

Notes Toward a Pragmatics-Based Linguistics

by

Peter Schorer
peteschorer@gmail.com

Dec. 28, 2019

The Relationship Between Syntax, Semantics, and Pragmatics

Language user: “In the present circumstances x (pragmatics) I have a task y (semantics) I want to perform using language w , and to perform it I choose the syntactic instance z (syntax) in w .”

“Semantics Is Pragmatics”

“For a *large* class of cases — though not for all — in which we employ the word ‘meaning’ it can be defined thus: the meaning of a word is its use in the language.”¹

“Language Has a Purpose”

Naive people, when asked what the purpose of language is, almost invariably reply, “Communication!” But this reply is so vague and general as to be all but useless. A much better reply is, “To accomplish certain types of task deemed important in a given language context.” Examples of these tasks are:

Table 1: Typical Language Tasks

| Language context | Agent | Type of task, i.e., semantic category | Syntactic instance to accomplish task |
|-----------------------------------|-----------|---------------------------------------|---|
| Everyday life in Berkeley, Calif. | Any adult | Request for information | “What time is it?”; “When is your birthday?” |
| | | Expression of feeling | “I am exhausted!” |
| | | Command | “Don’t forget to set the alarm!” |
| | | Declaration | “I despise the City government.” |
| | | Providing of information | “The police station is five blocks north, then two blocks west.” |
| | | Expression of agreement/disagreement | “I’m afraid you’re wrong, because you are not taking into account...” |

1. Wittgenstein, Ludwig, *Philosophical Investigations*, para. 43

Table 1: Typical Language Tasks

| Language context | Agent | Type of task, i.e., semantic category | Syntactic instance to accomplish task |
|------------------|--------|--|---|
| Restaurant | Patron | Request | “Could we have a table for four?” “Does the burger come with fries?” |
| | | Complaint | “This steak is under-done!” |
| | Waiter | Request for information regarding customer’s desires | “Would you like more coffee?” |
| Google | User | Request for information | “ ‘bloviate’, definition” |
| | | | “George Orwell on Ghandi” |

In contemplating the above table, the reader might be struck by the thought, “Since we read the table from left to right, it is literally true that, at least in the table, semantics precedes syntax.” And yet this is the reverse of how we normally think of the relationship between syntax and semantics: normally, we are given a word or phrase or sentence, and asked what its meaning is. The above table suggests that: *semantics is just a finite set of categories into which syntactic entities — words, phrases, or sentences — are distributed.*

Topology, or, “What Lies Near to What”

Topology can be viewed as the branch of mathematics that deals with relative “nearness”. A topology defines a set of *neighborhoods* in a space (almost always a space containing an infinite number of elements). Roughly speaking, elements in a neighborhood are closer to each other than they are to elements lying outside the neighborhood.

What happens if we apply topology to language? The application of topology to *computer languages* was done in the late sixties and early seventies by Dana Scott and Christopher Strachey, in a revolutionary, and completely general, theory called the *mathematical semantics of programming languages*.

Does topology have a useful application in natural languages?

Suppose we let the “neighborhoods” of a natural language (I put the word in double quotes because we have not made a formal definition of a topology, hence of neighborhoods, at this point) be the types of task — the *semantic categories* in the above table. A semantic category contains all the syntactic instances (strings, that is, words or phrases or sentences) that express the meaning defined by the category.

We observe that the mapping (function, table) from semantic categories to syntactic instances is what mathematicians call a *multi-valued* function: for each category there are typically many

syntactic instances that express the meaning defined by the category. (Natural language processing on the computer would be trivial if there were only one!) But we can reduce the number of values by breaking down the category into sub-categories.

What are some plausible types of “topologies” on words and phrases? Certainly the familiar grammar is one. Here all the forms of a noun or adjective or a verb lie close together. (See also “Appendix A — “Grammarless” Languages” on page 19.) Another is the alphabetical “topology”. Here, the closeness of words and phrases is determined by their alphabetical order: all the words beginning with “a” are closer to each other than they are to the words and phrases beginning with “b” or “c” or ... Still another “topology” is determined by usages of a given word in a text or group of texts. In other words, for each word, those phrases lie nearest to it which contain the word or a grammatical variant of it. By studying all these phrases, we get an idea of the meaning of the word, in accordance with Wittgenstein’s dictum, “the meaning is the use.” Another “topology” is determined by order of difficulty. We can categorize all the words and phrases in one of Cicero’s essays (in the original Latin) by how difficult they are for a beginning student.

Parsing and Semantics

Let us consider what by now must be a naive view of natural language processing. The processor (computer program) contains a complete grammar of a natural language. Given a string in the language, the program parses it and then, from the resulting grammatical structure, and the words in the string, somehow infers the meaning.

In contrast, consider a natural language processor utilizing the ideas set forth in the previous sections. Given a string in the language, the parser searches for key words — in the case of Google, these would certainly include “definition” and a long list of proper names (for example of authors). It then uses an internal table to find the semantic categories whose syntactic instances have the best match to the keywords in the string. If the string then matches a syntactic instance in the category to some degree of closeness, the meaning is taken as that specified by the category.

Observe that, in the first case, the processor “knows nothing” of actual usage. Each parsing task begins from scratch. In the second case, the processor is strongly based on usage. Each semantic category contains the most frequently occurring strings expressing that category, these strings having been placed there, at least initially, by the humans who wrote the processor. These humans believe that, fundamentally, *language is not so much rule-based as it is probability-based*. Which brings us to Chomsky.

