'Real Character': A New Focus

Ian Maun
University of Exeter, United Kingdom

Abstract

This paper gives a brief résumé of the papers thus far published in this series (Maun 2013, 2015, 2016) and goes on to examine a number of questions which will influence the development of a modern-day digital 'Real Character'. These include translation, *lacunae*, and idioms. Historical attempts at devising classifications of the world, or 'ontologies', are examined. In the light of differences between these attempts, together with the problems posed by translations, *lacunae*, and idioms, the idea of a universal Real Character is rejected in favour of a focused or tuned version, in which the language and culture of the recipient of a message are taken into account. The system of character composition is retained, but icons containing single characters may be replaced by others with supplementary elements which are necessary for

---

Ian Maun
University of Exeter, St. Luke's Campus, EXETER EX1 2LU, United Kingdom
Email: I.C.Maun@ex.ac.uk

comprehension by the recipient. The functioning of a Real Character Device is considered in the light of the proposed changes to the system of character-composition and syntactic formation.

1. A Modern 'Real Character': The Story So Far

This is the fourth in a series of papers which develops the basis for an international means of communication. The first three (Maun 2013, 2015, and 2016) have examined the possibility of re-introducing, in digital form, the 17th-century notion of a 'Real Character', that is to say, a pasigraphical ('read-only') system of writing that represents ideas not sounds. With such a system, it would, in theory, be possible for a speaker (or rather, writer) of one language to communicate with a reader who did not speak or read the writer's own language.

In the first paper (Maun 2013), the author deals with the theory of a 'Real Character' (hereafter, RC) and the possible application of such a system to computer-mediated communication. The need for increasingly efficient international communication is noted and the idea of 'iconicity' is incorporated into the argument, together with the narrower Peircean idea of the 'symbol'. Historical antecedents from the seventeenth century onwards are examined (George Dalgarno, John Wilkins, and Gottfried Wilhelm Leibniz), as are present-day artificial languages. Modern research into semantic primes is linked to work on visual primes, and the two are related through the possible use of straight lines, curves, points, and shapes to represent given meanings.

The notion of 'the sentence-as-character', is proposed, a format which dispenses with the need for linear syntax. This format consists of a T-shaped character, in which the predicator is placed on the top line, and two arguments (subject and complement) are placed in the angles. A sentence may thus be read, not in a pre-determined linear order, but in the order of the recipient's native tongue. The question of metaphorical meaning is discussed in the light of theories of embodiment. In order to facilitate speedy reading, on-screen Rapid Serial Visual Presentation (hereafter, RSVP) of texts with animation, colour, and 3-D is proposed. Technological issues of computer implementation and further development are raised, ways forward are examined, and cultural and linguistic problems tabled for further consideration.

Maun (2015) examines the question of a lexicon for such a system. The basis for a vocabulary is sought in work on simplified language and possible semantic universals, and the process of affixation in natural languages is examined as a way of identifying prominent semantic notions. It is suggested that Ogden's Basic English (1930) might form the lexical core of RC, as it provides words and lexical units which may be combined and which cover many of the concepts found in the world's languages, according to Swadesh (1950, 1971), together with a number of elements from Natural Semantic Metalanguage (Goddard 2010) and from Hogben's artificial language, *Interglossa* (1946). Estimates of necessary vocabulary size are made on the basis of previous studies and the conclusion is reached that a basic vocabulary for RC should consist of c. 2000-3000 words. A critique of a number of existing visual languages is given, including a brief look at the symbolic system developed by Karl Haag (1902), and methods are suggested for converting semantic notions into a consistent system of images and characters. Ways of giving

conceptual support (e.g., componential analysis of words) in the digital transmission of such characters are also examined.

Maun (2016) examines the work of Karl Haag (1860-1946) in greater depth and relates it to modern linguistics and RC. The paper places his work within the historical context of writings on 'universal language', artificial languages and the development of mathematics and logic in the early part of the 20th century. Haag's 1902 book develops a system to describe the logical structure of language and to represent it not by words but by symbols. The basis of the system is that language is predicated on the human body and it is through our perceptions of space (the vertical, the horizontal, the distant, and the enclosed) that we create both literal and figurative language. These perceptions form semantic primes and may be applied equally to a number of fields, e.g., the biological and the mechanical. Haag produces symbols consisting of lines and dots to represent the primes and use capital letters to represent the fields, e.g., M = Mechanics. He furthermore introduces the notion of 'Force Levels', by which a single concept such as 'in' may apply at five levels (state, transition, causation, propulsion, passive). Thus, *be in*, *go in*, *put in*, *force in*, *be inserted.*

Haag's work is then related to that of Lakoff & Johnson (1980, 1999) on the metaphorical nature of language. A relationship is also shown between Haag's system, which uses symbols, and that of Chilton (2014), which involves the use of 'vectors', shown on arrows. Suggestions are made for modifying the syntactic T-bar into an I-bar to incorporate the indirect object, and to represent Haag's Force Levels (usually shown by triangles) with symbols which are more consistent with his line-and-dot symbols.

2. The Present Situation: What We Have and What We Need

Summarising the above three papers, the following list demonstrates the principles to be incorporated into the design of a modern RC.

- The system will be developed for use by a generalist learner.
- Icons and symbols will be used, with occasional indices, and semantic primes will be incorporated, as well as identifiable semantic meta-units (affixes).
- The core of a lexicon may well be Basic English. Visual primes will be employed to convey meaning and it may be necessary to employ stylistic conventions in the design of all types of character.
- Haag's line-and-dot symbols will form the basis for both literal and metaphorical characters where icons cannot be used and his Force Levels may be included with appropriately modified forms.
- Haag's use of Fields to express metaphor will be incorporated, e.g., the symbols for Space when used to represent Time: *near* = *recently*, *soon*.
- The composition of RC characters may follow formational and reading parameters in the manner of Chinese characters and Maya glyphs, with conceptual support available digitally beneath the surface of characters.
- Syntax will be expressed through the use of the T-bar structure. This may be modified into an I-bar to incorporate an indirect object position.

- Presentation of messages will be though RSVP on computers and other digital devices.

It should be stressed that no fully worked-out RC has yet been developed, nor can such a system be designed until a number of questions and difficulties have been addressed. Given that an RC should function as an intermediary between two interlocutors who do not speak each other's language, we must now examine the question of whether such a system can function fully between *all* languages, or whether the conceptual differences in the world-view of different languages means that RC must be 'tuned' to the source and target languages. That is to say, formulated in such a way that a computer can produce in symbols a message from a writer of Language A that is capable of conveying the correct meaning to a reader of Language B.

3. From Language A to Language B

Consideration of the design of an RC has thus far focused on the code, i.e., the set of characters that will be used to formulate messages, and to a lesser degree, on the technology that will convey a message from a speaker-writer of Language A to a speaker-writer of Language B, where neither knows the other's language. It is the sharing of this code that will make the intended meaning of the sender clear to the recipient. This requirement, however, is only a part of the picture. We need to understand (a) the other elements of the communicative situation in which two users of RC will find themselves, (b) the difficulties caused by A and B having different means of encoding their meaning, and (c) the functioning of the

technology which will act as an intermediary in the act of communication.

Let us first take the writer, the reader, and their respective situations.

Speaker-writer A speaks, writes, and understands Language A. This language variety may be unique to a particular geographical area, class, caste or religious order, or even a few dozen people. Alternatively, it may be the common language in an area or even a world-wide language, such as English. It will be surrounded by, and form part of a culture, which may be narrow in scope or very broad. It may well possess words for particular aspects of that culture, e.g., customs and festivals, or for elements of the environment, natural and man-made, all of which are peculiar and unique to that situation, e.g., animals of a particular region or religious buildings. It may contain words which were once common and everyday but which now are used only by certain people, or in certain contexts, situations and places, e.g., religious terminology, courtroom language, the theatre.

