

Proceedings EUROCALL '95

7th, 8th and 9th September, 1995

*TECHNOLOGY ENHANCED LANGUAGE LEARNING:
FOCUS ON INTEGRATION*

Edited by
Ana Gimeno

UNIVERSIDAD POLITÉCNICA DE VALENCIA
SPAIN

1
EU

TECHNOLOG

© Los autores

© De esta edición: Ana María Gimeno Sanz

Diseño de cubierta: Juan Manuel Estévez Doménech

Edita: SERVICIO DE PUBLICACIONES
Camino de Vera, s/n
46071 VALENCIA
Tel. 96-387 70 12
Fax. 96-387 79 12

Imprime: REPROVAL, S.L.
Tel. 96-369 22 72

I.S.B.N.: 84-7721-421-2
Depósito Legal: V-3685-1996

UNIVERSIDA

Servic

**CANNOT WITHOUT PROCÉSS OF SPEECH BE TOLD:
LEARNING FROM THE FAILURES OF COMPUTATIONAL MODELING**

Willard McCarty
University of Toronto

"Ist es ihr Gesang, der uns entzückt oder nicht vielmehr die feierliche Stille,
von der das schwache Stimmchen umgeben ist?"¹

Franz Kafka, "Josefine, die Sängerin oder Das Volk der Mäuse"

"*Cannot without procéss of speech be told*": the sensitive ear, tuned to the cadences of English, will hear the lurching of this line to its viscous end, and perhaps before then a hint of the serpent's hiss. The line comes from John Milton's epic poem, *Paradise Lost*. In daring poetry, which even today seems risky, Milton has his God both appear and speak, and thus poses in a 17th-century mortal idiom the perennial enigma of language. Through the mouth of the angel Raphael to the ears of prelapsarian Adam and Eve, Milton tells his story of how God commanded the Son to create the world: "So spake th' Almighty", Raphael narrates,

...and to what he spake
His Word, the Filial Godhead, gave effect.
Immediate are the Acts of God, more swift
Than time or motion, but to human ears
Cannot without procéss of speech be told,
So told as earthly notion can receive (7.174-9).

Here Milton measures the scope of language by the ratio of divine to human speech and so locates our defining ability as *reordberend* (O.E. "man", lit. "bearer of language") in the transcendent. Few of us would do the same in quite those terms now, but when we attempt its measure, language is no less astonishing to us,² and so some figure of transcendence remains necessary, even in a world shaped by the computer. Without the paradoxical sense that

¹ "Is it her singing that entrances us, or is it nothing more than the solemn stillness with which her weak little voice is surrounded?" (my transl.).

² "Again, as in the matter of the prodigality of languages, the proper start is wonder, a tensed delight at the bare fact..." (Steiner 145). *After Babel* is a powerful articulation of that wonder and an effective introduction to the problems so briefly considered here.

language "demands more than humanity possesses" (as Dr. Johnson said),³ our understanding of it is blunted. Even our ability to compute language is crippled by a mundane view of speech, and the results rendered more or less trivial. Computing has of course most severe limitations, but though its scope of action is confined to the thin surface of the vast ocean of language, to get very far with it we must know the depths and steer by the stars.

Research and practice in computer-assisted language learning (CALL), I would argue, is no different. In small, it turns on a version of Milton's ratio, applying the machine to help the language student progress from stumbling reluctance toward an ideal native fluency. That, of course, is your concern. I am not here to tell you, as experts in CALL, what you already know, and know far better than I ever will. Rather, my assignment is to deliver a *key-note* speech, hence by definition to set "the prevailing tone of thought" for the conference (*OED* "key-note" 2.a). I cannot question the wisdom of Professor Davies in selecting me for the honour of being here, but the fact that I am an outsider largely ignorant of your field does raise the question of what, exactly, I can do. My only refuge is in the genre which I have been assigned. It bestows unusual liberties, and it demands that they be taken. What is not allowed in an ordinary, civilized conference paper, then, I must exercise on this occasion—in order to deliver to you this mental form known as a *key-note*.

