On Language

By

Stuart C Dodd

First Edition

Draft 1

August 2007

Compiled and edited
By
Burt Webb

Table of Contents

Editor's Introduction to this volume

In mid-2007, Dick Spady attended a conference in the mid-west. During the conference, he contacted me by phone and said that he had been talking to one of the attendees who was working on artificial languages. Dick recalled that Stuart Dodd had written some papers on that subject and asked me for references. I delved into the our archive of hundreds of Dodd's papers and found 6 that were explicitly about language. Of course they were in different type faces and degrees of legibility so I reformatted them, cleaned them up and compiled this volume. I decided that the best way to organize the papers would be in chronological order so the reader could follow the evolution of Dodd's ideas. There is some overlap between the papers and redundancy in the information presented. But each paper does include unique content. The following papers are included in this volume:

1. **On Measuring Languages - 1949**

    This paper proposes ten criteria by which the suitability of any language for use as an international language might be measured. These criteria fall into two classes. The first three are criteria of familiarity--that is, they measure the extent to which a candidate language is already familiar to the people who should have to learn it. The remaining seven are criteria of excellence, and are intended to rate languages according to such properties as their freedom from local idioms, from exceptions to the rules of grammar, from inflections, and so on.

2. **Model English - 1949**

    If basic semantic principles are applied to the overworked field of an international language some new and exciting results emerge. This paper applies a basic semantic principle and eight corollaries deduced from it to the leading candidate for a world language and comes up with a new and perfectly regular language which is yet familiar to more people than any other language today.

3. **TILP: A Ten Letter Alphabet of Meanings - 1951**

    Semanticists may expect advance in the science of symbolizing to come from many quarters. The particular advance described in this paper, of reducing about half of all language to ten elements of 'meaning', builds upon a number of different contributions. The analysis presented in this paper, and christened 'TILP' for convenience, reduces all the connective and operative half of any language to 10 primitive elements of 'meaning'. A considerable part of the remaining substantive half of any language is also analyzable into these 10 elements but the extent of this is still undetermined.

4. **Model English: An Example of A National Language Regularized For Machine Translation - 1955**

This chapter will describe some of the rules for a model language for purposes of machine transcription English will be the example in applying the rules. The purpose of this set of rules be to enable an electronic machine to translate swiftly from model English as an input language or, more importantly, into model English as an output language.

5. Ten Semantic Tangles and the Threat of War - 1955

Semantic tangles can contribute to the large complex of factors which lead to World War. International misbehavior in speech can add its bit to endangering the world's peace. If this is true, then we should try to straighten out these semantic tangles as fast as we can see and treat each one.

6. An Alphabet of Meaning for the Oncoming Revolution in Man's Thinking - 1956

The title of this paper suggests that such a profound, widespread, and rapid change is taking place in mankind's symbolizing habits as to be called a revolution. But what is the evidence that the peoples of the world are embarked on such a revolution of their semiotic habits - their customary inter-relatings of symbolizers, symbols and the referents symbolized?

On Measuring Languages

Abstract

This paper proposes ten criteria by which the suitability of any language for use as an international language might be measured. These criteria fall into two classes. The first three are criteria of familiarity--that is, they measure the extent to which a candidate language is already familiar to the people who would have to learn it. The remaining seven are criteria of excellence, and are intended to rate languages according to such properties as their freedom from local idioms, from exceptions to the rules of grammar, from inflections, and so on. Such criteria have three purposes. First, they would rank the candidate languages in familiarity and excellence. Second, they would diagnose weaknesses in each candidate: from this diagnosis a living language could be simplified towards the ideal regularity of an artificial language, while preserving more of familiarity to the world's population than an artificial language possesses. Finally they would indicate any progress that the would may be making from decade to decade towards achieving a single language.