Chomsky

In the late 1950s Noam Chomsky gained the attention of the linguistics community by pointing out what afterward seemed obvious, but which nevertheless refuted the assumption of many linguists that children’s minds were a blank slate on which the language-learning process during early childhood imprinted the native language. Namely, Chomsky pointed out that by the age of five, children are able to generate grammatically correct sentences that they have never heard. This, he claimed, proved that every human being is born, not with a blank slate, but with a universal grammar that is then filled in with the particulars of each child’s native language.

I have so far been unable to find if it was Chomsky or one of his followers who made the inference that, at age five, children are in principle able to generate an infinity of grammatically correct strings. In any case, we need to take a close look at these ideas.

There is no question but that Chomsky made a seminal contribution to the theory of formal languages. His ideas proved invaluable in the design of compilers for high-level computer languages. In particular, there is no question but that even the simplest of formal grammars, for example, one containing the rule “A can be replaced by bA”, can generate an infinity of grammatical strings (A, bA, bbA, bbbA, ...).

But until we have extensive *frequency-of-occurrence* data on the language of five-year-olds, and in fact on the language of all ages from two up to, say, eight, neither Chomsky nor any of his opponents will be able to talk meaningfully about exactly what goes on in the minds of children during their language-learning years.

I believe that what such data will reveal is that native speakers learn a relatively small set of strings which they then assemble in a limited number of ways, over and over, as daily experience demands (pragmatics again). Exaggerating and simplifying to make the point: *in virtually all language contexts, the same things are said over and over.*¹

An example: several years ago I began accumulating a list of things I needed to be able to say in French in order to communicate with college students from France who were guests in my house, some of whom had only a very limited understanding of English. I found that *fifteen word-processor pages* were enough for me to be able to “get along” in the language context. These pages contained the English sentence or phrase, followed by the French equivalent. The English was ordered alphabetically, for rapid lookup. Here are a few entries:

Day after tomorrow Après-demain.

Day before yesterday Avant-hier

Did [name] show you where you can buy some books about things to do in Berkeley and in San Francisco? Est-ce que [name] vous a montré où on peut acheter des livres à San Francisco ou à Berkeley, et ce qu'on peut faire dans ces endroits?

Did [name] show you where you can buy some food? Est-ce que [name] vous a montré où on peut acheter de la nourriture?

Did you buy it? L'avez-vous acheté?

Did you find it? L'avez-vous trouvé?

Did you have a good flight? Vous avez fait un bon voyage (en avion) ? /Avez-vous fait un bon voyage (en avion) ? Comment s'est passé votre voyage (en avion) ? [If formal, or you say it to several people/friends.] Tu as fait un bon voyage (en avion) ? /As-tu fait un bon voyage (en avion) ? Comment s'est passé ton voyage (en avion) ? [Informal, and to only one friend.]

Did you have a good trip? Vous avez fait un bon voyage ? /Avez-vous fait un bon voyage ? Comment s'est passé votre voyage ? [If formal, or you say it to several people/friends.] Tu as fait un bon voyage ? /As-tu fait un bon voyage ? Comment s'est passé ton voyage ? [Informal, and to only one friend.]

Did you have a nice day [today, yesterday]? Est-ce que vous êtes bien amusée [aujourd'hui, hier]? (or) Est-ce que tu as passé une bonne journée? (hier, aujourd'hui)

Did you have a nice walk? Vous avez fait une bonne promenade ? / Avez-vous fait une bonne promenade ? [If formal, or you say it to several people/friends.] Tu as fait une bonne prome-

1. A year or so after I wrote this, I learned about Zipf's law, which states that in any language, “a few hundred [most-frequently-occurring] words do all the work, while the rest languish in desuetude.” — Holt, Jim, “He Conceived of the Mathematics of Roughness”, review of Mandelbrot, Benoit, *The Fractalist: Memoir of a Scientific Maverick*, in *The New York Review of Books*, May 23, 2013, p. 46.

nade ? / As-tu fait une bonne promenade ? [Informal, and to only one friend.]
Did you read the newspaper today? Avez-vous lu le journal aujourd'hui?
Did you sleep well? Est-ce que vous avez bien dormi?
Do you have any of your photographs with you? Est-ce que vous avez porte des photographies avec vous?
Do not open the front door to anyone you don't know. N'ouvrez pas la porte a un inconnue.
Do not open this door any more than to about here. The door jams very easily. Please move it very slowly and carefully. At night you can close it all the way. S'il vous plait, n'ouvrez pas la porte plus que cela. La porte n'ouvre pas facilement. S'il vous plait, bougez-la tres lentement. A la nuit vous pouvez la fermer tout a fait. [Or, s'il te plait, je prefere si tu ne bouge pas ce porte car c'est tres fragile; je te conseille de la laisser moitie ouvert pendant la journee, comme ca tu ne dois pas la bouger beaucoup, et puis pendant la soiree, avant de te coucher, tu peut la fermer. mais ne l'ouvre pas jusqu'a le mur.]
Don't forget. N'oubliez pas.
Don't forget to lock both locks of the front door when you leave. S'il vous plait, fermez les deux serrures de la porte d'entree quand tu quitte la maison.
Don't forget your umbrella. N'oubliez pas votre parapluie. [If person is familiar - N'oublie pas ton parapluie.]
Don't hesitate to ... N'hésite pas a ...
Don't leave yet. Ne partez pas encore [or, instead of encore, tout de suite]
Don't lose it [e.g., a loaned umbrella]. Ne le perds pas, c'est un parapluie d'emprunt.
Don't tell me! [expression of surprise] Ce n'est pas possible! [Also] Je n'y crois pas!