The term *key-note* is musical. It originates from the writings of Guido d'Arezzo (c. 991-1033), the medieval music theorist from whose principles modern Western musical notation is derived.⁴ Guido used the term *clavis* (L. "key") to mean 'the note or tone on which a scale or sequence of notes is based', hence the idea of a *key-note*. The English term *key* carries the anterior sense of "that which serves to open up, disclose, or explain what is unknown, mysterious, or obscure" (*OED* n1 6.a). A *key-note speech* is thus supposed to provide the harmonizing idea by which the various papers and discussions of a conference may constitute concordant, polyphonic expression. It opens up the secret of their unity. Its success is measured by how often subsequent discussion (which until delivery it has had no opportunity to influence) shows an answering resonance or returns to the central idea, like a musical composition to its *key-note*. Unusual risks, hence the unusual liberties.

The first of these is the freedom to make unproved and perhaps unprovable assertions as stimuli to insight and thought. I will venture immediately, then, past the practical operations of CALL (which I am

³ Johnson's remark, from the Preface to his edition of Shakespeare, is about the extreme demands put on the editor who would emend the text; see the Yale edition, 7.94-5. My thanks to Professor Peter Seary (English, Toronto) for tracking this quotation to its source.

⁴ *New Grove Dictionary of Music and Musicians* 7.803-7.

incompetent to discuss in any case) to that ideal fluency which animates it. This ideal suggests to me that CALL has the same aim as all human technologies: namely, to allow us to construct or repair ourselves according to an image of desire. The constructive role of desire in human life is of enormous importance. As Gaston Bachelard says,

Aussi haut qu'on puisse remonter, la valeur gastronomique prime la valeur alimentaire et c'est dans la joie et non pas dans la peine que l'homme a trouv   son esprit. La conqu  te du superflu donne une excitation spirituelle plus grande que la conqu  te du n  cessaire. L'homme est une cr  ation du d  sir, non pas une cr  ation du besoin.⁵

For example, Mircea Eliade notes that in several ancient societies megaliths and other burial monuments seem to have constituted a means for the soul of the dead to put on an immutable stone body and so to achieve immortality on earth (216-20). An exhortatory slogan I saw recently –"Monolingualism can be cured!"– gives us the corresponding image of desire for CALL. Since in biblical tradition sickness is commonly a metaphor of lapsarian existence, sickness for want of languages suggests as its cure a return to life before Babel, when we could all understand each other.

You may wish to object that developing CALL is hard enough without making it a religious problem and existential quest! My purpose, however, is not to complicate an already difficult task –though its true complexities must be understood– rather to connect it with poets' and philosophers' visions of language so that we can get some idea of where we might be going. As George Steiner remarks throughout his seminal book *After Babel*, it is from them we must get our linguistic star-maps and ocean-charts.

First, however, allow me to turn back from omni-lingual fulfilment more or less to our current state of affairs as Milton saw them. The question I wish to raise is in a more basic sense what computers have to do with Milton's ratio of divine to human language, and so with our present condition. The answer was suggested about 15 years ago by Roberto Busa, the Jesuit scholar who in 1946 began work on the first application of an electronic computer to language,⁶ his

⁵ "As far back in time as we can go, gastronomic value has always been more highly prized than nutritive value, and it is in joy and not in sorrow that man discovered his intellect. The conquest of the superfluous gives us greater spiritual excitement than the conquest of the necessary. Man is a creation of desire, not a creation of need." *La Psychanalyse du feu* 39, trans. Alan C. M. Ross, *The Psychoanalysis of Fire* (Boston: Beacon Press, 1964): 15-16.

⁶ Strictly speaking Busa is not the first "humanist" to be involved with computing, although his is the first computer-assisted concordance project. The honour probably goes

Index to the writings of St. Thomas Aquinas. With characteristic insight, Busa notes his providential survival through many years of work in humanities computing, under conditions few of us would survive, then finds consolation in an analogy: "Since man is child of God and technology is child of man," he muses, "I think that God regards technology the way a grandfather regards his grandchild" (87-8). Of course this is humorously intended, but its cultural resonances are worth the most serious attention, as I will attempt to suggest later. I cite Busa's charming thought, however, because *mutatis mutandis* it provides the beginnings of an answer to the question of what computers have to do with language. The version I offer is this: *that the language of God is to the language of man as the language of man is to the verbal data of computing.*

My purpose in adapting Busa's analogy is, again, to approach the paradoxical idea of a language beyond the language we know, and to press it into service for humanities computing. I am interested in the paradox of speech, that is, because it gives us the ability to formulate, however broadly, the aim of language and so the overall goal of computing it.