Criteria for a "Best" Language

The problem of an international auxiliary language has become in part a problem of selecting it from among the three hundred candidates which have been proposed in the last seventy years. To select the best candidate requires prior agreement on what is "best". What are the criteria which specify the "best"? This paper proposes ten criteria. It further proposes ten indices which measure the degree to which each criterion is satisfied by a given candidate language. A weighted sum of these indices can then rank the candidates into a relative order.

The criteria for the best world language may be put into two classes: the practical and the ideal. These are also called the natural vs the schematic types when applied to artificial languages. They specify, respectively, what is most likely to be adopted by the world and what is intrinsically the most excellent as a language. These will be referred to here as the criteria of familiarity and the criteria of excellence.

For it should be obvious that the most practical proposal is one which involves the least change or the least amount of new learning for the world population. Thus the candidate language which has the largest proportion of elements which are already familiar to the maximum number of people will encounter the least resistance. We can measure the degree of familiarity of a candidate language and thus intake a crucial comparison to prove which candidates are most practical.

The Familiarity Criteria

What language is most familiar to the users of any one language in that it has the largest percentage of its words, grammatical forms or other elements the same between their language and the candidate world language? In order to measure this, several indices, varying in completeness are as fellows: A first index of familiarity, which may be labeled F_1, is calculated by taking as a first step a representative sample of the elements of the candidate language. One such sample might be the 1000 semantic words which occur most frequently, as determined from a semantic word count. (A "semantic" word is defined as a word or phrase with a unit 'meaning i.e. "to look out for" meaning "to protect".)

The next step is to assign to each of these thousand "most frequent" words, which serve as a representative sample of the candidate language, a value of 1 if it is exactly the same in the national language and a value of ½ if it is partly the same (as in having a root or an affix in common). These unit or half unit values are added up and, since they will give a thousand points at maxima, this total will serve as a percentage of common vocabulary between one candidate language and one national language. This percentage is F_1. $F_1 = (\Sigma\ V/N)$ where V = value of 1 or ½ and N = number of words in sample studied.

Next there will be other values of F_1, one for each national language paired with each candidate language. That is, for one candidate language there will be as many F_1s as there are national languages or important groups of national languages deserving consideration in the world. This number of F_1 indices will then be multiplied by the number of candidate languages for which research provides these data and which are considered important enough to be likely candidates for a world language. It is obvious that this is an immense project of research for many scholars for many years.

These F_1 indices of familiarity next must be combined into a net index of familiarity for each candidate language for the whole world. That is, the F_1 index for one candidate language must be weighted or each multiplied by the number of people speaking the national language corresponding to that index. This gives greater importance to the familiarity index of a language spoken by one million people. From this democratic process of weighting each index by the population to which it applies, there will result a single net index of familiarity, which we may call F_2, for each candidate language. These indices will rank the different candidates in order of familiarity to the world. The indices reveal the languages at the top of the list which deserve further study and also reveal the languages at the bottom of the list which may be dropped from further consideration.

A number of problems of method will have to be solved in computing these indices. For example, in determining the number of people speaking each national language, in order to fix upon a weighting of its F_1 index, allowance must be made for bilingual people or that fraction of a population which may speak more than one language. Such persons might be counted as ½ for each of the two languages they may speak thus giving them a total weighting the same as for any person speaking but one language. Again a further refinement of the indices might be to weight each word in calculating the F_1 index in proportion to the frequency of occurrence of that word. A priori, it seems probable, however, that if the thousand most frequently used words are taken as the sample in calculating F_1, the differences in the frequency of individual words would not greatly change the relative size of the F_1 indices.

A third problem is whether to take the total population speaking a given national language as a weighting factor in calculating F_2 or whether to take some part of it which is more relevant for international purposes. Thus the <u>literate</u> part of each national population is probably a more suitable number to take as a weighting coefficient. This index might be called F_3, measuring the degree of familiarity to literate people. The literate population represents those who are communicating in international affairs more adequately than the host of illiterates. To include the illiterates would give the 400 million or more illiterates of China or India an importance greater than all the Western European nations combined. To weight each nation in proportion to its literate population would probably be fairer basis, since part of learning an international language is learning its written forms. An index of familiarity should apply in part to the people who have already learned some written form of language and might have to unlearn and relearn an international language more than to the people who have no written form and to whom learning a new word would be little more difficult from learning their own national written forms.