This language will also have conventions about forming texts, both written and spoken. It will be used in various physical situations. It will have different registers, both spoken and written, and conventions about taboo words and topics that may or may not be used and discussed. It will have conventions about turn-taking and the formulation of speech acts such as promising, swearing an oath, or apologising. It will have slang, dialects and idiolects, neologisms, and new ways of using morphology and syntax.

In short, Language A will be a pre-existing but still-developing system embedded in a culture and being used by a particular speaker in a given social situation, principally with other speakers of Language A who know the vocabulary and the grammar as well as

the social conventions of usage.

Language B may be totally different in terms of syntax, phonology, morphology and vocabulary and a set of usages, conventions and constraints, both linguistic and social, will surround it. These may be (and almost certainly *will* be) totally different from the linguistic, social, and cultural confines of Language A.

Thus in a message from Writer A to Reader B, almost everything about linguistic formulation and social usage could be different. Nevertheless, Writer A must create a message in a given physical and social situation using the linguistic tools available to him or to her. RC transfers not words or sounds, but ideas. When we formulate a message, however, it is in language, not concepts or ideas. Writer A must start from Language A.

Reader B who receives this message may not understand Language A, nor its social and cultural conventions. In normal circumstances today, such a reader would be dependent upon a translator who knew both Language A and Language B and their respective conventions. The point of an RC is that Reader B does not need a human translator. The technological device that processes RC will be able to interpret the input-symbols and either transmit them directly, or make the necessary social and cultural changes, such that Reader B receives a symbol-message which is comprehensible linguistically, socially, and culturally.

4. Translation and *lacunae*

Translations are probably as old as language itself. When a speaker of language B does not understand a speaker of language A, an intermediary who speaks both languages is sought, and

communication is established. To find a means by which a speaker of Language A can transfer a written version of a message to a speaker of Language B, or any other language, is the problem at hand.

The fact that translations exist, and that ways can be found around difficulties of transfer, does not mean however, that the message from speaker A may be accurately and exactly transferred to speaker B. This is, in part, because there is usually no systematic match between the vocabulary of a source language and that of a target language. Studies of colour (Berlin & Kay 1969, Deutscher 2010) and kinship (e.g., Lounsbury 1956) have been fertile grounds for examples of non-matching. Despite the possibility of reasonably accurate translations between languages, it is rarely the case that there is an exact match between a word or expression in one language and a word or expression in another, as shown in the following examples.

Let us take four widely differing languages. German is an Indo-European language closely related to Dutch and English, and more distantly related to the languages of Scandinavia (Danish, Norwegian, and Swedish). Pitjantjatjara and Yankunytjatjara (hereafter, P/Y) are regional varieties of the Western Desert Language of Australia (Goddard 1993). Seneca is a minority Iroquois language of the United States. Navajo is a Southern Athabaskan language of the Na-Dené family from the south-west of the USA.

In English one can say, *I'm going to the shops*, *I went to London*, *I shall be going to America in a month,* etc. The single verb *to go* and its morphological and suppletional variants cover all these cases. In German, the concept of *to go* is not so simple. The method of travel must be specified. Thus *to go on foot* is *gehen*, whereas *to go in a vehicle* is *fahren*. Then, if one goes by plane, *fliegen* must be used (as in English, *to fly*). To go by ship is *reisen*, which is also the verb used

for general travelling over a long distance. However, where English would naturally use *to go*, German does not necessarily do so. Thus, *I must go to the shops* is *Ich muß einkaufen* (= I must make purchases). *I went to my sister's* is *Ich besuchte meine Schwester* (= I visited my sister), and *Where did you go on holiday?* translates as *Wo warst du in Urlaub?* (= Where were you on holiday?). The linguistic travel-map is different for the two languages.

Then there are lexical differences. Where there is no single word for a word in another language, this is known as a *lacuna* (= 'gap'), plural, *lacunae*. German has many single words which cannot be translated by a single word into English. Many are casual or idiomatic, such as *Fahne* (= 'flag') for *the smell of alcohol on someone's breath*. Some words are nut-like condensations of several concepts into one, such as *verschlimmbessern*, meaning *to damage something when trying to repair it* or *to make a situation worse while trying to improve it*. *Fingerspitzengefühl* (= 'finger-end feeling') is used for sensitivity in dealing with other people. The fact that there is no single word in English for a German lexical item does not mean, however, that the German is unstranslatable. It simply means the German and English conceptual worlds are differently divided. *Verschlimmbessern* is a narrow slice of the world. The English version, *to make a situation worse while trying to improve it*, is a wider slice, requiring for its expression a broader range of semantic concepts from various fields, which require expression in the surface form, e.g., [SITUATION], [TIME], [SIMULTANEITY], [BAD], [GOOD], [COMPARATIVE], ETC.

Let us turn to P/Y. Goddard (1993) notes that there is no P/Y word that perfectly matches the English 'to hit'. If the striking is done with the hand, then the word is *punganyi*, but 'to hit with a stick or an axe' is *atuni*, and 'to hit with a thrown stick' is *rungkani*. One cannot

translate from English to P/Y without knowing the manner of striking. Translating from P/Y into English presents other problems. As stated, 'to hit with a (thrown) stick' is *rungkani*, but this can also mean 'to grind' or 'to knead'. To make the situation even more complicated, the verb meaning 'to cough', *kuntjulpungyani*, contains *punganyi*, the verb meaning to 'hit with the hand'.

The single word *tjukurpa* (or in Yankunytjatjara, *wapar*) may mean 'something that has been said', or 'a word', but extends to 'a story' and 'Aboriginal Law'. Context and social circumstances must be taken into account, whichever way one is translating. The world does not map neatly from one language to another.

P/Y also has a category of nouns which may be described as Actual/Potential. Since most animals are potentially meat, *kuka* is used for 'edible animal' and 'meat'.

Again, context will determine the translation into English. A *kuka* running about is an '(edible) animal'. The English for its flesh is 'meat'.

A similar semantic problem for translators lies in P/Y verbs. The future tense may express not only futurity, but also possibility, since neither result has yet manifested itself. Where English has two separate 'slices of reality' ('This *will* happen' v. 'This *may* happen'), P/Y makes do with a single syntactic termination, and context determines the meaning.

In Seneca, our third language, *ha?nih* ('my father') may, in fact, refer to males who have a connection to *ego* (the speaker) through a relationship one or more generations above him/her. Thus: 'father', 'father's brother', 'father's mother's sister's son', 'father's father's brother's son', and many more (Leech 1974). Thus the translation of the single word *ha?nih* may equal 'father', 'uncle', 'first cousin once removed (male)', etc.

In another Native American language, Navajo, there is no single verb for the concept [CARRY]. Where English has a simple verb 'to carry', Navajo requires at least ten different verbs, according to the shape and physical properties of the object carried. As Baker (2000: 6-7) notes: ''*Aah* means to carry a solid roundish object such as a ball, a rock, or bottle; *kaah* means to carry an open container with its contents, such as a pot of soup or a basket of fruit; *lé* means to carry a slender flexible object like belt, a snake or a rope; and so on. For this reason, finding the right words to use in a translation to or from Navajo involves much more than simply substituting one string of letters for another. The "Replace All" command on your word processor will never be able to do it properly.' Such linguistic phenomena have profound implications for RC, as we shall see.