Milton's is only one of numerous statements in the poetics of many languages and cultures. James J. Y. Liu has shown, for example, how in the Taoist and Buddhist traditions this same paradox surfaces as an insistence on the ultimate eloquence of the nonspeech towards which all speech tends. "Endless meanings are conveyed by limited words," writes Zhu Guangqian, the leading aesthete of modern China, "therefore many meanings all lie in nonspeech. The reason why literature is beautiful does not lie merely in limited words, but even more in endless meanings."⁷ If this is so, then for those of us with an eye to the data the question is, where are these meanings situated? Hence Maurice Merleau-Ponty's question, "Mais si le langage exprime autant par ce qui est entre les mots que par les mots? Par ce qu'il ne 'dit' pas que par ce qu'il 'dit'?"⁸ Paul Valéry answers, that the aim of poetic language is exactly to reach beyond itself, into that interstitial silence, toward "l'expression de ce qui est inexprimable en fonctions finies des mots...."⁹ This same

to Wilhelm Schickard (1592-1635), whose calculator was in existence by 1623 and so antedates Pascal's, of 1642. Schickard started his academic career at Tübingen in 1619 as "professor hebraeus" and lectured widely on other subjects in the arts and sciences; see 350 *Jahre Rechenmaschinen*. I am indebted to Professor Dr. Wilhelm Ott (ZDV, Tübingen) for introducing me to Schickard's work and for the reference.

⁷ *Wuyan zhi mei* ["Beauty without words" or "The beauty of nonspeech"], trans. by Liu 91.

⁸ "But what if language expresses as much by what is between words as by the words themselves? By what it does not 'say' as by what it 'says'?" (*Signes* 56), trans. Richard C. McLeary, *Signs* (Evanston, Ill.: Northwestern University Press, 1964): 45.

of striving to ex
 limited frustration at
 T. S. Eliot

has now beginning, a c
 With steady equipment
 In the general sense of a
 Undisciplined spends of
 By strength and volume
 Once or twice, or seven
 To combat - but there i
 There is only the fight i
 And found and lost again
 That seem unpropitious
 For us, there is only the

Our business, Eliot
 (for whatever else i
 the impossibility of the ta
 ago were it not for th
 biblical tradition, to find in
 to reach, or at least
 language, we strive after the
 success than through the my
 is a question to which I an
 beyond themselves through
 refer to as their "context".

If I return to Milton
 cannot be known except i
 Valéry suggests, the opposi
 breath without a myth or
 Perhaps it is even true that o
 but here I reach my limit an
 to the fascinating suggestio
 may have created man, ra
 Genesis and the Gospel of J

⁹ "the expression of what cannot
 II 1985

¹⁰ See Viegelin and Auerbach.

¹¹ The term is Steiner's (144).

¹² From *Bild und Wort*, Walter Sal
 1966: 15-16, quoted by Steiner

paradox, of striving to express the inexpressible, also takes shape as the poet's enlightened frustration at the impossibility of his or her task. "...And so each venture", writes T. S. Eliot in "East Coker",

Is a new beginning, a raid on the inarticulate
With shabby equipment always deteriorating
In the general mess of imprecision of feeling,
Undisciplined squads of emotion. And what there is to conquer
By strength and submission, has already been discovered
Once or twice, or several times, by men whom one cannot hope
To emulate –but there is no competition–
There is only the fight to recover what has been lost
And found and lost again and again: and now under conditions
That seem unpropitious. But perhaps neither gain nor loss.
For us, there is only the trying. The rest is not our business.

Our business, Eliot may be taken to suggest, is triangulating on the infinite (or whatever else we choose to call it) with our finite minds and tools. The impossibility of the task might well have stopped the species cold a long time ago were it not for the apparently unquenchable impulse, for us rooted in biblical tradition, to find in the seeming chaos of events the traces of what we need to reach or at least dream beyond them.¹⁰ Thus in the domain of language, we strive after the eloquence of nonspeech, to which we have no other access than through the mysterious operations of finite words. Somehow –this is a question to which I will return– these finite words communicate meaning beyond themselves through interaction with each other in what we roughly refer to as their "context" (see note 22).

If, to return to Milton's vocabulary, the immediate language of God cannot be known except through our human "process of speech", then as Valéry suggests, the opposite is also true: that human speech is not worth our breath without a myth or "axiomatic fiction"¹¹ of transcendent language. Perhaps it is even true that *our speech is powerful in proportion to that fiction* – but here I reach my limit and must stop. Not without, however, pointing beyond to the fascinating suggestion of the biologist Jacques Monod, that "language may have created man, rather than man language".¹² Echoes, perhaps, of Genesis and the Gospel of John?