The Excellence Criteria

In analyzing next, the excellence of any language for international communication the following criteria are proposed as hypotheses. Some combination of criteria such as these would define what is meant by "the most excellent language" for international communication.

The eight proposed criteria of excellence are that: its sentences should be idiomless and ordered in wording; its words should be univocal in meaning, flectionless, phonetic in spelling and unique in pronunciation; its letters should be unique in sound and shape.

<u>A world language should be idiomless</u>. It should not have phrases which are local and

peculiar to one nation and cannot be literally translated into other languages. A world language should have all its phrases so logical as to enable literal translation into any national language. To measure the freedom from idioms of any language, a list of its idioms as found in a frequency count of a representative sample of perhaps a million words of prose should be made. The index E_1 would be the ratio of the million words of prose examined divided by that million plus the number of idioms (including repetitions) found in that representative body of prose. $E_1 = (N/N+M)$ where N = number of words in a sample studied; M = number of idioms. If many idioms are found, this ratio would be a small percentage. It would become a 100%, indicating a language entirely free of idioms, only when no idioms are found.

To detect an idiom, three tests are available. The first test is the definition of an idiom as a phrase different in meaning from its constituent words. Another test it to try translating each phrase into each of some dozen other representative languages and see whether that phrase can be translated literally. Another test is to see if each phrase can be expressed in the symbols of modern Symbolic Logic. This new science, grown up in the last half century, develops an algebra for words and sentences, so that these qualitative symbols can be handled in equations with all the precision of mathematics.

A world language should have the order of the words in its sentences obey the rules without exception. The rule, of course, may be very rigid; or very flexible as in stating that certain words may occur anywhere in the sentence depending on the emphasis desired. Ideally, it is possible to conceive of a language in which all word order is determined by one rule such as that "modifiers follow that which they modify". This rule would mean that a verb followed the subject and that the object of the verb followed the verb whose meaning it completes. This rule would mean that adjectives followed the noun they modify and adverbs follow the verbs they modify and every phrase or clause follows whatever it modifies.

The index, which we may label E_2, which measures the excellence of a sentence in having the order of its words abide by rule could be computed as a ratio of the number of words in a representative sample of prose, (perhaps one million words), to the number of these words when each is multiplied by its "frequency rank". This "frequency rank" needs explaining. It is determined as follows: For each word in a sentence the rule that determines its position in the sentence is decided upon. The frequency of occurrence of each rule must then be determined in a large sample of prose, and the rules put into a rank order so that the most used rule will be given a rank of 1, the next most used rule will be given a rank of 2, etc. Each word is then multiplied by the rank (whether 1, 2, 3, etc.) of the rule which governs its position in the sentence. This multiplication by a rank is "weighting" the word according to the frequency rank (in this case the frequency rank of the rule governing the position of the word in the sentence). $E_2 = (N/\Sigma R_2)$ where N = number of words in sample studied; R_2 = frequency rank of rule governing the position of each word. By this index, if there is but one rule for all words the weighting factor is 1 and the index will be a 100%, as it will be a million words divided by a million words. If, however, a second rule appears then some of the words in the denominator of the index will be multiplied by two and the index will be less than 100%. If a third rule appears the index will become still smaller in proportion as the number of rules becomes greater.

A high index, therefore, measures simplicity of language in this respect and a low index measures its complexity or irregularity. The index also is proportional to the frequency with which each rule occurs in the representative sample that is studied. It should be obvious that

having a definite word order makes sentences which are clear and unambiguous in meaning. If the order of words in a sentence always follows some rule, there is little possibility of different people interpreting the sentence in different ways. Thus a rule-abiding order of words is an objective way of measuring and controlling the degree of ambiguity in the sentences of a language. This is especially so in a language whose words are not inflected (as explained below).