5. Idioms

A further problem in matching the 'worlds' of different languages is that of the *idiom*. An idiom, also known as a proverb, saying, axiom, or aphorism, is a type of metaphor, but not of the usual type. Where common metaphors such as 'at the end of the day' (i.e., *ultimately*) or 'grasp the nettle' (i.e., *decide to do something difficult, although it may cause pain or hardship*) are merely a vivid way of expressing one's meaning, idioms (proverbs, sayings, axioms, aphorisms) tend to express a general truth about some aspect of life or give a warning based on such truths. The whole expression does not mean what a literal interpretation would lead one to believe. Thus 'Don't count your chickens before they're hatched' has nothing to do with poultry-breeding, but warns the listener not to do something too early, but to take all relevant factors into account and await a real not

an imagined outcome.

Idioms are peculiar to cultures. Thus, if one has had a bad experience which one does not wish to repeat, the idiom to use in English is *Once bitten, twice shy*. This expresses the resultant behaviour based on the bad experience and may be taken as a general truth or as a warning. In German, however, the idiom is *Ein Gebranntes Kind scheut das Feuer*, i.e., 'A burnt child shys away from the fire.' This is sometimes reduced to *Ich bin ein gebranntes Kind* (i.e., 'I am a burnt child'). In French, the same idiom is rendered as *Chat échaudé craint l'eau*, i.e., 'A scalded cat fears the water.'

In each of these idioms, the *ground* of the metaphor is different, i.e., the particular concept or image on which it is based. In English, it is 'dangerous animals', in German, 'fire', and in French, 'boiling water'. All three idioms, however, convey essentially the same message. The semantic core could be written as: [RESULT] = [NOT] + [AGAIN] + [HAVE] + [BAD] + [EXPERIENCE].

The recipient of a message containing this idiom in its original linguistic form must (a) recognise that the message is not literal, and (b) interpret the underlying meaning, as if expressed by the semantic elements above. Thus a French person who does not speak German and who wishes to send a message in RC to a German recipient a message would not use the symbols for [SCALDED] + [CAT] + [FEAR] + [WATER], since this might not be understood, but those for [RESULT] = [NOT] + [AGAIN] + [HAVE] + [BAD] + [EXPERIENCE].

In Maun (2015) it was suggested that the basic lexical elements for RC should be drawn from Ogden's Basic English (hereafter, BE), as this covered much of the vocabulary of the Swadesh lists and Natural Semantic Metalanguage (Goddard 2010). Using BE, the semantic

core of our idiom may be rendered as [RESULT] = [NOT] + [HAVE] + [TWO] + [TIME] + [BAD] + [EXPERIENCE]. This message could then be converted into RC characters and given a syntactic form using the T-bar, with '[RESULT] =' in Subject position, and '[NOT] + [HAVE] + [TWO] + [TIME]' in Predicator position, as it stands for 'do not repeat', and '[BAD] + [EXPERIENCE]' in Complement position. If this statement were being used as a general truth, there would be no need to mark the Predicator with a Force Level, of the type used by Haag (Maun 2016). If it were used as a warning, however, then it could be marked with the causative (Level III) marker, namely '[', to show that the result would the causation of a non-repetition.

Where the original idiom would be largely iconic in form (animal/fire/cat), the RC version becomes much more abstract and therefore symbolic in form. As noted in Maun (2013), metaphors could be marked in colour to show that they are not literal.

If, however, *idioms* are reduced to their semantic core, as above, then there is no need to use colour, as the actual meaning of an idiom is brought out through the use of semantic primes. Colour, however, could be used to indicate to the reader that he/she could find an idiom in his/her own language. This would not be strictly necessary, as the semantic core should provide all the meaning necessary for understanding.

This means that a sender will type into the RC-device (hereafter, RCD) the original idiom, the device will recognise it as such, because it contains an ontology of such expressions, and will convert the idiom into its semantic core, ultimately producing a coloured syntactic form in RC symbols. The reader will interpret the symbols, mentally converting them into his/her native idiom.

6. Ontologies

Not only will the RCD require an ontology of idioms, it will also require an ontology of the world as expressed through its RC lexicon. This is what Leibniz proposed in 1715 for his *Lingua Characteristica*, 'a sort of general algebra in which all truths of reason would be reduced to a kind of calculus' (Leibniz 1969). As Moro (2016: 112) notes: 'In this case, obviously, it isn't a matter of finding the minimal elements common to a group of established languages, but instead the minimal elements common to all human knowledge.' While such an enterprise still remains a distant ideal, we can at least work with our given lexicon, Basic English, whose elements have been shown to share much with other languages.

Given that we are taking as our basic lexicon the 850 words of Basic English, how can we classify the symbols of RC in such a way that a writer of another language is able to find the ones that he or she needs for his or her message, especially when *lacunae* or idioms are involved? Clearly, the alphabetical list of vocabulary used by Ogden will be of no use to a speaker of any language other than English. Some kind of categorisation will have to be found which will function over all languages. But is this even possible? While photograph-like icons, such as symbols for animals, might be clearly recognisable, it will be far more difficult to know how to classify abstract words and syntactical function words. And how, for instance, are relationships between family members to be listed, when different cultures use different linguistic classifications? Is there some way of categorisation that can be found that is universal, or nearly universal? An examination of historical attempts to divide the world, i.e., to create ontolgies, will go some way to answering this question.

'The world is all that is the case. The world is the totality of facts, not of things.' Thus writes Wittgenstein (1922), at the opening of the *Tractatus*. Despite the fact that the world presents itself to us as a spectrum of images, sounds, and sensations, human beings have, since the dawn of history, attempted to analyse the data before them and to classify its constituent parts into clearly divided discrete concepts as well as classes, categories, types, genera, and many other divisions besides. The ability to classify is universal. 'We have evolved to create and store concepts through signs and to recognize relationships between the signs so formed. […] No culture in the world, no non-pathological human being anywhere, will be found without the ability to generalize.' (Everett 2012: 242)

In order that we may be able to talk about the world, make judgments, and manipulate 'reality', we are obliged to *assume* that such divisions actually exist. Only through categorisation and individuation can we attempt to match our own internal world to that of an interlocutor with whom we are trying to communicate. If the analyses of the world that we make are correct, then we might expect every language in the world to use the same divisions and to have exactly matching words and expressions for the various individual phenomena and the classes, categories, etc. into which a world shared by us all has been divided by the human brain. This of course, is not the case, either for natural languages or for international auxiliary languages. The possibility of the creation of a 'universal' language, or a 'Real Character' in which to write such a language, is thus faced with a number of problems.

6.1. Aristotle and Beyond

An examination of the ways in which various philosophers,

encyclopaedists, and linguists have attempted to classify the world will reveal a number of commonalities and a number of problems for the development of RC.

Any consideration of such attempts must begin with Aristotle, but his thoughts on categorisation could not possibly be compressed into a journal paper of this length. Suffice it to say that his division of the world, knowledge, and the way we talk about these into *substance* (material containing the essence [nature] of a being or 'that which cannot be predicated of anything or said to be in anything', i.e., nouns) and *predication* (that which we can be said of a substance) is perhaps the earliest attempt of any importance at going beyond simple categorial divisions of the world such as *animal* v. *plant*, divisions which he also, but separately, addressed. In Aristotle's system, predication relating to some being may include *quantity* (how much, how many), *quality* (the nature of the being in question), *relation* (how one being is related to another), *place* (the position of a being in relation to another), *time* (position in a chronological sequence rather than in space), *attitude* (e.g., standing, sitting, lying), *habitus* (being in a particular state as the result of action), *action* (doing something), and *passivity* (lit. 'suffering', i.e., being on the receiving end of an action, e.g., 'being struck').