“Syntax Is Pragmatics”

*Observe that the above list requires no knowledge of grammatical rules! (Although, after repeated use of the fifteen pages, one cannot help but start to infer grammatical rules. Which suggests an idea: *grammar comes after, not before, we learn a language* (or strings in a language). Or, in other words, “syntax is pragmatics”. Modifying Wittgenstein’s assertion (see ““Semantics Is Pragmatics”” on page 2), we can say, “For a *large* class of cases — in fact for all — in which we employ the word ‘syntax’ (grammar) it can be defined thus: the syntax governing a word is established by its use in the language”. Grammar is something we “gather” from frequent use of a language. A grammar is a report, not a rule book. *When we learn our native language, we don't learn grammatical rules first, or while we are learning words and phrases. If you are a native speaker of English, you were quite capable of speaking English before you began studying grammar in school.**

I am confident that 50 pages like the above would suffice to enable one to get along quite comfortably as a tourist in France. (Test: take any elementary book professing to teach French for tourists, for example, Berlitz’s, and count the number of lines given to specific examples of spoken or written French.)

There are now electronic devices available on the market that allow the user to type in or otherwise select a phrase in the native language, and the equivalent words in the foreign language are then heard through a loudspeaker. Such devices have been used by the military in the Iraq war..

Why Semantic Categories At All?

The previous section now raises the question, “But if we can, in effect, make a natural language processor that is essentially a big lookup table, why bother with semantic categories at all? Just have the semantics for each entry in the table follow the entry. As computer programmers might say, ‘Hard wire everything!’”

This is an idea definitely worth pursuing, especially if the processor is designed to be updateable: all input strings would be recorded and sent back electronically to the main programming facility. Those that were not already in the processor’s table would be added, with appropriate semantics.

I would remind the reader who objects that the lookup table would have to be too big, that computers with multi-gigabyte memories (that is, with memories capable of holding tens of *billions* of letters) are now available at the local computer store.

Information Theory

Whenever we talk about frequency-of-occurrence of strings, we are ultimately talking about information theory. Information theory offers a way to encode strings in the most efficient way, namely, by assigning the shortest codes to the most frequently occurring strings.

It is legitimate to ask what the information content of any language context is. The answer begins with a frequency-of-occurrence table for all the strings occurring in the context. By our remarks above, the table would *not* be infinitely long! In any case, we can always limit the table to strings that occur at least, say, 5% of the time.

Note 1: “information content” here does not refer to meaning, as when people say, for example, “His talk had no information content.” Information content here refers solely to frequency-of-occurrence of strings. (If every Christmas you get the same card from your grandmother, and it always contains the same message, then every Christmas you get zero information from your grandmother (at least via her card).)

Note 2: The term “strings” must allow for so-called “wild-cards”, that is, unspecified elements between the given elements of the string. For example, there is in principle an arbitrary number of values of the wild-card *x* in “Please tell *x* that dinner is ready,” but we would normally count the string as one, not many. This is even more important in languages like German in which the subject typically comes first and the verb last in a sentence.

More on “Syntax is Pragmatics”

Applied to textbooks and books for learning languages, another way of expressing the title of this section is, “the organization of the book should be governed by its purpose”.

This idea arises from William Curtis’s, *How to Improve Your Math Grades*, which is accessible on the web site www.occampress.com. Curtis suggests that, regarding the organization of a textbook for any technical subject, we ask ourselves, “What does the organization of this book imply about the way it is intended to be used?” If we consider the organization of a book as a kind of higher syntax of the language, then Curtis’s idea is an example of use (pragmatics) determining syntax.

Curtis’s book arose from his asking himself a naive question, namely, “Why are technical textbooks organized the way they are?” His answer is: “So that they can present the logical struc-

ture of their subject.” But he then asked another question, not quite so naive, namely, “Is organization by logical structure the most efficient organization from the point of view of problem solving?” His answer is “No”, and the purpose of his book is to present a radically different structure which, he argues, is far more efficient for problem solving.

We may apply Curtis’s first question to grammar books — in particular, grammar books aimed at teaching a natural language. We may ask why these books invariably include tables of the conjugations of verbs and of the cases of nouns and adjectives. We may recall our jr.-high- and high-school Latin courses in which we had to memorize, and even be able to recite aloud, the cases of the adjective “hic”: “hic, haec, hoc, huius, huius, huius, huic, huic, huic, hunc, hanc, hoc, ...” We may recall the idle question that passed through our minds at the time, namely, “How often did the Romans say these *sequences* of words?” And our resigned answer: “Probably never, unless they were grammarians.” And there we probably left it, even though we were bothered that, although we could recite conjugations of verbs and cases of nouns and adjectives, we never did learn how to say things that Romans said every day: “How are you?”, “I am tired”, “Which way is the forum?”, and so forth.