     <u>A world language should have words which are uninflected.</u> This criterion means that no word should ever change its form to express a grammatical inflection such as masculine or feminine gender, persons, or number, tense, voice or mood of a verb or degrees of an adjective. This is the trend of evolution of language. Languages grow up with these grammatical inflections in primitive thinking as when man ascribed masculine and feminine gender to all nouns, simply because man thought of his own difference in sex as existing in everything else around him. But as people developed towards greater maturity and flexibility in language they dropped these grammatical inflections. Some of them are entirely unnecessary. Others are expressed in uninflected "participles" such as the prepositions and conjunctions and adverbs like "to", "as", "and", "or", "not", etc. Chinese has gone furthest in developing a completely uninflected language of root words which can be flexibly combined in different orders to make a great variety of meanings.

     The flexibility of uninflected words can be compared to the flexibility of the alphabet where the clumsy symbols for whole syllables were replaced by a few letter symbols for elemental sounds. These letters can be flexibly combined to make any word in any world's language. Somewhat similarly, root words and particles yield more flexible sentences with a greater range of possible meanings than inflected words can do.

     To measure the degree to which a language has progressed towards the ideal of complete absence of inflections, an index, which may be called E_3, may be calculated from the same representative body of prose of perhaps a million words which may be used for calculating most of these indices discussed in this paper. The formula for the index of freedom from inflections is the ratio of one million words divided by those million words each weighted by its "frequency rank of inflections". This 'frequency rank of inflection" is determined in a way similar to the frequency rank of rules in the preceding index, E_2. To get it, the number of times each inflection occurs in the million words is counted, and the frequency of the inflections with one grammatical meaning is given the ranks of 1, 2, 3, etc. Each word is then multiplied by one if it is uninflected, by 2 if its first inflection is the most frequently occurring one, by 3 if its first inflection is the next most frequently occurring one, by a weight of 4 if its first inflection is the next most frequently occurring one, etc. If the word has more than one inflection, it will be multiplied by more than one such rank. $E_3 = (N/\Sigma R_3)$ where N = number of words in sample studied, and R_3 = frequency rank of each inflection of each word. By this index, a language will be a 100% flectionless only when it uses root words and particles, only. It will be less than a 100% flectionless in proportion as:
    a. It has many words which are inflected
    b. The inflected words are frequent in occurrence, and
    c. There is more than one inflection to express one grammatical meaning.
Thus a language which has four conjugations for its verbs instead of one conjugation will have a larger weighting in the denominator of the index and, therefore, a lower index of excellence in respect to being flectionless.

A world language should be phonetic in spelling. This criterion of excellence that every word should be spelled exactly as it is pronounced implies the criterion mentioned below that every letter should represent only one sound. When the words of a language are spelled as they are pronounced, learning to read that language becomes very simple. If there is much literature and reading matter in one's environment, a child will learn to read without schooling as automatically as he learns to speak by merely being surrounded by people using the written language and by his wanting to know what others are writing and to write things himself. A phonetic spelling is perhaps the greatest aid to make the population a 100% literate. All languages which use letters were phonetically spelled at one time of course, but in the case of many languages the spelling of a previous century has become standardized while the pronunciation has changed. Another source of un-phonetic spelling, however, is that there are more sounds in a language than letters, so that some letters will be used to mean more than one sound. Thus English uses 40 sounds, but has only 26 letters in its alphabet with a result that its irregularity of spelling is greatly increased.

