, a hundred years ago the concept of lexical cloning was, to put it mildly, unknown. So George Bernard Shaw has to construct a new slang impromptu, creating it for me.

Compatibility.

The closeness of the relationship between the predicative definitions from the search-obtained text excerpts and those from the question is defined according to a standard formula, called Compatibility:

⁸ http://lexiclone.com/conversation_with_bernard_shaw_extra.html

$$\text{Compatibility} = \left(\frac{\text{Sum} \left(\text{Weight-of-the-same-predicative-definition-User} * \text{Weight-of-the-same-predicative-definition-Text} \right)}{\text{sqrt} \left(\text{Sum} \left(\text{Weight-of-each-predicative-definition-User}^2 \right) * \text{Sum} \left(\text{Weight-of-each-predicative-definition-Text}^2 \right) \right)} \right) * 100$$

In the terminology of Reinforcement Learning⁹ Compatibility is called the "controller".

That is, the number with the percentage that appears at the beginning of the answers quoted above represents the Compatibility of the search-obtained sentence in relation to the question.

Tautology.

From my questions and the clones' answers it is evident that Compatibility arrives at maximum significance -- at 1 (100%) -- only in the case where the search results in a tautology -- only when what was asked is repeated literally. For example, one arrives at a complete tautology when the lexical clone is asked: "How do you do?" Shaw's clone's answer is: [100% Bernard Shaw. Pygmalio_15] "How do you do?"

Ludwing Wittgenstein was right: "...the tautology and the contradiction that they say nothing. The tautology has no truth-conditions, for it is unconditionally true. Tautology and contradiction are, however, not nonsensical; they are part of the symbolism, in the same way that "0" is part of the symbolism of arithmetic¹⁰."

The use of Compatibility as a controller leads to the creation of a tautology: the asker of the question formulates in his answer a subjective Truth, which he would like to hear either confirmed or refuted. And such a confirmation is the literal reiteration of the asker's Truth, with or without an added refutation.

Other Examples of Lexical Cloning.

When I asked: "How should we prosecute terrorists, murderers who kill innocent people? What should we do with them? Should we kill them in return?", Shaw's lexical clone replied, with 19.8% Compatibility: "Who are we that we should judge them?"

Plato's lexical clone answered, with 25.5% Compatibility: "How should we answer him, Socrates?"

Fedor Dostoevsky said, with 14.7% Compatibility: "And so we did."

In continuing the conversation with George Bernard Shaw's clone and asking him "Are you trying to say that my judgment is wrong and it is my fault that I cannot understand you?"¹¹, the following reply was obtained: [11.3% Bernard Shaw. You never can tell_31] "My specialty is being right when other people are wrong" - in this specific case, after a synonymic expansion of the question, a sentence was found in the book *You Never Can Tell*.

⁹ <http://www-anw.cs.umass.edu/r/r/>

¹⁰ [4.4] Ludwig Wittgenstein. Tractatus Logico-Philosophicus. <http://www.kfs.org/~jonathan/witt/ten.html>

¹¹ http://lexiclone.com/conversation_with_bernard_shaw.html

Or, in answering my question: "Is Saddam Hussein our enemy? Should we destroy him?", President George W. Bush said; "This is our calling", with 23.6% Compatibility.

To another question: "How should we prosecute terrorists, murderers who kill innocent people? What should we do with them? Should we kill them in return?", Bush replied, with 4.7% Compatibility: "And the rules are, we got to do everything we can to protect you."

And here is an example of Truth-finding in a conversation with the lexical clone of that well-known demagogue, Plato. I said: "Do you think that I need to study more from you before I can become a Sophist?"
Original Result

[9.3% plato phaed_3] (13)

Then I must try to make a more successful defense before you than I did

Another example:

Question: I am pretty sure that there are no gods and I must take care of myself by myself.

Original Result

[21.1% plato protagora_2]

And what am I doing

[21.1% plato philebu_16]

PROTARCHUS What am I to infer¹²

START and InfoLab Group at the MIT Computer Science and Artificial Intelligence Laboratory.

The START Project presents a technology quite similar to the technology put forward in this article. The results obtained by START are of serious practical interest. But I would like to point out a series of substantial deficiencies in the project:yh

1. "Kernel sentences" are predicative definitions, the combination of a Name with a verb-adjective pair.
2. Texts have contexts and subtexts: the search for information is more effective if one knows which texts use the slangs that relate to the context and the subtext of the question.
3. START isn't a dialoguing system; although it could probably become one.

Summary and Discussion.

The term "Natural Language" should be withdrawn from circulation for uselessness. Indeed, if we can work with Living Language and slangs, what need do we have of such a corpse unfit for life?

References.