⁹ "the expression of what cannot be expressed in the finite function of words", "Poésie" II.1085.

¹⁰ See Voegelin and Auerbach.

¹¹ The term is Steiner's (144).

¹² *From Biology to Ethics*, Salk Institute, Occasional Papers 1 (San Diego: Salk Institute, 1969): 15-16, quoted by Steiner 133.

On the other side of the Busan analogy, the situation is different, since clearly we have direct access to human language without going through verbal data.¹³ How intimately, then, are verbal data and human language interrelated?

It is obvious, I hope, that in any given instance computing a set of verbal data depends crucially on knowledge of the particular human language in question.¹⁴ What may not be quite so obvious is the universal proposition with which I began, that without an *adequate* conception of human language as a whole, our efforts to compute it as data are doomed to the laughably inadequate. Literary scholars are painfully familiar with the limitations of software that result from simplistic notions and plain ignorance about language, but the problem is not confined to great literature. As you well know, and as Steven Pinker has shown recently, casual speech is a highly complex and subtle medium. Granted, there are some purely utilitarian interchanges into which complexities of expression do not enter, but someone confined to these is largely condemned to silence. (Think of yourself alone for a week in a city whose language you know only from a guidebook.) Hence *all* of us who compute words must understand how important it is that we keep the vanishing point of language in view, even if we work only in the immediate foreground. A myth of transcendence is thus useful for allowing us paradoxically to work with this vanishing point, and working with it we begin to acquire a more adequate idea of language on which to base our use of the computer.

This vanishing point has, however, an opposite effect too, as the Miltonic ratio suggests. If it gives us something to work toward, it also puts computing into perspective, and so reveals our analogue to the poet's "shabby equipment always deteriorating". Everyone who owns a computer understands Eliot's

¹³ In a practical sense, even now many of us do not in fact access some bodies of text directly but get at them only through operations with verbal data, e.g. when searching a large textual corpus. As large corpora become more and more important for academic study and other aspects of life, the gulf between direct and indirect knowledge will become much more of a problem and appear increasingly like the absolute divide Milton posits between divine and human language—unless, of course, access mechanisms improve.

¹⁴ Note, however, that most analytic software seems to have been written with English in mind, and that attention to character sets is not the only problem. Retrieval techniques that use proximity as a measure of relation, for example, do not work well with highly inflected languages such as Latin, in which words may be grammatically related across a considerable distance. Steiner asserts more radically that "the 'languages' of computers, the meta-linguistic codes and algorithms... are founded on a sub-text, on a linguistic 'pre-history', which is fundamentally Anglo-American.... Computers and data-banks chatter in 'dialects' of an Anglo-American mother tongue" (xvii). A tempting thought. I would be glad to know of any studies on the relation between computing and the linguistic structures of English.

phrase immediately, but I, beyond hardware or even software, to conceptualize our material equipment is indeed shabby our understanding of language.

The moral of this story is that the problem is both persistent and coming to terms with it no response is driven by two: equipment is, second, that there is only the trying.

Now there are two equipment. The first, which engineering. Without a doubt just suggested, it seems to be evolutionary cycle of improvement, is the way of life. Of course we can learn from scholarly attention must be credit of our tools is made is meant as an exhortation to

I make failure my stake make lasting sense of our works well for technology, of CALL are yearly visible of the humanities, progress relatively brief period of software development can only be rendered embarrassing will improve. Let us make authentic languages will even instructional materials with how do we come to terms taking refuge in impractical learning about language as engineering solutions? It is

¹⁵ See Steiner's remarks on the xvii.

phrase immediately, but I am speaking of the mental equipment of computing beyond hardware or even software, i.e. the ways that computing now allows us to conceptualize our materials. From the vantage point of language, this equipment is indeed shabby, and each formulation in it seems to grow worse as our understanding of language improves.

The moral of this story is that since literary and linguistic complexity is inexhaustible, and the evolution of technology apparently unstoppable, the problem is both persistent and always changing, so we had better have a way of coming to terms with it now. Again the poets guide us. As Eliot suggests, our response is driven by two imperatives: first, that we realize just how shabby our equipment is; second, that we nevertheless use it with all our strength. "For us, there is only the trying."