So, in light of our previous remarks regarding the application of “topology” in natural language, and of Curtis’s questions, we say: the “topology” of a book whose purpose is to teach a language, should reflect that purpose: words that are close together (on the page) should be words that are close together in actual usage (pragmatics again).

The Damnyankee¹ Theory of Language — Not Grammar Books But Frequency-of-Occurrence Books!

As far as learning to understand a language is concerned — either as a listener to what is spoken, or as a reader — I think a compelling case can be made that the traditional grammar book should be replaced by a frequency-of-occurrence book, i.e., a book that lists:

- the most-frequently-occurring single words;
- the most-frequently-occurring pairs of words in a sentence, where the two words are not necessarily consecutive;
- the most-frequently-occurring triples of words in a sentence, where the three words are not necessarily consecutive);
- ...
- etc.

Available space in the book will necessarily determine where these lists stop and how long each list can be — for each n , the list would consist of the most frequently occurring n -tuples of words, the second-most-frequently occurring n -tuples of words, the third-most-frequently occurring n -tuples of words, etc.

Such a book, I believe, would be a far greater aid to the student than the current grammar book.

Noam Chomsky’s observation that in principle, children by the age of five can generate an infinity of words in their native language because they have internalized the grammar, must be

1. The term is used because it is a well-known example of words that “go together” (in certain parts of the U.S.)

countered with the observation that, in fact, over the course of a lifetime, we only utter a *finite* number of sentences, exclamations, etc., and, more important, that the majority of these are uttered many times. It would be an eminently worthwhile research project to create frequency-of-occurrence tables (histograms) for words and sequences of words in various contexts, e.g., everyday living, restaurants, low-level white collar jobs, etc.

I speculate that the number of words and sequences of words that occur significantly often will be surprisingly low. It is true that some individuals (authors, poets) are creative in the use of language, but the vast majority of people are not: they essentially say the same things over and over. These words and sequences of words can be listed alphabetically, each followed by its translation into a target language. (See the chapter, “Instant French”, in the book, *Thoughts and Visions*, on thoughtsandvisions.com.) A few examples, in alphabetical order in each context, are the following:

In the context of the life of a family with children:

“Did you have a good time?”;
“Have you done your homework?”;
“Tell your father that dinner is ready”.

In the context of TV political talk shows:

“I’m afraid I cannot accept that...”;
“In my opinion”;
“I think that recent events strongly suggest that...”

In the context of a business letter or a business email:

“Dear ...”;
“How much do I owe you for...”;
“I am wondering if you have received...”;
“I have received/not received...”;
“In regard to your letter/email of ...”;
“In what form do you prefer payment?”;
“Please give me your estimate for the following work/products ...”;
“Please send me a bill when the work is completed.”;
“Please find enclosed my check for...”;
“Sincerely”;
“You may recall that on ... [date] we discussed...”;
“Yours truly” ;

A Criterion for Correctness of Linguistic Theories

Every theory worthy of the name must carry with it, implicitly or explicitly, a criterion by which its truth can be established. I propose as a criterion for testing linguistic theories, the effectiveness of the theory in improving natural-language processing on the computer.

The use of machines for measuring the success of ideas is well-known in science. In thermodynamics, the ideal heat engine is described by the Carnot cycle, and the performance of any proposed or actual heat engine can always be compared to this ideal. In computer science, the Turing machine makes it possible to formalize and measure the properties of algorithms and mechanical procedures. (A Turing machine is an idealized computer consisting of one or more finite or infinite tapes containing a succession of cells, each of which contains a symbol or a blank. There is a tape head positioned over one cell at any given time. The “finite control” (program) is a sequence of instructions each of which reads the symbol under the current position of the tape head, then replaces the symbol with another symbol and moves the head one cell to the left or the right.)

But the criterion “effectiveness of the theory in improving natural-language processing on the computer.” is not as precise as the Carnot cycle or the Turing machine. Experience will have to provide ways of tightening this criterion.

Linguistics Under the New Paradigm

What would linguistics look like if the above ideas proved to have some validity? It would look much as it does today, with one change (in the treatment of semantics) and one addition (frequency-of-occurrence tables). That is, each language would be characterized by: phonetics, grammar (in the traditional sense), morphemics, semantics (structured by semantic categories as described above), description of geographical distribution of native speakers, with native speakers categorized by primary occupations and by their cultures, frequency-of-occurrence tables of their language, and a list of existing natural-language processors with the language as source or as target language. A similar breakdown would apply to language sub-contexts within the language.

Lecture to My High School Latin Class

The following are the opening remarks I would make to a first-year high school Latin class if I were allowed (1) to teach such a class and (2) to teach the class as I wanted to.

Welcome to first-year Latin. Let me begin by saying that this course will be unlike any other Latin course currently being taught in high schools around the country — or in Europe, as far as I know.

The main difference is a *de*-emphasis on grammar. The typical Latin course — like the ones I took when I was young — aims at making students knowledgeable about the grammar of Latin in the belief that only in this way can students understand Latin texts and write Latin sentences.

But I have always had a hard time with that idea, because nowhere in the process of teaching grammar to students did teachers or textbook writers ever explain the *process* we were supposed to go through in order to translate a Latin sentence, or write one. Given a Latin sentence, were we supposed to first be sure we knew what the meaning of each word was — and if we didn’t know the meaning, look it up? But looking it up usually wasn’t possible, first, because Latin textbooks had very meager indexes, and second, because Latin uses cases, unlike English. Let me explain what I mean.