1. **Abdur Chowdhury, M. Catherine McCabe.** Improving Information Retrieval Systems using Part of Speech Tagging.
2. **Aleksandrov P.** An introduction in the theory of sets and general topology.
3. **Aristotle**, A new Aristotle reader, Princeton University Press, Princeton, 1995
4. **A.J. Ayer.** Language, Truth and Logic.

12 Plato. Phaedo, Protagoras, Philebus

5. **Bertrand Russell**, On some difficulties in the theory of transfinite numbers and order types; in *Essays in Analysis*, edited by D. Lackey.
6. **Bertrand Russell**, Scientific Method in Philosophy.
7. **Bertrand Russell**, Introduction to Mathematical Philosophy.
8. **Bertrand Russell**. Principles of mathematics.
9. **Bertrand Russell**. My Philosophical Development.
10. **Boris Katz**. Annotating the World Wide Web using Natural Language.
11. **Brill E., Susan Dumais and Michele Banko**. An Analysis of the AskMSR Question-Answering System.
12. **Cantor G.** Works on the theory of sets, Moscow, 1985.
13. **Dmitry Zelenko, Chinatsu Aone, Anthony Richardella**. Kernel Methods for Relation Extraction..
14. **Erik F. Tjong, Kim Sang**. Memory-Based Shallow Parsing.
15. **Fernando Pereira**. Formal grammar and information theory: together again?
16. **George Moore**. Selecting Writings, International Library of Philosophy, Edited by Thomas Baldwin.
17. **Gilbert Ryle**. Dilemmas.
18. **H.P. Gricce and Peter F. Strawson**. The Philosophical Review 65 (1965).
19. **Hong Yu, Vasileios Hatzivassiloglou**. Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences.
20. **Hegel G.** Science of Logic
21. **Hegel G.** Phenomenology of Mind
22. **Hegel G.** Lectures on the History of Philosophy
23. **Hegel G.** On Art, Religion, Philosophy
24. **Hegel G.** The Encyclopedia Logic
25. **Huma Lodhi, Craig Saunders, John Shawe-Taylor, Nello Cristianini and Chris Watkins**. Text Classification using String Kernels.
26. **Hume D.** Dialogues Concerning Natural Religion
27. **Hume D.** Enquiries concerning the human understanding and concerning the principles of morals
28. **Hume D.** Treatise of Human Nature
29. **Ilya S Geller**. The Role and Meaning of Predicative and Non-predicative Definitions in the Search for Information.
30. **James W.** Pragmatism.
31. **Jeremiah**. Bible.
32. **Jimmy Lin, Dennis Quan, Vineet Sinha, Karun Bakshi, David Huynh, Boris Katz, David R. Karger**. The Role of Context in Question Answering Systems.
33. **Jimmy Lin, Dennis Quan, Vineet Sinha, Karun Bakshi, David Huynh, Boris Katz, and David R. Karger**. What Makes a Good Answer?
34. **John Chen, Srinivas Bangalore, Owen Rambow, and Marilyn Walker**. Towards Automatic Generation of Natural Language Generation Systems.

35. **Joyce Chai, Jimmy Lin, Wlodek Zadrozny, Timing Ye, Margo Stys-Budzikowska, Veronkia Horvath, Nanda Kambhatla, and Catherine Wolf.** The Role of a Natural Language Conversational Interface in Online Sales: A Case Study. *International Journal of Speech Technology*
36. **John R. Searl.** Internationality.
37. **Ludwig Wittgenstein.** Tractatus Logico-Philosophicus. <http://www.kfs.org/~jonathan/witt/ten.html>
38. **Noam Chomsky.** 1957. *Syntactic Structures*.
39. **Norman Malcolm.** Knowledge and Certainty: Essay and Lectures by Norman Malcolm.
40. **Paul.** Bible.
41. **Roderick Chilsholm.** Theory of Knowledge.
42. **Steven Bethard, Hong Yu, Ashley Thornton, Vasileios Hatzivassiloglou and Dan Jurafsky.** Automatic Extraction of Opinion Propositions and their Holders.
43. **Turney P.,** Mining the Web for Lexical Knowledge to Improve Keyphrase Extraction: Learning from Labeled and Unlabeled Data.
44. **Matthew W. Bilotti, Boris Katz and Jimmy Lin.** What Works Better for Question Answering: Stemming or Morphological Query Expansion?
45. **Vincent Claveau, Pascale Sébillot, Cécile Fabre, Pierrette Bouillon.** Learning Semantic Lexicons from a Part-of-Speech and Semantically Tagged Corpus Using Inductive Logic Programming.
46. **W.V. Quine.** Two Dogmas of Empiricism.
47. **Walker, Marilyn; Rambow, Owen; and Rogati, Monica.** Training a Sentence Planner for Spoken Dialogue Using Boosting Computer Speech and Language
48. **Wesley Hildebrandt, Boris Katz, and Jimmy Lin.** Answering Definition Questions Using Multiple Knowledge Sources.