Now there are two ways to realize good effect from such imperfect equipment. The first, which I will call the "path of success", is the way of engineering. Without a doubt, engineering produces useful products and, as I just suggested, it seems to mean for CALL and computing in general an endless evolutionary cycle of improvements. The second, which I will call the "path of failure", is the way of science (in the etymological sense of "knowledge").¹⁵ Of course we *can* learn from success, but the edge of knowledge, where scholarly attention must be focused, is where things fail, and just where the crudity of our tools is made apparent. Hence the second part of my title, which is meant as an exhortation to regard the intellectual value of failure.

I make failure my subject because facing it is the only way I know to make lasting sense of computing in the humanities. The myth of progress works well for technology in the short-term –improvements in the machinery of CALL are yearly visible and audible– but in the long-term, the time-frame of the humanities, progress is a highly dubious notion. Even within the relatively brief period of an academic's career, genuine accomplishments in software development come at an enormous professional and institutional cost, only to be rendered embarrassingly obsolete in short order. Arguably, things will improve. Let us make the safe assumption that, for example, high-level authoring languages will eventually allow the teacher to put together or modify instructional materials without quite such a cost. My point, however, remains: how do we come to terms with the blunt crudity of our tools *now*, without taking refuge in impractical and evasive promises? What, in other words, are we learning about language and language-learning from the endless round of engineering solutions? If there is no intellectual gain, or if we do not pay

¹⁵ See Steiner's remarks on the problematic nature of any "science of language" (118-9; xv-xvi).

attention to it, I would argue that the cost of development is too high.¹⁶ Perhaps this is a commonplace of CALL research; it does not seem to be always so in other areas of the humanities.

My title is more than an exhortation to value failure, however. It specifically draws attention to the failure of *computational modeling* and so suggests a certain way of thinking about how we use computers, what they represent to us, their makers –or, as Busa would have it, their parents.

What, then, is modeling? To clarify the notion and how we might exploit it, we need to put the computer briefly into an historical perspective. Although it is a relatively recent invention, as a kind of device the computer is very old. Culturally speaking, it is only the latest in a long line of automata, or “self-moving machines”, the first of which appear for the Western European tradition in Homer (e.g. *Il.* 18.376-7).¹⁷ Throughout this history, the automaton shows a persistent tendency to anthropomorphism, especially in the form of the robot or android. Although one may question how important human shape remains for the automaton, current American popular culture, for example, provides considerable evidence that literal anthropomorphism is vital: witness the thoughtful Data in *Star Trek*, and the numerous progeny, in *Robo Cop*, *Blade Runner*, and so forth, of the sinister Maria in the German Expressionist masterpiece, *Metropolis*. Even when the outward shape is not human, however, the thought-form is deeply anthropomorphic. For me one of the most eloquent and telling images of this underlying form is the microphotograph, taken by John K. Stevens and Judy Trogadis, of a human brain cell growing on top of a Motorola 68000 CPU chip (see the figure, page X).¹⁸ The juxtaposition of cell and chip implies, of course, the narrowing functional gap between them, but more importantly, it is an image of the desire, old as Orpheus, to transform the inert world into human form. It shows us what we want, what we are trying to do.

¹⁶ Note in passing that the scientific, even philosophical aspects of computing in general and CALL in particular are essential for the results of our expensive involvement to be *practical*.

¹⁷ See the discussion and references in McCarty, “Language, Learning, and the Computer”. The history of automata makes clear that computers are very much the concern of humanists. Though we may leave it to others to build them, we should never look on them as alien nor fail in our responsibility to say how we think they should be designed and used.

¹⁸ Originally, Playfair Neuroscience Unit, now Eye Research Institute, Toronto Western Hospital and University of Toronto. Image used with permission. See Stevens; Trogadis may be contacted via e-mail at judy@camtwh.eric.on.ca.

...without Process of Spe

Let us suppose, then, incidentally, that we are showing us what we cannot see. The man-made world is what we would, or whether our reflected image is accurate or not, shows its original something like an idealized

The fact that computers are mortal faults to it. Let the trivial kind, and so we computer to construct simulation. We learn from it something. That is the learning from the failures of

Modeling is a highly allow me to simplify by physicists. Their situation studying a reality inaccessible constructing a more or less reality. This device or model experiment, or actually in reality is not, and its behavior suggest new approaches. failures are more likely to be not assumed to be true. knowingly employed despite known as “tinkertoy” model.