Aristotle's way of looking at the world and his attempts to analyse both the things making up 'reality' and the way in which we look at them influenced every form of intellectual thought up to the seventeenth century, when a more empirical, experimental, and data-driven approach to scientific theory entered the theatre of the mind (Slaughter 1982). Indeed, one of the most important questions to be asked is: Are we classifying concepts (i.e., knowledge) or words (i.e., language)?

In the 17th century, Francis Bacon (1561-1626) in his *Novum*

Organum Scientificarum (Bacon 1620) attempted a categorisation of knowledge, dividing it into (1) External Nature, (2) Man, and (3) Man's Acting upon Nature. Under External Nature he placed astronomy, meteorology, geography, minerals, plants, animals; under Man, anatomy, physiology, powers, and actions; under Man Acting upon Nature, medicine, chemistry, the visual arts, the senses, the emotions, the intellect, architecture, transport, printing, agriculture, navigation, arithmetic, and others (McArthur 1986).

This early attempt at the creation of an ontology essentially formed the basis for future efforts at an analysis of the world. See Rossi (2000) for a thorough examination of ontologies and mnemonic techniques through the ages.

6.2. Education and Language

Johann Amos Kominksy (1592-1670), more generally known as Comenius, included language-learning in his vision of a broad, universal education. To this end, he devised a pedagogical system that would, in his own words, 'follow the footsteps of nature'. He created a system of topics for study which bears striking resemblance to that of Bacon's analysis of reality, without being identical thereto. McArthur (1986) summarises these topics as Comenius listed them in *Ianua Linguarum Reserata* (Kominsky 1631):

Table 1. Comenius's Topics

1.	The origin of the world	11.	The mechanical arts
2.	The elements, the firmament, fire, and meteors	12.	The house and its parts
3.	Waters, earths, stones, minerals	13.	Marriage and the family

4.	Trees, fruits, herbs, and shrubs	14.	Civic and state economy
5.	Animals	15.	Grammar, dialectic and theory
6.	Man and his body	16.	Arithmetic and geometry
7.	The qualities of the accidents of the body	17.	Ethics
8.	Diseases, ulcers, and wounds	18.	Games
9.	The senses (external and internal)	19.	Death and burial
10.	Mind, the will, and the affections (emotions)	20.	Providence, God, and angels

While there are some differences, it is possible to fit Comenius's categories into those of Bacon with relative ease, e.g., Topics 1 and 2 fit into astronomy and meteorology; Topic 5 matches Bacon's 'animals' perfectly, while Topic 8 will fit into Bacon's 'medicine'. Topics 13, 14, and 15 do not fit so easily, but might perhaps be accommodated under 'powers and actions'.

6.3. *A Priori* and *a Posteriori*

The fact that there is no absolute match between Aristotle, Bacon, and Kominsky already suggests that they were looking at the world from different perspectives. It was in part for this reason that the 'language projectors' (Dalgarno, Ward, Wilkins, *inter alia*) devised the idea of a 'Real Character', that is to say, a form of writing which reflected reality. Such a system would, of course, require an *a priori* analysis of nature before the language to reflect it could be produced.

In this, their artificial languages differed from many which followed, which were *a posteriori*, i.e., they were based on existing, natural languages. Libert (2000: 2) points out: 'The distinction between *a priori* and *a posteriori* languages is not a strict dichotomy, but a spectrum; many, if not all, *a priori* languages have elements drawn from a natural language. […] [I]f we insist that an *a priori* language have no *a posteriori* elements, there will be very few or no *a priori* languages […]'.

Wilkins's (1668) *Essay* produced a categorical summation of the world in 40 genera, set out in 'Philosophical Tables'. These were (using McArthur's (1986) simplified version):

> 1. God; 2. the elements, meteors, stones, metals; 3. plants, herbs, flowers, shrubs, trees; 4. animals, fishes, birds, beasts; 5. parts of bodies; 6. quantity, magnitude, space, measure; 7. quality (of natural power, habit, manners, the senses, diseases); 8. action (spiritual, corporeal, in motion, in operations); 9. relations (in the family, regarding possessions, and provisions); 10. public relations (civil, judiciary, naval, military, ecclesiastical).

In theory, if this analysis of 'reality' were correct, it should have been the last word in scientific exactitude and all Real Characters created thereafter would have necessarily been based upon it. This was not the case. Language does not follow science.

Stillman (1995: 243-244) notes: '[Wilkins's] philosophical tables detail relations between individual things and groups, between subordinate groups and major groups, in such a way as to configure (ideally) a complete network enumerating the sum total of relations among the things and notions. […] Knowledge about any one thing leads directly, through the predictive agency of the tables, to a

knowledge of its relationship to all other things.' That, at least, was the theory.

In order to make his Real Character logical and readable, each of Wilkins's categories was given a letter, which, in ideographic form, assumed a symbolic shape on which could be marked Aristotelian differences, of which there were, quite arbitrarily, a maximum of six allowed. This convention of marking categories was much adopted in early artificial languages, in which prefixes or suffixes served the function of identifying catgories. Thus in Vidal's *Langue universelle et analytique* (1844) there are 20 major categories, the words in each one being indicated by a different capital letter. Thus: N – measure, matter, form, movement; Z – plants; B – animals; etc. (from Couturat & Leau (1903: 44) – my translation).

In part this corresponds with some features of certain natural languages. For instance, Bantu languages, such as Luganda, have categories of nouns marked according to reference (e.g., people, animals), their shape (e.g., round, cylindrical), or their size or material (e.g., small, large, liquid). Mandarin also has category markers. DeFrancis (1984: 47) notes: 'Just as foreign students of English have to memorize phrases like "a flock of sheep", "a herd of cattle", "a crowd of people", so students of Chinese must memorize that *zhang* is the appropriate measure word for flat objects like paper and tables whereas *tiáo* is the measure word for long narrow things like snakes and roads.'

Two problems arose with the system of markers in artificial language systems. Firstly, all words within a category looked similar, and secondly, the categories did not necessarily correspond with Wilkins's (allegedly universal) top-level analysis of 'reality', but, rather, they picked out sub-categories such 'agent', or 'male'. Esperanto, for instance, has no major semantic category markers in

each word (being an *a posteriori* language, not an *a priori* one) but uses affixes such '-ar-' to indicate 'a collection' (*arbo* = tree; *arbaro* = forest). The various analyses of reality attempted by the authors of artificial languages were thus all out of focus with each other, none giving an absolute and definite picture of the world. This, of course, was in part because each author was either seeing the world through his own native or academic language and dividing it accordingly, as in the case of Bacon (Latin), Wilkins (English), and Vidal (French), or dividing the world according to a particular root, e.g., Indo-European languages in the case of Zamenhof (Esperanto).

6.4. Roget: Then and Now

The work of Peter Mark Roget (1779-1869) differs from that of the 'language projectors' and the inventors of *a posteriori* artificial lanaguges in that the purpose of his classification was completely different. Moving away from simple alphabetical lexicography, Roget's desire was to present a list of words in the English language 'not in alphabetical order, as they are in a Dictionary, but according to the ideas which they express'. The object of his *Thesaurus of English Words and Phrases* (1852) was expressed thus: 'The idea being given, to find the word or words, by which that idea may be most fitly and aptly expressed. For this purpose, the words and phrases of the language are here classed, not according to their sound or their orthography, but strictly according to their signification.' (Roget 1852: v)

Roget began with a six-fold division of vocabulary: Abstract relations, Space, Matter, Intellect, Volition, and Affections (i.e., emotions). Each classification was then divided into sub-headings, thus:

- Abstract relations: existence, relation, quantity, order, number, time, change, causation
- Space: generally, dimensions, form, motion
- Matter: generally, inorganic, organic
- Intellect: formation of ideas, communication of ideas
- Volition: individual, intersocial
- Affections: generally, personal, sympathetic, moral, religious

Each of these sub-heading was further divided into sub-sub-headings and these latter into sub-sub-sub-headings. An alphabetical list of the words of the language which were covered in the work was provided in the second half of the book to enable a reader to search for any particular term.