To measure the degree to which a language is phonetic in spelling an index of this criterion of excellence, which we may call E_4, may be defined as a ratio calculated from a large sample of perhaps 100 thousand letters as they occur in the representative sample of prose referred to above. The index might be 100 thousand letters divided by the number of those letters when each one is multiplied by its "frequency rank of pronunciation." This frequency rank of pronunciation is again calculated similarly to the frequency rank of rules in E_2 or frequency rank of inflections in E_3. To calculate it, the frequency with which each pronunciation of each letter recurs must be counted. Then for any one letter, its most frequent pronunciation is given a rank of 1. Its next most frequent pronunciation is given a rank of 2 and so on. Each letter in the denominator of the ratio is multiplied by its rank and these products are added to make the denominator of E. $E_4 = (N/\Sigma R_4)$ where N = number of letters in sample studied; R_4 = frequency rank of pronunciations of each letter. By this index, a language will be 100% phonetic only when each letter has one pronunciation and when every word is spelled in a single and phonetic way. The phonetic index E_4 of a language will decrease in proportion as its letters, as they occur in words with current spelling, have more than one pronunciation.

A world language should have words which are univocal in meaning. This criterion of excellence means that every word should ideally have only one meaning and every meaning should have only one word to symbolize it. There should be no words with multiple meanings nor should there be any synonyms which mean exactly the same. (Synonyms with slightly different meaning are desirable to express shades of differences in meaning and to make a language rich, but words between which no differences in meaning can be detected are merely confusing.) This is a fundamental principle of symbolism--that each symbol should represent one and only one "referent" or meaning. Obviously, our living languages as they have grown up in folk usage have acquired multiple meanings for many of their words. Only artificial languages such as Esperanto approach the ideal of "one word, one meaning" as they can start out afresh by assigning a word or phrase for every meaning listed in the dictionary.

To measure the excellence of language in respect to its words being unique in meaning, an index, E_5, may be defined as a ratio calculated from the same representative sample of a million words of prose which has been used previously. This index is one million words divided by the number of those words when each is multiplied by its "frequency rank of meaning". This

frequency rank of meaning is similar to previous frequency ranks. It would require a semantic word count, i.e. a count of the frequency of occurrence of each meaning (as listed in the dictionary for each word) in the million word sample of prose. (See Eaton's Semantic Frequency List for English, French, German and Spanish.) Each meaning of each word will be given a rank of 1 if it occurred most frequently, of 2 if it occurred next most frequently, and so on. Each word would be multiplied by this frequency rank and all these products would be added up to get the denominator of E_5. $E_5 = (N/\Sigma R5)$ where N = number of words in sample; R_5 = frequency rank of meaning of each word. Since no complete semantic word counts have been made as yet in the world to the author's knowledge, (although a scientific committee is at work on this in the United States) a similar index of uniqueness of meaning of words may be E_{5A}*E_{5A} might be the number of words in the most complete dictionary of a language divided by the number of meanings listed in that dictionary. This variant index is easily computed but has the disadvantages of giving great weight to unusual and archaic meanings with which a language may be burdened and ignores the important factor of the frequency of the use of a word with multiple meanings.

The world language should have uniform pronunciation everywhere. This sixth criterion of excellence means that there should be no difference in different countries in the way any word of the world language is pronounced. By means of the international standardized phonetic alphabet the standard pronunciation of every word can be fixed. Phonograph records and radio recordings can also fix the pronunciation. Some people may comment that since languages have changed their pronunciation in the past would not the new international language also change as a whole or in regional dialects? This is highly improbable as the modern forces such as radio and other agencies of mass communication would increasingly tend to unify and standardize and preserve pronunciation. Dialects grow only where people are separated with little communication between them.

To measure the degree to which any candidate language approaches this ideal of universally uniform pronunciation an index of uniformity of pronunciation, E_6, may be developed. For this index, a survey of a sample of persons representative of the various regions, social classes, etc. within each national language might be asked to read standardized prose into a recording machine. From these recordings, the frequency of each pronunciation of each word could be counted, and each pronunciation of a word given a rank. Then the index E_6 would be that million words when each has been multiplied by its frequency rank of pronunciation. $E_6 = (N/\Sigma R_6)$ where N = number of words; R_6 = frequency of pronunciation of each word. This index, like the previous ones, becomes 100%, showing complete uniformity of pronunciation, when the rank of every word is one so that the index is one million divided by one million. In proportion as there is more than one pronunciation for each word, the denominator increases and the index of uniform pronunciation shrinks. For example, if there were two pronunciations only on the average for every word the ranks of 1 and 2 would occur equally often as weights, in the denominator which would then have the average value of 1.5, giving an index of one million divided by a million and a half which is 67% of uniform pronunciation. Again, if there were, in general, three pronunciations of every word so that the ranks of 1, 2, and 3 occurred about equally often, then the denominator would be twice the size of the numerator and the degree of the uniformity of pronunciation would be only 50%.