Imported into human recommends itself because fabricate and perfect many particular helps us come reality and simulation, that

¹⁹ See McCarty, “The Shape of

²⁰ See the brief survey in McCarty, *Metropolis* 285-6.

²¹ I am not arguing that my accepted among physicists, to physics is the field in which to

Let us suppose, then, that like all automata the computer is essentially, not accidentally, *catoptric* (f. Gk. κατοπτρον, *mirror*), i.e. that it mirrors its maker, showing us what we cannot otherwise see, an image of ourselves reflected from the inanimate world.¹⁹ Whether in fact the computer goes about its tasks as we would, or whether our brains work like a computer –whether, that is, the reflected image is accurate, or like those on the walls of Plato's cave (*Republic* 514ff), shows its original only in a shadowy way– the intent is clear: to make something like an idealized form of ourselves.

The fact that computers “die”, hard disks “crash”, and “link rot” threatens the World Wide Web ironically suggests that we tend to assign our own mortal faults to it. Let us assume, however, a technology without failures of the trivial kind, and so without these lineaments of mortality. If we use the computer to construct simulacra of ourselves, e.g. in order to teach languages, what do we learn from its inherent failures due to the crudity I have been pointing to? That is the question implied by my title: how can we go about learning from the failures of computational modeling?

Modeling is a highly complex topic that I cannot even survey here,²⁰ but allow me to simplify by focusing on the particular kind prevalent among physicists.²¹ Their situation is not dissimilar to our own. Faced with the task of studying a reality inaccessible by direct means, they adopt the technique of constructing a more or less crude device that embodies some theory about the reality. This device or *model* may exist only in the mind, as part of a thought-experiment, or actually in the lab, but in either case it is manipulable, as the reality is not, and its behaviour can be observed, then used to refine the research or suggest new approaches. Typically model-building is recursive, as initial failures are more likely to be due to omissions and errors of design. The model is not assumed to be true, rather it is used as an expedient heuristic. Those knowingly employed despite their crudity are, in a charming Americanism, known as “tinkertoy” models (from the popular wooden forerunner of legos).

Imported into humanities computing, the notion of modeling recommends itself because it draws attention to our use of the computer to fabricate and perfect manipulable forms of our ideas. Tinkertoy modeling in particular helps us come to terms with the inevitable discrepancy between reality and simulation, that is, with the shabbiness of our intellectual equipment,

¹⁹ See McCarty, “The Shape of the Mirror”.

²⁰ See the brief survey in McCarty, “Encoding Persons and Places” 278-80. See also Saumjan 285-6.

²¹ I am not arguing that my rough model of modeling is universally or even widely accepted among physicists, for there appears to be no standard meta-model, only that physics is the field in which to look for a notion suitable to humanities computing.

at the same time that it extends our use of it. The profoundly interesting question it raises, and to my mind the point of the entire exercise, is again occasioned by failure. Where and how does the model fail? Investigating that question brings a lens to bear on the fuzzy and moving boundary where mechanical precision leaves off and imaginative precision begins.

Let me illustrate the process of recursive tinkertoy modeling in the humanities with an example from my own research in Latin literature. The research problem is to define in some useful way what is meant by the term *vocabulary* when applied to a subject, idea, motif, or theme, such as the "erotic vocabulary" of the Roman poet Ovid, or the "military vocabulary" in *Don Quixote*. The meaning is of course quite clear, until one wishes to compute with a given vocabulary, since the computer requires a list of word-forms, or what I call a "finite vocabulary". In kinds of writing that depend on a strictly defined set of words, such as we find in theology and to a lesser extent in philosophy, a finite vocabulary is easily accommodated. Wherever the imagination is allowed to play, however—and this includes ordinary conversation—the paradox of meaning beyond language resurfaces and subverts the attempt to make the vocabulary finite. The problem is easily observed, for example, with the erotic, which is probably the most elusive vocabulary of all. As J. N. Adams observes, "Almost any object or practice can acquire a sexual symbolism in a suggestive context" (vii, my emphasis). Thus the problem of meaning escapes from words into something constructed from and evoked by words—into the verbal interstices, as Merleau-Ponty suggests. What do we mean by context?²² How does it work—and how can it be computed?

Although ultimately doomed to failure, the attempt to build a finite vocabulary from imaginative language is highly instructive, as we students of failure would expect. It begins with isolated word-forms, in a recursive process of adding these one at a time to a list, returning to the text with retrieval software to test the list, discovering and adding additional terms, and so forth. Eventually, one reaches the point at which it becomes impossible to add more terms without special pleading. Then, when the finite vocabulary fails to locate passages of evident importance,²³ the attention moves beyond isolated word-forms to context.