While the philosophical purpose of this work was entirely different from that of Wilkins (and by implication, Aristotle), Roget acknowledged his debt to the former, and expressed the hope that further work could be undertaken to create 'a *Polyglot Lexicon* constructed on this system' (Roget 1852: xxiii). Such a lexicon has yet to be created but will ultimately be a necessary element of an RC, given the lack of overlap between languages.

Roget's system of classification, was, of course, constrained by the philosophical and scientific methods and mind-sets prevalent at the time. The 1996 edition of *Roget's International Thesaurus* (Roget 1996) reveals a wider and more science-oriented classification in its main headings:

The body and the senses; Feelings; Place and change of space; Measure and shape; Living things; Natural phenomena; Behaviour and the will; Language; Human society and institutions; Values and ideals; Arts;

Occupations and crafts; Sports and amusements; Mind and
ideas; Science and technology.

Some lexical areas have clearly moved from a sub-category to a
category in the more modern version. Thus Class II, 'Space', with
four subcategories in the original work, is covered in the modern
version by the heading 'Place and change of space' and 85 finely
divided minor categories. The modern category of 'Science and
technology' takes in concepts such as *radar* which did not even exist
in Roget's time. As the world changes, so ontologies must change.

6.5. Sivartha, Haag, and Hartrampf

In 1884, Alesha Sivartha (a.k.a. Holmes W. Merton) published, in
a somewhat pretentiously-titled work, *The Book of Life* (Sivartha
1884/1912), a synthesis of his religious, mystical, social, and
educational views. According to the title page, this was, 'a collection
of Discoveries of 1859 to 1878'. During this period, Sivartha made a
number of attempts at classifying human knowledge and ultimately
proposed the sketch of an *a priori* universal language named Vesona,
based on what he described as 'a Universal Synthesis of human
knowledge'. The circular diagram (see Figure 1) 'is arranged so as to
display those relations and analogies which unite each branch to the
rest'. (Sivartha 1912: 325)

Sivartha's four basic divisions are those of Life, Object, Property,
and Motion, each category being given a prefix to identify its
members: *ano-*, *ako-*, *ado-*, and *aso-*. Each category is sub-divided
and given an affix, and then further sub-divided, with an affix being
given to these sub-sub-divisions. Sivartha (ibid.) states, 'The first
two or three letters of any word give the general meaning. And the

Figure 1. Sivartha's Categories.

added letters specialize these meanings'. Needless to say, these divisions and sub-divisions do not match those attempted by previous 'language projectors'. The world has been cut into a different set of slices.

Sivartha attempts no apportionment of concepts or vocabulary to the categories and their sub-divisions. We do not know how he would treat the problems that we have encountered such as verbs of travel (German), kinship (Seneca), or verbs of hitting (P/Y) or how we would have assigned the various lexical items to his categories. Sivartha's analysis, unlike that of Haag (1902), lacks any sound theoretical basis. While it is historically interesting, this ontology is both empty and unproductive. It would certainly not serve as a basis for RC.

Haag's analysis of concepts and divisions of the world (1902) has been dealt with in some depth by Maun (2016). Starting from bodily perceptions of the world, Haag uses the basic concepts that he discovers, e.g., near, far, inside, outside, as metaphors within a number of categories, namely: *Space*; *Time*; *Degree*; *Type*; *Logic, Causality*; *Mechanics*; *Chemistry, Material*; *Life*; *Feeling*; *Thought*; *Volition*; *Action*; *Physical Geography*; *Astronomy*; *Anatomy and Physiology*; *Zoology*; *Botany* and *Economics*.

Haag's analysis was based on a firm theoretical foundation, namely that knowledge starts with physical awareness of the world. This lends strength to his arguments, but his book *Towards a Logically-based Graphical Language* (1902) was only a tentative sketch, an outline which was never expanded, even in his later work *The Cognitive Basis of Language* (*Das Denkgerüst der Sprache*) (1935).

Hartrampf's *Vocabularies* (Hartrampf 1929) more closely resembles the work of Roget (1852) than that of Sivartha or Haag. Its

aim is to improve people's use of language by providing a volume of synonyms, antonyms, and relatives. The theoretical and organisational basis more closely resembles that of Haag (1902) than that of Roget. As illustrated on the circular *Idea and Word Chart* (see Figure 2), Hartrampf divides up the human experience of the world (in its largest sense), starting from *Sense Conduct or Quality of Sense Reaction.* He takes the 5 sensory causes, *Food*, *Odour*, *Touch*, *Show-Light*, *Speech-Sound*, and the effect that these cause on the human being: *Tasting*, *Smelling*, *Feeling*, *Seeing*, and *Hearing*. Readers are then directed from the *Chart* to the various sections of the book (529 pages).

His other divisions relate to *Character of Experience*, and these are arranged in pairs of opposites or relatives on the chart. Thus: *Passage-Passageways*, *Opposition-Desire*, *Give-Take*, *Division-Unity*, *Change-Stability*, *Reduction-Promotion*, *Disorder-Order.* As with the upper division of the chart, readers are directed to the appropriate section of the book which deals with these concepts. Thus a reader seeking words on *Seeing* (section 119) will be directed to *Vision*, *Eye*, *Search*, *Discovery*, and *Watchfulness*. Under *Vision*, he/she will find verbs, adjectives, and nouns, relating to vision (119A). Under 119B are to be found adjectives and nouns relating to *Defective Vision*. Since this is a thesaurus rather than an encyclopaedia, the entries for each section give alternative words rather than descriptions or definitions, many entries being circular e.g., *blink = get a glimpse; glance = look hurriedly; glimpse = glance.*

Hartrampf's approach to language, and thereby the knowledge that it expresses, differs from all other such analyses in both its method of analysis and its form of presentation. Like all such analyses, however, its divisions are, ultimately, arbitrary.