English Word Order vs. Latin Word Order

In the sentence, “The boy hit the ball”, “boy” is in the subjective, also called the nominative, case — it is the subject; the boy is performing the action of hitting. “Ball” is in the objective case — it is the object of the verb “hit”; the ball is being hit. So, because “boy” comes first and “ball” comes second in this sentence, you know that the boy was doing the hitting and the ball was being hit. On the other hand, in the sentence, “The ball hit the boy”, “ball” is in the subjective (nominative) case and “boy” is in the objective case. The rule is simple: the thing referenced by the word that comes first is doing the hitting, the thing referenced by the thing that comes second is being hit. But note that the words “boy” and “ball” are the same words in either case!

In Latin, however, cases are indicated by endings on words, and sometimes by variations in the words themselves. In Latin, “The boy hit the ball” is “Puer percutit globum”. (“Puer” means “boy”, “percutit” means “hit”, “globum” means “ball”.) But the word order doesn’t matter. You could say “Globum percutit puer” or “Percutit globum puer” and a Roman would understand that you meant that the boy hit the ball each time. “Puer” (boy) is in the nominative case, and “globum” (ball) is in the objective case, regardless of where these words appear in the sentence. Now to say “The ball hit the boy” we have to change endings. That is, we have to say “Puerum percutit globus” “Globus” is in the nominative case and “puerum” is in the objective case. And here too you can vary the order of the words, and still mean the same thing: “Globus percutit puerum”, “Percutit puerum globus”, etc.

There are typically ten different endings for *each noun* in Latin. There are something like 65 different endings for *each verb* in Latin! So it’s understandable that the writers of Latin textbooks have not put every word in their indexes. (Although now, with huge computer memories available at anyone’s desktop, it is entirely possible to put all these variations of each noun and verb in an index.)

So what the teachers of typical Latin courses in essence say to their students is: memorize all the nouns and all the verbs and all the endings — there are similarities in endings for different nouns, and that makes the job a little easier — then translate the sentence from all this knowledge you have memorized. In other words, if you want to read the sentences in, say, Chapter 6, you have to have gone through Chapters 1 through 5 and memorized just about everything.

John Franklin, in his very interesting autobiography, *Genius Without Genius*¹, describes the kind of memorization required in a traditional Latin course:

“What we learned in ... class wasn’t a *language*, if by that you mean something you can use to communicate with, but instead a series of poems constructed of the forms² of Latin words. Here is the one I remember best, along with a few of our pronunciations of some of the lines:

“hic, haec, hoc [hick, hike, hoke],
huius, huius, huius [hooyus, hooyus, hooyus],
huic, huic, huic [hooick, hooick, hooick],
hunc, hanc, hoc [hunk, honk, hoke],
hoc, hac, hoc [hoke, hock, hoke].

1. www.thoughtsandvisions.com, Vol. 1, chapter 4, section “Junior High School”

2. For “hic, haec, hoc”, the adjective meaning “this”, the columns are, from left to right, the masculine, feminine and neuter forms; the rows are the cases: nominative, genitive, dative, accusative, ablative. The first section contains the singular endings, the second contains the plural. For “agricola”, the noun meaning “farmer”, only the cases are given, since the noun has only one gender, which is feminine. The first section contains the singular endings, the second contains the plural.

“hi, hae, haec [hee, high, hike],
horum, harum, horum,
his, his, his [heese, heese, heese],
hos, has, haec [hohss, hass, hike],
his, his his [heese, heese, heese].

“Another poem was:

“agricola [agricolah],
agricolae [agricol-eye],
agricolae [agricol-eye],
agricolam agricolahm],
agricolā [agricolahhh (longer on the *a*, because this was the ablative)],

“agricolae [agricol-eye],
agricolarum,
agricolis [agricoleese],
agricolas [agricolahss]
agricolas,[agricoleese]

“These and numerous other similar poems we had to memorize.”

But the only place that we find “poem”s like the above is in Latin grammar books! No Latin classic and I daresay no everyday conversation in ancient Rome contained these “poem”s. But if students are going to be forced to memorize strings of words, shouldn’t those strings of words be some of those that occur frequently in the passages that the students will read from the classics?

Let me say it briefly: *Latin grammar is just a complicated way of helping students to figure out the meanings of words. It would be far simpler if students were given a dictionary containing all the words in the texts that students would be reading — and by “all the words” I mean a separate entry for each word-with-ending.* Thus the student could look up, alphabetically, “molior” and find that it means “I work at”; “moliebantur” and find that it means “they worked at”; “molitus erit” and find that it means “he shall have worked at”, and similarly for nouns. Every word in every Latin passage could be looked up directly. No figuring out meanings.

What Good Is Learning Latin?

To answer the question, “What good is learning Latin?”, we can first decide what is most important about Latin in the 21st century, and recognize that this is no longer a time when every educated young person is not only expected to be able to read Latin fluently, but also to write it — in fact, to write verses in it! Those times are gone, and therefore I think that we teachers have to decide what will be the most beneficial, the most useful, for students living at the start of the 21st century.