²² Helmut Schnelle defines *context* as "a comprehensive term for internal configurations conditioning at each time the function of the language behavior processing of a linguaton [language user]" and provides the outline of a detailed analysis (332-4). I make a primitive attempt to sketch the problem of computing a context in "Encoding Persons and Places" 270-1.

²³ The discipline of interpretation under these circumstances reveals much about the limitations and powers of computational thinking. In "Encoding Persons and Places", I argue for the translation model as a means of understanding this discipline (268-76).

A typical strategy for follows. The first step is to but not limited to gra automatically.²⁴ Many a inflection, word-order, sy (KWIC) concordance, so following provide a surp The second, closely relate non-contiguous words, wh is to retrieve according distribution.²⁵ Beyond knowledge, as it is call because we do not yet kno it and in most areas we d form with which to expe judgement is required, for sense erotic or is simply arbitrary, since meaningful loss, but again that loss or l

Thus my point: at ea meaning leads to a new n spiral. To take the path of mechanisms that result. To mechanized knowledge e imagination, in no way diu trying."

Other areas of huma tinkertoy modeling, for e cannot catch?²⁶—or, I hazar of knowledge because we horizon cautions us agai progress. The most importa language, the important bi

²⁴ The software to find repeate part of the TACT textual-analy <http://www.ech.epas.utoronto> published by the Modern Lang see Lancashire and Wooldridge

²⁵ Also by means of C. J. Ger.

²⁶ See Potter for a summary of

²⁷ See McCarty, "Encoding Per

A typical strategy for defining and exploring context mechanically is as follows. The first step is to add repeating sequences of word-forms (including but not limited to grammatical phrases), which can be discovered automatically.²⁴ Many approximate repetitions escape because of differing inflection, word-order, synonymy, and the like, but a keyword-in-context (KWIC) concordance sorted by words preceding and another by words following provide a surprisingly powerful means of finding them manually. The second, closely related step is to add frequently repeated collocations of non-contiguous words, which may also be discovered automatically.²⁵ A third is to retrieve according to various kinds of more sophisticated statistical distribution.²⁶ Beyond that lies the frontier of research. "Real-world knowledge", as it is called, intrudes to an increasing degree, uncertainly because we do not yet know, at least in literary studies, how to define and apply it, and in most areas we do not yet have sufficient data accessible in electronic form with which to experiment. Typically at each stage of each process judgement is required, for example to decide whether a word "is" in some sense erotic or is simply made so by its context. Decisions must often be arbitrary, since meaningful ambiguity cannot be successfully resolved without loss, but again that loss or failure is the meat of the exercise.

Thus my point: at each stage, a failure of the tinkertoy model of textual meaning leads to a new model, which in turn fails, and so on in a heuristic spiral. To take the path of engineering is to note the slowly improving retrieval mechanisms that result. To take the path of science is to watch the horizon of mechanized knowledge extend while the vitalizing *terra incognita* of the imagination, in no way diminished, recedes before us. "For us there is only the trying."

Other areas of humanities computing exhibit the same behaviour under tinkertoy modeling, for example, encoding texts for phenomena algorithms cannot catch²⁷ –or, I hazard a guess, CALL. We can speak about the progress of knowledge because we continually learn more, but the irony of a receding horizon cautions us against thinking uncritically in terms of discovery or progress. The most important discovery is in any case that when one computes language, the important bits are not just those that do not compute but those

²⁴ The software to find repeated sequences is *CollGen* ("Collocation Generator"), which is part of the *TACT* textual-analysis system. *TACT* is currently available online, at the URL <http://www.cch.epas.utoronto.ca:8080/cch/tact.html>; the manual is scheduled to be published by the Modern Language Association of America. For applications of *CollGen*, see Lancashire and Wooldridge.

²⁵ Also by means of *CollGen*. See the previous note.

²⁶ See Potter for a summary of work to 1990, and esp. the work of Burrows.