Figure 2. Hartrampf's Idea and Word Chart

IDEA & WORD CHART
FROM
HARTRAMPF'S VOCABULARIES
by
Gustavus A. Hartrampf

APPETIZING
TOOTHSOME
DELECTABLE
EXQUISITE
SWEET
LUSCIOUS
AMBROSIAL
DELICIOUS

EAT
BANQUET
FEAST
DINE
GORMANDISE
GORGE
GLUT
SATE

BRIGHT
FLASHING
GLEAMING
RESPLENDENT
BRILLIANT
SCINTILLANT
SPARKLING
DAZZLING

FOODS	48A
SAVOURINESS	48P
UNSAVOURINESS	48Q

EATING	51A
DRINKING	51B
INHALING	51A

ZEST
RELISH
PLEASURE
ENJOYMENT
GUSTO
DELIGHT
ECSTASY
THRILL

| FRAGRANCE | 122 |
| STENCH | 123 |

| SMELL (DETECT ODOUR) | 120 |

CONTACT	124
CARESS	98C
INFLUENCE	135B

SENSIBILITY	121
THRILL (JOY)	71A
PAIN	31

SHOWING	95A
PICTURES	100B
BEAUTY	108
BRILLIANCE	101A
COLOURS	102A

VISION	119A
EYE	119C
SEARCH	119E
DISCOVERY	119F
WATCHFULNESS	119G

MELODY	68A
LAUGHTER	71B
BRAG	76
PROPHECY	77
TRUTH	61A
PRAISE	63

| HEARING (TO HEAR) | 118 |

DEPARTURE	15
SPEED	9
ASCENT	17
DESCENT	22
CONVEYANCE	24A

OPENINGS	25
PATHS	25B
STEPS	25C
ORGANIC PASSAGE	25E
ARTIFICIAL PASSAGE	25F

PROHIBITION	35
BANISHMENT	35A
HATRED	28
ANGER	29
DEFENCE	36

LIKING	38
LOVE	39A
PASSION	39B
PERSUASION	40A
DEMAND	41B

GENEROSITY	43A
REWARD	45
PERMISSION	46A
FORGIVENESS	47

SEIZE	49C
ACQUIRE	49A
ACCEPT	49B
ACHIEVE	50

SEPARATION	145A
SPREADING	145I
QUANTITY	147A
TIME	148A
PLACE	149A

UNIFICATION	146A
MIXTURE	146C
WEDLOCK	146H
PARTNER	146L
ASSEMBLAGE	146K

CONVERSION	3A
TENDENCY	2
START	4

| CONSTANCY | 11A |
| REST | 11B |

DECREASE	14
DIFFICULTY	13
PREVENTION	12

CAUSE	1
PROMOTION	6
AID	5A

LIBERAL
LAVISH
BOUNTEOUS
UNSTINTED
GENEROUS
MAGNANIMOUS
MUNIFICENT
PRINCELY

CONSTANT
LASTING
EVERLASTING
ENDLESS
PERPETUAL
DEATHLESS
IMMORTAL
ETERNAL

CONFUSION	125A
EXCESS	125C
ERROR	126C
CLUMSINESS	126D
DIRT	129A
BASENESS	129B
CRIME	130A

CONCORDANCE	132A
TEMPERANCE	132C
ACCURACY	134A
SKILL	134B
PURITY	138C
EXCELLENCE	140B
GOODNESS	140A

BAD
BASE
CORRUPT
VILE
EVIL
ABOMINABLE
AWFUL
HEINOUS

EXCELLENT
EXCEPTIONAL
MATCHLESS
PEERLESS
INCOMPARABLE
EXTRAORDINARY
SUPREME
SUBLIME

Inner wheel: QUALITY OR — FOOD — OF SENSE — REACTION — CHARACTER OF EXPERIENCE — 5 SENSES — KEY WORDS — CAUSES — SENSE — CONDUCT — TOUCH — ODOUR — TASTING — SMELLING — FEELING — SEEING — HEARING — PASSAGEWAYS — DESIRE — TAKE — UNITY — STABILITY — PROMOTION — ORDER — DISORDER — REDUCTION — CHANGE — DIVISION — GIVE — OPPOSITION — PASSAGE — SPEECH · SOUND — SHOW · LIGHT

KEY WORDS TO ALL SPEECH
SYNONYMS — ANTONYMS — RELATIVES

When you have read the preceding pages you will have all the information needed for immediately increasing the sparkle, the charm, and the effectiveness of your expression. But if you wish to achieve absolute mastery of thoughts and words, you will find the Idea and Word Chart indispensable. This presents 12 word-pairs (24 keywords) that miraculously reflect, in words, the fundamental qualities in all ideas. If you give this magical chart indelible residence in your mind, you will never again have to grope hopelessly for words.

7. Syntactic Problems

We can already see that the design of an RC meets problems when one attempts to divide the world. There simply is *no* universal classification system that would enable the designer to satisfy all speakers of all languages, and no way to divide Basic English which will similarly match all languages. But that is not the end of the problem. There is the question of syntax.

The way in which speakers of various languages convey a message may vary syntactically rather than semantically. Imagine a film-clip in which a man steps into a river and then reaches the other bank under his own propulsion. An English speaker summarising this clip would say:

(1)   *He swam across the river*
      (Verb + Preposition + Noun)

A French speaker, however, would say 'He crossed the river by swimming';

(2)   *Il a traversé la rivière à la nage*
      (Verb of directional movement + Noun + Adverbial phrase of manner)

Similarly, if the film shows a woman very rapidly climbing stairs, the English speaker would summarise the clip as:

(3)   *She ran upstairs*
      (Verb of motion + Adverb of place)

The French speaker would say 'She mounted the stairs by running';

(4)   *Elle a monté l'escalier en courant*
      (Verb of directional movement + Noun + Adverbial phrase of
      manner)

In the first example, there are three semantic elements in the grammatical predicate are: [CROSS], [RIVER], and [SWIM], and in the second example the three elements are: [RUN], [UP], [STAIRS]. If RC is to be read by a speaker of any language, then the symbols used to represent these three concepts must be readable in the order of the reader's native language.

This suggests that the RCD must be able to unscramble the semantic elements i.e., [CROSS] and [SWIM] and place them in the correct order for the recipient. Such a procedure, however, would only be possible if the RCD knew what the target language was. Thus an English speaker would write his/her message thus:

(5)   [HE] + [SWIM] + [PAST] + [CROSS] + [RIVER]

For a German recipient, the RCD would retain this order (*Er schwamm über den Fluss*), but for a French recipient the order of elements would be:

(6)   [HE] + [CROSS] + [PAST] + [RIVER] + [SWIM]

Choosing in advance the necessary options in the RCD to enable a particular reader or group of readers to read a message is something of a paradigm shift in the conception of RC. It is a major departure

from the original conception of RC as being able to create messages readable by a speaker of *any* language. We see from such examples as the above, from *lacunae*, from idioms, and from attempts at encyclopaedic classification, that it will be necessary to 'focus' or 'tune' RC for specific readers.

8. Real Character – Modifying Intentions

The plethora of 'world-divisions' and 'category-divisions' revealed in analyses of natural languages and international auxiliary languages shows that there is no single 'map' of the world which will suit all languages, all linguists, all philosophers, and all scientists. The best that we can do is settle on a set of elements and symbols, combined with an agreed syntax, which will permit the interpretation of a message by its destinee and the interpretation of the core of a message by most people, as well as and enabling writers to compose a message for a given recipient.

It will remain to complete Roget's *Polyglot Lexicon* in which central concepts in RC will be defined, with additional or unnecessary elements defined, depending on the language in focus, i.e., the language of the intended reader. It will not be possible to cater adequately for readers beyond the intended recipient of an individual language, although other readers may be able to understand the passage, either partially or in full, since RC characters will be employed.

The original concept of a Real Character as the 'language projectors' saw it was based on the Aristotelian belief that 'the relationship between the mind and the world was natural, and common to all human beings; the relationship between the mind and

language was conventional, and differed to the extent languages differed from one another. […] Wilkins believed that it was possible accurately to map the order of thought (and therefore of the things which this represented), and it was thus possible to devise a language which might represent this. Such a language would be commonly understood.' (Lewis 2007: 157)

Our brief foray into translation, idiom, and ontologies has shown that an international language commonly understood by all speakers of the world's languages is, in fact, an impossibility. What emerges from the studies undertaken is that the means of *coding* messages is possible, through the use of RC elements and characters, but that an RCD will be required to adjust any source-language message in accordance with the language of the recipient. One cannot create a single message which will be understood by everybody. The conceptual, semantic, linguistic, and cultural differences between various peoples preclude such an idealistic vision.

Since no language can perfectly mirror 'reality', it is impossible to create an artificial language which does so. This means that an artificial language must be strictly governed by conventions. This more realistic vision of a set of conventionalised images and symbols that can be combined into an adjustable set by means of an RCD gives a much more promising future for RC. Thus, elements of semantic metalanguage will be chosen, such as icons, to which will be added chosen symbols based on BE, Haag, and NSM. Symbols for abstract notions will be composed of the latter according to formational parameters, e.g., nouns above adjectives.