Understanding Latin Phrases That Are Still Used

What I think is most important is to know the meanings of the Latin words that are still part of our lives. For example words like “etc.” and “e.g.” and “i.e.” which, as some of you probably

know, are frequently-used abbreviations for “and so forth”, “for example” and “that is”. Latin is also present in words like “nil” (nothing), “stat” (immediately, as you no doubt know from watching hospital shows on TV), “et al.” (short for “et aliter”, meaning “and other things”), “per diem” (per day), “status quo” (the situation as it is), and in many other words. There’s a lot of Latin written on our money! On the back of the penny is “e pluribus unum” (from, or out of, many, one — in other words, out of many states, one nation); on the back of the dollar bill is “annuit coeptis” (he has smiled on our beginnings, “he” being God), and “novus ordo seclorum” (a new order of secular things). Until a few years ago, Latin was used in services in the Catholic Church, and we will be learning some of the best-known words and phrases from these services. We will also be learning the meanings of the Latin mottos of some of our states and some of the Latin inscriptions on buildings and statues and some Latin legal terms.

And since I want this course to be of lasting benefit to you, I will be handing out sheets listing, in alphabetical order, all the Latin phrases that are still used. That way, weeks, months, years after you finish this course, you will be able to quickly look up the meaning of any of these phrases you come across.

Seeing How Many English Words Are Derived from Latin

What is second most important, I think, is learning just a few of the many words in English that are *derived* from Latin: words like school, student, competition, stadium, victory, radio, television, automobile, bus, negative, positive, solar, lunar, universe — the list goes on and on. In this course, we will take apart these words and look at their Latin roots. Knowing the Latin roots, you will find, is an enormous help in spelling.

Enjoying a Little of Latin Literature

Third most important is to become acquainted with Latin literature. Most students, if they have to learn a dead language like Latin (in other words, a language that is not spoken or written any more), would like to be able to read things written in Latin, preferably interesting things. And there are interesting things — Latin was spoken by the residents of the Roman empire, the most powerful in Europe for several hundred years (before and after the birth of Christ).

But becoming able to read Latin in the way we read English is a daunting, long-term task — one that I don’t think any student could accomplish in a single high-school course. Furthermore, the struggle to understand Latin texts, what with the difficulty imposed by not knowing the meanings of all the words, and the difficulty figuring out what grammatical rules to apply to each word and phrase, would make the texts far less understandable than they would be in English translations.

So, in this course, we will “enjoy a little of Latin literature” by reading English translations of Latin texts. In some cases, the Latin texts will be side-by-side with the English translations, in other cases not. But we will be reading the translations.

Actually Reading Some Latin

Fourth most important, and separate from the previous goal, is to become able to read some actual Latin. This will be passages chosen by traditional textbook authors from the works of Latin authors. The goal will be to learn the meanings of the words and the associated grammar (from a grammar book I will recommend) so that you can actually *read* each passage: that is, go word-for-word along the Latin sentences, understanding the meaning of each sentence, just as you do with

an English text. This will unquestionably involve repeated readings and partial readings of each sentence.

You will be helped in this task by having *side-by-side Latin-English translations*, and also, where available, *interlinear translations*. An interlinear translation is one in which immediately below each word in the Latin text, there is, in smaller type, the equivalent English word. Now you may be wondering, “Why should we need a side-by-side translation if we have an interlinear translation?” And the answer relates to the problem of word order I mentioned above. In Latin, the order of the words doesn’t matter very much — except for literary effects — whereas in English it does. So an interlinear translation of a Latin sentence in the word order that the author wrote it, is hard to figure out even if the English word is under each Latin word. For example, an interlinear translation of “Percutit globus puerum” would have “hit” under “percutit”, “ball” under “globus” and “boy” under “puerum”. “Hit ball boy”. You have to be up on your endings to know that this means “the ball hit the boy”. But if you also have the side-by-side Latin-English translation, then it will say, “The ball hit the boy” and you understand the sentence right away.

What scholars have done to get around the word order problem is to change the Latin word order of the original so that it more closely approximates that of the English. And in fact you can buy interlinear translations of the great Latin authors and “almost” read them because the English words are right below the Latin words *and* the word order is roughly the same as it would be in English. But as far as I’m concerned, that is cheating. You are not reading what the Latin author wrote. However, I think you will find that an interlinear translation *and* a side-by-side Latin-English translation will enable you to “almost” understand the original Latin more easily than you might have thought.

Of course, *far better* would be an interlinear translation with the word order *the same* as in the original Latin, and with the English word (or words) below each Latin word. That alone would make it difficult to understand the Latin, because you would have to figure out what the equivalent word order in English would be. *But*, if you also had a side-by-side translation, that would enable you to read the Latin text almost as if you understood the Latin as printed. Sadly, however, no publisher that I know of has produced such a translation.

By the way, if you are wondering what good all this struggle with word order will do you, let me tell you that it is one of the best ways to gain flexibility in your own writing. The more ways you can think of saying something, the better the writer you will be.

We *will* have a grammar book and we will refer to it often, but my main purpose is not to make grammarians out of you.

In addition, for certain texts we will read, I will provide you with **frequency-of-occurrence word lists**. These will actually be a hierarchy of lists:

first,
the most frequently occurring, second most frequently occurring, third most frequently occurring, ... *single* words;

second,
the most frequently occurring, second most frequently occurring, third most frequently occurring, ... *pairs* of words (not necessarily adjacent);

third,

the most frequently occurring, second most frequently occurring, third most frequently occurring, ... *triples* of words (not necessarily adjacent);

...