A world language should have every letter unique in shape and sound. These two final criteria of excellence of any language apply only to its written and printed forms. The second means that every letter should have only one pronunciation and every elemental sound in the language should have a letter to represent it. The index to measure this is included in the index of phonetic spelling, E_4 above.

For each letter to be unique in shape it means that every letter would have one and only one visual form, regardless of whether it occurs-in print or in hand writing, or whether at the beginning of the word (where capitals are used in some languages), or in the middle or end of a word. Thus English has four forms for many of its letters and Arabic has three forms for many of its letters. To measure the degree of uniqueness of shape of letters more exactly a seventh index of excellence, E_7, may be defined as the ratio from a sample of a hundred thousand letters in the representative samples of written and printed prose. This numerator would be divided by the sum of those letters each multiplied by its frequency rank and shape. This frequency rank of shape, like the preceding frequency ranks, would be determined from a count of the frequency of each shape of each letter. Putting them into rank order and multiplying each letter by its rank and adding these products gets the denominator of the index E_7. $E_7 = (N/R_7)$ where N = number of letters in sample studied and R_7 = frequency rank of shape of each letter. It will be 100% only when every letter (including its connection to an adjacent letter) has only one shape.

Seven indices of excellence for any language have been defined above. The next scientific step is to combine them into a single index of excellence for any one candidate language. There are various possible ways of combining them. The simplest way is to draw a profile graph. This means to draw a column showing the percentage value of an index and placing the seven columns for the seven indices of one candidate language side by side. The broken line across the tops of these seven columns is the "profile" for that language. By drawing and superposing profiles for the different candidate languages it might be obvious that one or two are far superior in most respects to all the others (or possibly far inferior to the others and so may be dropped from further consideration).

If the profile, however, shows several candidates with overlapping profiles a more exact method of combining the seven indices into one must be used. The simplest way is to get the simple unweighted average, E_A, by adding them together and dividing their sum by 7. This gives a simple average index of excellence for one language permitting its excellence to be compared with the excellence of other candidate languages. If more refined weighting is desired, it can be secured by having panels of judges who are experts in the science of language distribute a 100 points to the 7 criteria so as to show the relative importance of each. The average number of points assigned by the judges to each criterion would then be a weighting factor for that criterion. This weighting factor for each criterion would be multiplied by the value of its index before adding the. 7 indices and dividing to get the weighted average index of net excellence for one language, E_w. $E_w = (\Sigma wE/w)$ where E = each of the preceding indices of excellence E_1 to E_7 in turn; w = weight of each index.

Still more refined weighting schemes could be developed such as one based upon the number of man-hours, or the amount of human energy, required to learn and use whatever each index measures. Thus if unphonetic spelling adds 20% of letters to the words of the language in general then writing, typing, and type setting of that language requires 20% more time than *a* language having phonetic spelling. Similarly, the number of hours required to learn an otherwise equivalent but flectionless language would yield a weighting factor for the third criterion

dealing with flections.

The Purpose of the Criteria

As a result of the researches outlined above there would be an index of familiarity and an index of excellence (such as F_3 and E_w) for each of the languages, whether an artificial or a living one, which are candidates to become the auxiliary world language. These indices will serve three purposes. First they would rank the candidate languages and tell which was the most familiar and the most excellent. Thus the problem of choosing the "best" world language would find a scientific answer (based on rules in which the subjective element has been minimized).