9. Icons and Cultures

Once an ontology of such icons and symbols has been created, it will serve as a box of building-blocks from which all other symbols are made. With regard to learning, Unger (2004: xv) notes that, for Japanese, 'If you want to read and write Japanese, then *kanji*, as the Japanese call them, *simply must be learned*' (my emphasis). The same is true of RC. It will be necessary to accept conventionalised characters and the formational parameters that govern them.

This limitation is the price that must be paid for finding a way through the difficulties recognised by earlier 'language projectors', who concluded that a universal language could not possibly exist. What we have instead is a universal code which can be employed to make a message from a writer of Language A comprehensible to a reader of Language B who has no knowledge of Language A.

10. Vocabulary – Modification and Clarification

It has been established that the 850-word vocabulary of Basic English will form the foundation of RC. Ogden, however, devised his system in the 1920s, and although motor-propelled vehicles existed for both land and air at that time, there are no words in BE for *car*, *lorry*, *truck*, *motor-bike*, etc., nor do we find *aircraft*, *aeroplane*, *plane*. More specialised vehicles that we are familiar with today such as *tanker*, *artic* (articulated lorry), *taxi*, *gritter*, etc. are also absent, as are all military vehicles, and even the ubiquitous *bicycle*. In the 21st century, we cannot make do with Ogden's sole vehicular words, *carriage* and *cart*. Nor can we do without basic terms of information

technology such as *computer*, *mobile phone* (*cell-phone), i-Pad*, or *tablet*. Once we have established what Basic English really is today, we shall be in a position (a) to create symbols for all words, (b) to decide on semantic divisions so that users can locate icons and symbols, and (c) to devise sub-lexicons for individual target languages.

Having established that any semantic classification or ontology of the world is, ultimately, arbitrary, we can now decide on what will be an entirely conventionalised classification of BE. Ogden's original categories of BE are *Operators* (verbs, prepositions, conjunctions); *Things - 400 general words* (e.g., account, act, addition); *Things - 200 picturable things* (e.g., angle, ant, apple); *Qualities - 100 general* (e.g., able, acid, angry) and a category consisting of opposites (e.g., angry, awake, bad). In this latter category the opposite is not given. It is obviously assumed that the user knows the opposite meaning.

11. Are Icons Enough?

A first move might be to insist that all picturable objects become icons. Simple as that sounds, however, problems arise when cultures become involved. *Bread* does not look the same in Britain and the Middle East. A *cow* in England does not look exactly like a cow in India. It will therefore be necessary to have a basic but modifiable icon according to the target language. Thus a basic *cow* symbol will need to be modified for the Indian version, with broader horns and a deeper dewlap. *Bread* will appear in a cuboid form for English, a long form for French and a flat form for Middle Eastern cultures. Thus if the sender is an English-speaker and the target language is Arabic, when 'bread' is typed into the RCD, the device will

automatically choose the flat-bread icon.

Ogden's categories show some overlap, e.g., some foodstuffs such as *bread*, *butter*, and *milk* appear in the *Things-general* category, while others appear in the *Things-picturable* category. Given the variety of shapes of manufacture and packaging of the various foods, it is unclear why *bread* is not given as picturable, while *cheese* is. Nevertheless, we are not bound by Ogden and must move on to a system of categorisation that will work better for RC.

Without defining all possible categories into which Basic English may be sorted, we can find a few large-scale ones, e.g., FOOD, KINSHIP, BUILDINGS & THEIR SOCIAL INSTITUTIONS, TOOLS. While these may be 'obvious' categories, problems still remain. While a pig may be seen as FOOD in some cultures, in others its meat is forbidden. The same is true of the cow, which is seen as holy in Hindu cultures, and may therefore not be harmed in any way. There must therefore be a switching mechanism between languages where these categories are incompatible. Thus if an English message contains the word 'pig', and the target-language is Arabic, Hebrew, or any other language whose culture regards this animal only as an animal, then the RCD must remove any specification of [MEAT] from the RC symbol.

12. Symbols and Supplementals

An analysis of target languages and cultures will be needed before an RCD can be fully programmed with icons. The same is true of symbols.

Given that, for instance, the word *go* in BE has no exact equivalent in German, it is necessary to form a sub-lexicon for German. When

an English-speaker wishes to send a message to a German-speaker, he/she will select 'German' from the Target-Language Menu. Thus, if an English-speaker writes to a German-speaker *I went to Manchester*, the RCD will (a) recognise *went* as the past of *go*, and (b) show prompts to the writer illustrating means of travel. The RCD will then show the sender a choice of supplementary characters. Using symbols and icons, these will represent [GO] (symbol) + [MEANS] (icon). Thus, [GO] + [FOOT], [GO] + [VEHICLES], [GO] + [PLANE], etc. The writer chooses the appropriate symbol to add to *went* and the RCD then composes the necessary symbol to enable the recipient to understand.

Such a way of looking at the notion of *to go* might seem somewhat unnatural to an English-speaker, but conventionalisation means that particular characters *must* be established for particular meanings.

An English-speaker will probably never have thought about the shape or property of the object that he/she is carrying, but if a message containing the verb *to carry* is to be sent to a speaker of Navajo, then a selection from the symbolic representations of *'aah*, *kaah*, *lé*, etc. will have to be made. If we adopt the letter W to represent bent arms carrying something, then when an English speaker types in *carry* (or *carried*, *will carry*, etc.) the RCD will select a symbol such as Wo or W__ to represent the carrying of a round object, a long object, depending on the syntactic Direct Object given. If the message was destined for a speaker of French or German, then no such adjustment would be required, a fact recognised by the RCD, which would simply choose the W symbol for *to carry*.

For P/Y, an English-speaker writing *hit* will be offered an icon menu showing a fist, a static stick, a thrown stick, etc. Once the precise meaning of the message has been specified, the RCD

composes the necessary symbol accordingly.

Thus the RCD would be provided with information as follows. (At this stage, all symbols are provisional, of course.) For individual languages, the additional symbols would be mandatory.

Table 2. Specifying Target-Language Symbols

Verb	Target language	RC Foundation symbol	Additional symbols
CARRY	French	W	None
	German	W	None
	Navajo	W	+ [o] / + [-] / + [□], etc.
GO	French	→	None
	German	→	+ [CAR icon] / + [PLANE icon], etc.
HIT	French	-|**	None
	German	-|**	None
	P/Y	-|**	+ [FIST icon] / + [STATIC STICK icon] / + [THROWN STICK icon], etc.

13. Kinship

The problems connected with the linguistic expression of kinship have been very briefly outlined above, using Seneca relationships as an example. Terminology varies from one culture to another. There is a possible solution to the expression of kinship terminology for RC in the visual format of a 'kinship chart'. This resembles a traditional Western family-tree with an iconic code to assist the reader.

Everett (2012: 247) explains how to read such a chart, in which 'ego', the speaker or person spoken about, is marked centrally with a shaded black and white circle: 'The equals sign connects two individuals that are married. Women are represented by circles and men by triangles. Vertical lines indicate children. A vertical line descending from an equal sign indicates a child of that marriage.' Thus, in RC, a small part of such a diagram would serve as an icon. Whether for instance the Seneca term *ha?nih* means 'father' or 'father's brother' would be indicated by a line tracing a route from 'ego' to the relative in question on this mini-map. No complicated symbolic explanation would be required. The recipient of the message would *see* the relationship, not read it.