You will then be encouraged to learn the meanings of all the words in the first list, then all the words in the second list, then all the words in the third list ... *in that order*. Or, you can just learn the meanings of just the most frequently occurring, second most frequently occurring, and third most frequently occurring words in each list.

In any case, this will make the learning of the meaning of words much more efficient than the process in typical Latin courses.

We will be reading not only excerpts from serious stuff but also a few things that I hope you will find amusing. Initially, we will be all over the map, reading simple sentences from a wide variety of Latin authors. Then, later, we will settle down to one author whom I think you might find interesting, namely, Cicero (106 bc - 43 bc), a Roman lawyer in a day when public speaking (oratory) was a major art, and a lawyer whose writing I think you will find to be quite clear (with the two types of translation at hand). He wrote about many things, and among them I think you might enjoy reading some of his discussions of divination, that is, the prediction of the future, which was very important activity in his time. Generals and politicians always had to read the signs before undertaking a campaign. One source of divination was dreams, and dream interpretation was very popular in Cicero's time, as it is in our day. You will see how remarkably modern he seems in his way of balancing the pros and cons as to whether divination really works or not. And I must remind you that this was written about 45 - 50 years before the birth of Christ — something like 2050 years ago!

As should be clear by now, this is not a course that requires that you do a lot of memorizing. In fact, I have done everything I can to reduce the amount of memorization to a minimum. But I think you will find that your reading of the Latin excerpts goes faster if you have memorized a few of the most commonly occurring words and phrases, and so I will be handing out, for each author, a list of these in the passages that we will be reading, along with their English translations.

We Will Do a Lot of Reading Aloud!

Speaking of oratory (speaking of speaking), I am going to ask you to read *out loud* to yourself the sentences I assign — and then out loud in class. That's right. I am going to ask you to pretend that you are reading to a small group of illiterate Romans, so that you have to speak very clearly and emphasize appropriate syllables. For example, one sentence you will be given is "Etiam fortes viri subitis periculis terrentur" for which the literal translation is "Even brave men by sudden dangers are frightened." (Here, you hardly need the side-by-side translation.) When you read this aloud, you should read it something like, "Etiam *fortes* viri...subitis periculis...terrentur." In other words, with expression, including hand movements, dramatic gestures, and whatever else helps you to put yourself *inside* the meaning of the words and convey the meaning to your audience of eager Romans. I will ask that you not read aloud *the words*, but that you read aloud *the meanings* of the words. I will explain what I mean by that distinction in subsequent classes. As you do your reading aloud, keep in mind that you are speaking a language that was spoken more than 2,000 years ago. You are bringing it to life! You are experiencing, to a certain

extent at least, what an ancient Roman experienced when he spoke his language. You are him! You are there!

If you are still skeptical about reading aloud, let me remind you that the way we learn our native language is by hearing it spoken, and by speaking it. We learn by hearing the *sounds* of words. Reading doesn't enter the picture until we are fluent, at, say, the level of six-year-olds. The more I think about this, the more I am convinced that learning a language via a grammar is fundamentally wrong. We learned our native language — the one we know best — by hearing it.

Perhaps some day it will be possible to access spoken Latin via the computer. Not only would a speaker read passages in the style I have described above, but the Latin words, with their English meanings below them, would appear as subtitles on the screen. Everyday conversations in Latin between several persons, could be presented the same way. The viewer (student) would be encouraged to speak the words in each scene, along with the speaker.

We Will Study All or Many Of the Usages of a Given Word

In order to improve your ability to read Latin with only “a little” help from the translations, we will at times take a frequently-used word and study numerous sentences in which the word, or a grammatical variant of it, is used. I think you will find that this helps considerably in enabling you to rapidly understand the meaning of the word in your reading, *and* its syntax (grammar).

Speaking Everyday Latin

Fourth most important in my opinion is that you learn to say a few ordinary, everyday things in Latin. You will not be asked to write very much in the language, but I don't see why you shouldn't be able to say, by the time you're done with this course, things like “Hello” and “How are you?” and “My name is ...” and “Where are you going?” I think you should be able to write a simple email in Latin. This task will be made much easier for you because I will hand out a list of sentences and phrases in English with the corresponding Latin, just like the list of English/French sentences and phrases I described above under “Chomsky” on page 4 . The list will be in alphabetical order, and so it will be easy for you to look up, for example, how to say “I am tired.” or “The weather is too hot.” or “I am going to the rock concert tonight.” You simply look up the English alphabetically and get the Latin right next to it! And the list will also contain things like “How much do you think I paid for this watch?”, and “Darn, there goes my beeper”, and “I'm outa here.” Incidentally, the Latin for these last three sentences is given in an amusing book by Henry Beard, called *Latin for Everyday Occasions*¹. Copies will be handed out after this lecture.

OK, that's it! I hope this year of Latin study will not be totally boring to you. I will certainly do my best to keep it interesting.

On the Process of Translating Latin Sentences Into English

I said in the opening paragraphs of the previous section that in no Latin course that I had ever taken were we students given a *procedure* for translating Latin sentences into English. I did not

1. Villard Books, N.Y., 1990. Available through www.abebooks.com.

discuss this process in my lecture because such translating was not to be a major part of my course. But for students in traditional courses, I want to say a few words about this procedure.