Secondly, these indices would diagnose and measure weakness and the degree of strength in each candidate language whether *in its* freedom from idioms, its regularity in word order, its freedom from flections, its phonetic spelling, its uniqueness of meaning of words, its uniformity of pronunciation or its uniqueness of the sound and the shape of its letters. From this diagnosis any living language could be simplified towards the ideal regularity of an artificial language while preserving more of familiarity to some part of the world's population than artificial languages possess.

This has been done for English. The resulting "Model English," constructed by the author, has the regularity of the most ideal artificial language coupled with greater familiarity to more people than any rival national or synthetic language. Its indices of excellence are all 100% and its indices of familiarity are well above all rivals. This gives Model English first rank by all criteria for a world language.

A third purpose of these indices is to measure any progress that the world may be making from decade to decade towards achieving a single world language. The relative degree of gain among the rival languages may be measured from period to period partly by techniques of representative sampling as in Gallup polls. The degree of a person's knowledge of a language, averaged for a population, must be also included in any accurate measurement. Any trend, however, slight, towards a single language eventually sweeping the field and becoming the sole auxiliary language would be shown, and its spread could be facilitated.

Model English - a world language

(stated in model english)

a familiar and perfectly regular international auxiliary language

Foreword

If basic semantic principles are applied to the overworked field of an international language some new and exciting results emerge. This paper applies a basic semantic principle and eight corollaries deduced from it to the leading candidate for a world language and comes up with a new and perfectly regular language which is yet familiar to more people than any other language today.

A basic semantic principle calls for one-to-one correspondence between a symbol and its referent. This is never perfect but is a goal to be sought in spite of shifting conditions where a symbol's meaning may vary among speakers, among hearers, at various times, in various contexts, and with other conditions. This principle of one-to-one correspondence when applied to the different aspects of a language calls for a single word order for each sentence, a 'single meaning', grammatical form, spelling and pronunciation for each word, a single sound and shape for each letter and no idioms. The correspondence principle also calls for using the language which is familiar to most people. For in proportion as a language is unfamiliar to many people it lacks referents to them; an unknown language has no referents to its hearers; a familiar language has more words with referents to more people.

Those corollary principles can be measured and as percentage indices each being a percent of the maximum of that principle. This has been done and published in 'on measuring languages' in the journal of the american statistical association, march, 1949. This gives a set of indices, or even a single scale if desired, to guide the construction of a model language. To illustrate its familiarity the paper from this point on will written in Model English (except for phonetic spelling) and will tell its own story.

what is the best language for the one world of the future? will it be one of the thousand present folk languages or one of the three hundred artificial languages? a new answer to this old question may come from putting several sciences together using principles from linguistics, semantics, psychology, sociology, and statistics. such a set of principles can be briefly stated in ten rules or tests for 'the best' language for world use. this set of ten rules will be applied here to the english language to test how good a world language it can become.

this paper will build on another paper[1], developing measures for each of the ten tests. the measures can be percents which show how well any language, natural or artificial, may satisfy each test. to show this the english language will be tested by its percent of the possible hundred percent on each test. this description here, to be brief, will follow the summary of it in plate 1. the reader should study plate 1 to understand the remarks below. in plate 1, column 1 will state the ten rules or tests; column 2 will state how a percent can measure each test; column 3 will state the application of the rule to english to make it a model world language; and column 4 will state the degree to which this model english can satisfy each test. the ten rules will be applied here to english because it can satisfy the rules more completely than any other language, natural or artificial, as the percents in column 4 of plate 1 prove. more people understand english in the world today than understand each other in speaking any other language. several hundred million people speak it out of the two hundred billion people in the world. it thus will satisfy rule 8 in having the most users. as a second reason for choosing english, it can become a completely regular language with no more changes than the reader will note in this paper. by choosing the rule-obeying forms out of the many different grammatical forms that exist in english, it can become as completely regular as the most

perfect artificial language. for proof, all the sentences in this paper use only this regular english, obeying the rules as in plate 1 (except that the spelling will not be phonetic). this model english can satisfy the rules 100% as the reader will see from column 4 of plate 1.