14. Writing It All Down

The procedure for writing a message in RC will follow a definite sequence. The language of the destinee of the message will firstly be specified before writing begins. This will enable the RCD to detect adjustments that will need to be made, just as Chinese speakers of Mandarin writing a text-message today type in Romanized pin-yin and are offered a choice of Chinese characters where there are tonal differences (e.g., *ma* means 'mother' or 'horse', *inter alia*, depending on intonation). The writer then selects the appropriate character for the message (Maun 2013). In RC, however, the RCD will offer an appropriate symbol to the recipient according to the target language, e.g., for Navajo, Wo if the syntactic Direct Object is round by nature or W__ if it is long. If the message is to be copied to speakers of other languages, then those languages will also have to be specified to allow processing by the RCD.

As outlined in Maun (2013), the RCD contains drop-down menus that enable the writer to drop symbols into semantico-syntactic slots on the T-bar, e.g., Subject. Frequently occurring lexical items such as 'I', 'you', etc., are also offered. In cases of synthesis in which the subject is incorporated into the verb, e.g., 'Voy' (= I go) in Spanish, the RCD will recognise the source form. If the target language of the RC reader is analytical, e.g., French ('Je vais'), the RCD, having information about the TL, will provide symbols for 'I' and 'go' in the RC representation, placing them in Subject and Verb positions.

15. Conclusions

Lacunae, idioms, and ontologies show that there is no perfect match between languages. All map the world slightly differently. The idea of a 'universal' Real Character, laudable as it is, is simply not practicable. Indeed this conclusion was reached by many 'language projectors', and the consequence of this was a move towards international auxiliary languages, such as Esperanto, Ido, and Interglossa.

A modified RC will be based on BE, not so that it has an occidental basis, but because Ogden's system contains many elements which seem to be present in most of the world's languages.

Given the peculiarities and quirks of individual languages, it now seems more likely that a focused version of RC will be possible. Since such a means of communication will be transmitted digitally, using icons and symbols, an RCD may be programmed to take account of individual semantic and syntactic features of particular languages. At the same time, much of the iconic and symbolic system will be common to all languages, and the formational

parameters for both symbols (e.g., Haag elements) and syntax (e.g., T-bar or I-bar) will be defined as a central tenet of the system.

As international relations become ever more complicated, a truly international means of communication becomes ever more desirable. It is to be hoped that the development of a modified RC will go some way towards endowing people with just such a system.

References

Bacon, F. 1620/1902. *Novum Organum Scientificarum*. Published in English 1902. New York: P.F. Collier, Inc.

Baker, M. 2000. *The Atoms of Language. The Mind's Hidden Rules of Grammar*. New York: Basic Books.

Berlin, B. & P. Kay. 1969. *Basic Color Terms*. Berkeley & Los Angeles: University of California Press.

Chilton, P. 2014. *Language, Space, and Mind. The Conceptual Geometry of Linguistic Meaning*. Cambridge: CUP.

Couturat, L. & L. Leau. 1903. *Histoire de la langue universelle*. Paris: Hachette. Reprint on demand.

DeFrancis, J. 1984. *The Chinese Language. Fact and Fantasy*. Honolulu, HI: University of Hawaii Press.

Deutscher, G. 2010. *Through the Language Glass. How Words Colour Your World*. London: Heinemann.

Everett, D. 2012. *Language: The Cultural Tool*. London: Profile Books.

Goddard, C. 1993. *A Learner's Guide to Pintjantjatjara/Yankunytjatjara*. Alice Springs: IAD Press.

Goddard, C. 2010. NSM. Available at URL <http://www.griffith.edu.au/humanities-languages/school-languages-linguistics/research/

natural-semantic-metalanguage-homepage>.

Haag, K. 1902. *Versuch einer graphischen Sprache auf logischer Grundlage*. (*Towards a Logically-based Graphical Language*.) Stuttgart: W. Kohlhammer.

Haag, K. 1935. *Das Denkgerüst der Sprache. Die Bestandteile des Denkens als die Grundlagen des Sprachbaus.* (*The Cognitive Basis of Language. The Elements of Thought as the Foundations of the Structure of Language*.) Heidelberg: Carl Winters Universitätsbuchhandlung.

Hartrampf, G. 1929/1956. *Hartrampf's Vocabularies. Synonyms, Antonyms*, *Relatives*. 35th Printing. Marple: Psychological Publishing Co. Ltd.

Hogben, L. 1946. *Interglossa*. Harmondsworth: Penguin.

Kominsky, J. [Comenius]. 1631. *Janua Linguarum Reserata*. Leszno, Poland.

Lakoff, G. & M. Johnson. 1980. *Metaphors We Live By*. Chicago, CA & London: The University of Chicago Press.

Lakoff, G. & M. Johnson. 1999. *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought.* New York: Basic Books.

Leech, G. 1974. *Semantics*. Harmondsworth: Penguin.

Leibniz, G. 1969. *Philosophical Papers and Letters*. 2nd Edition. Translated by L. Loemker. Dordrecht: Reidel.

Lewis, R. 2007. *Language, Mind, and Nature: Artificial Languages in England from Bacon to Locke*. Cambridge: CUP.

Libert, A. 2000. *A Priori Artificial Languages*. Munich: Lincom Europa.

Lounsbury, F. 1956. A Semantic Analysis of Pawnee Kinship Usage. *Language* 32.1, 158-194.

Maun, I. 2013. Towards a 'Real Character' for the Computer Age.

Journal of Universal Language 14.1, 143-184.

Maun, I. 2015. 'Real Character'. Towards a Lexicon. *Journal of Universal Language* 16.1, 87-130.

Maun, I. 2016. Karl Haag, Modern Linguistics, and 'Real Character'. *Journal of Universal Language* 17.2, 53-110.

McArthur, T. 1986. *Worlds of Reference. Lexicography, Learning, and Language from the Clay Tablet to the Computer*. Cambridge: CUP.

Moro, A. 2016. *Impossible Languages*. Cambridge, MA: MIT Press.

Ogden, C. 1930. *Basic English*. Cambridge: The Orthological Institute.

Roget, P. 1852. *Roget's Thesaurus of English Words and Phrases*. Reprinted facsimile of the first edition, 1992. London: Bloomsbury Books.

Roget, P. 1996. *Roget's International Thesaurus*. Glasgow: Harper Collins.

Rossi, P. 2000. *Logic and the Art of Memory*. *The Quest for a Universal Language*. London: Continuum. Originally published as *Clavis Universalis: Arte della Memoria e Logica Combinatoria da Lullo a Leibniz*, 1983. Bolgna: Società editrice il Mulino.

Sivartha, A. 1884/1912. *The Book of Life. The Spiritual and Physical Constitution of Man*. 7th Edition. Reprint on demand. New York: Holmes W. Merton.

Slaughter, M. 1982. *Universal Languages and Scientific Taxonomy in the Seventeenth Century.* Cambridge: CUP.

Stillman, R. 1995. *The New Philosophy and Universal Languages in Seventeenth-Century England. Bacon, Hobbes, and Wilkins*. London: Associated University Presses.

Swadesh, M. 1950. Salish Internal Relationships. *International*

Journal of American Linguistics 16, 157-167.

Swadesh, M. 1971. *The Origin and Diversification of Language*. Edited post mortem by Joel Sherzer. Chicago, CA: Aldine.

Unger, J. 2004. *Ideogram. Chinese Characters and the Myth of Disembodied Meaning.* Honolulu, HI: University of Hawaii Press.

Vidal, E. 1845. *Langue universelle et analytique.* Paris: Sirou. Quoted L. Couturat & L. Leau 1903: 44.

Wilkins, J. 1668/2002. *An Essay Towards a Real Character and a Philosophical Language*. London: Sa. Gellibrand. Reprinted, 2002. Bristol: Thoemmes Press.

Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus.* Revised Edition, 1974. London: Routledge.