The nearest thing I have ever come across to such a procedure is in the most maddeningly frustrating Latin text I have ever read. It is Gavin Betts' *Teach Yourself Latin*¹. One thing I can say without a moment's hesitation is that the book was never tested on self-teachers. The student is given the usual, tedious, boring pages on grammar, then is expected to translate advanced sentences using vocabulary that has not appeared up to that point in the book. It is true that in the back of the book there are translations of the exercises, which include some enormously difficult passages from Latin poetry. (Ironically, the best Latin textbook I know of, namely, Frederic M. Wheelock's *Latin: An Introductory Course Based on Ancient Authors*², has no such translations!) But the student is instructed not to use these translations until he has given up in attempting to translate the exercises on his own, or else as a check on the results of his efforts.

But let me reproduce the process that Betts gives, and suggest that the student use it as a starting point from which to develop his or her own procedure by making changes based on his or her experience.

- (a) Examine each word with reference to its meaning and ending (if it has one) and parse it fully.
- (b) Mark all finite verbs³. This will indicate the number of clauses.
- (c) By observing the punctuation, the position of the finite verbs (which normally come last in their clauses) and conjunctions, define where each clause begins and ends.
- (d) Taking each clause separately, see how each word relates to the finite verb in its clause.
- (e) See from the conjunctions how the clauses are related to each other, and work out the overall meaning of the sentences. — Betts, *op. cit.*, p. 13.

What Good Is Studying Latin Grammar?

Think of the complexity of Latin grammar: the five declensions into which nouns and adjectives are grouped, the four conjugations in which verbs are grouped, and then, in the case of nouns, the cases (nominative, vocative, genitive, dative, accusative, ablative), with a singular and plural form for each, and then, for verbs, all the forms, singular and plural, for each of the three moods... The student may well wonder to him- or herself, "Why is all this necessary?" (and then quickly suppress the question).

The answer that I am sure the vast majority of Latin teachers would give is, "So that you can translate from Latin to English and from English to Latin." But let us look at each task more closely. The process of translation consists of an input and an output. The input is, say, a Latin word or phrase or sentence. The output is the equivalent word(s), or phrase, or sentence in English. Or the input is an English word, phrase, or sentence, and the output is the equivalent word(s), phrase, or sentence in Latin.

But most of this labor can be performed by simple look-ups! Why shouldn't the meanings of *all* the Latin words that are encountered in, say, high school Latin classes be directly look-up-able? I come across the word "regebamus" in a sentence. Why shouldn't I be able to look it up (in

1. Teach Yourself Books, McGraw, Chicago, Ill., 2000.

2. Barnes & Noble, Inc., N.Y., 1960.

3. Although the term "finite verb" appears throughout the book, it is not in the index. If the reader does not notice in the table of contents that there is a section called "Glossary of grammatical terms", he must start searching at the first page.

the R's) and find that the meanings are "we were ruling", "we used to rule". To find that out, there was no need for me to know that "regebamus" is the first-person plural form of the Imperfect Indicative Active tense of the Third Conjugation verb "regere".

Similarly, to translate an English word or phrase or sentence into Latin, why shouldn't the student be able to look up the Latin translation in an alphabetical list of English words, phrases, and sentences? (Obviously, the student might have to do some piecing-together of several items in the list.)

You might reply that such lists of words would "take up too much space". But such lists are in large part just a rearrangement (alphabetical) of the content of the Latin textbook!

"Gradual" Translations of Books in Foreign Languages

I cannot conclude without proposing a type of translation that could prove unusually helpful to a person wishing to improve his or her skills at reading a foreign language. The first page or two of the book would be entirely in English. Then, the most frequently occurring word or phrase would be replaced by the equivalent in the foreign language, followed by the translation in square brackets. This would be done for that word or phrase *throughout the book*. Then the next-most-frequently occurring word or phrase would be treated the same way. And the next, and the next, so that, as the book progressed, the reader would be exposed repeatedly to the foreign language versions of the most frequently occurring words and phrases in the original. I cannot predict how beneficial this type of translation would be, but I certainly think it should be tried.

Appendix A — “Grammarless” Languages

An aid to our thinking about grammars is the rather bizarre concept of the “grammarless” language. There are at least two types. In the first, the grammar consists solely of a list of all finite grammatical strings of words in the language. To learn the language, one must memorize the entire list, and so the list is necessarily finite. There are no grammatical rules, although there is one word for each thing to be represented. (The question immediately arises, of course, if such a language is possible.)

In the first type, word order is of paramount importance. In the second type, word order matters not at all. Thus, each finite set of words has one meaning, regardless of the order of the words. It is an amusing exercise to attempt to answer the question, “How would different meanings involving the same set of things be indicated in such a language?” Thus, for example, whereas in English “Man bites dog” and “Dog bites man” have different meanings, the sequences of words would mean the same thing in the second type of language. Perhaps one way of expressing the two different meanings would be by the addition of a number: “Man bites dog (1)” or “Bites dog man (1)” or “Dog bites man (1)” or any combination of the three words and the number “(1)” would mean, say, “Man bites dog”. Whereas any combination of the three words and the number “(2)” would mean “Dog bites man”.

The obvious next question would be, “Why not eliminate the words and just use the numbers?”