²⁷ See McCarty, "Encoding Persons and Places".

that escape the data altogether into the interstitial silence of nonspeech. We have been hearing, or almost hearing, this silence for a very long time, of course. Arguably, however, our verbal machine makes two differences: it shows us words as data, then gives us an exacting means of examining their fuzzy edges. Ironically, the most important contribution of the exacting computer is not, as so many have claimed, to make our work more "scientific" (in the N. American sense) but to renew and perhaps deepen our awareness of our own imaginative powers. How is it, we are forced to wonder, do we get poetry, or the undervalued subtleties of common speech, out of mere words? The computing humanist's approach to this ancient question is to transcribe these words as character-strings, then by modeling human cognitive processes however crudely with the computer, to hold up a mirror to our mental life. To paraphrase the title of Vannevar Bush's seminal article published at the beginning of the computer-revolution, the machine thus shows us *how we may think*.

I have suggested here that the cycle of software development is ultimately not about the products we make, but that it is primarily an heuristic activity whose purpose is improved self-understanding. I know, such a statement may sound like the love-sick murmurings of Narcissus to himself, and so a poor if not outrageous key-note for such a practically-oriented gathering as this. Allow me to remind you, however, philosophically that the difference between Narcissus and Socrates is precisely a matter of self-awareness; psychologically that, as Freud said more wisely than many since, "dieser Narzißmus der allgemeine und ursprüngliche Zustand ist, aus welchem sich erst später die Objektliebe herausbildete, ohne daß darum der Narzißmus zu verschwinden brauchte" (431).²⁸ In other words, I derive my key-note from what appears to be a fundamental human characteristic. Thus, if the end we put in sight is to answer Bush's question of how we may think, we will be on common ground with our students, who are just as interested in the answer as we are, although they may be somewhat less focused in their pursuit of it. Meanwhile, of course, the *sine qua non* of the quest is the practical activity of making good software, so like Thales we constantly demonstrate that we know what's what.

Hence I would argue that the practical success of the CALL enterprise will be based only partially on the success of our language engineering as such. Computing systems are now beginning to reach a level of complexity at which, like other human artefacts, they are able to show traces of the personality, corporate or individual, that made them. I suspect that CALL will succeed in drawing our students toward and into the promised land of multilingualism as

²⁸ "Narcissism is man's universal original condition, out of which object-love develops later without thereby necessarily effecting a disappearance of narcissism". The English translation is by Joan Riviere, in *A General Introduction to Psychoanalysis* (New York: Liveright Publishing Corp., 1935): 360.

²⁹ Said of Clerk of Oxenford in
to Gary Shawver (Medieval Stu

much or more because of the persona in the software as its technical quality. As it is with books. "For books are not absolutely dead things", wrote Milton in the *Areopagitica*, "but do contain a potency of life within them to be as active as that soul whose progeny they are." The final question is, again, about our own creative potential and what we do with it.

I am not qualified to discuss creative potential. My concern here is rather to remove some impediments to its realization by clarifying what the computer is as a cultural object and how it stands in relation to us, in particular us humanists. Roberto Busa's analogy, though of course not true, usefully implants the historical truth that the computer belongs to us and is our professional and moral responsibility to develop. For the computing of language, fulfilling this responsibility begins with a deep transcendental *philology*, by which I mean a love of nonspeech manifested through passion for language, and is fostered by a combination of stubborn persistence and the playful curiosity of liberated desire. In other words, the qualities of a good teacher are the qualities we attempt to embody in the teaching machine. There is already evidence that CALL can give the good teacher – a rare individual at the best of times – the ability to propagate his or her talents through software. My key-note, however, is as Busa said, that "the use of computers in the humanities has as its principal aim the enhancement of the quality, depth and extension of [our understanding] and not merely the lessening of human effort and time" (89), the reducing of budgets, or any other such efficiency measure. My key-note, echoing Chaucer, is that we are here "gladly [to] lerne and gladly teche"²⁹ the inexhaustible, ultimately ungraspable wonders of nonspeech. With the computer!

²⁹ Said of Clerk of Oxenford in the *Canterbury Tales*, General Prologue 308. I am grateful to Gary Shawver (Medieval Studies, Toronto) for finding this echo for me.

WORKS CITED

- Adams, J. N. *The Latin Sexual Vocabulary*. Baltimore: The Johns Hopkins University Press, 1982.
- Auerbach, Erich. "Figura". Trans. Ralph Mannheim. In *Scenes from the Drama of European Literature*. Theory and History of Literature, vol. 9. 1959; rpt. Manchester: Manchester University Press, 1984. Pp. 11-76.
- Burrows, J. F. *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford: Oxford University Press, 1987.
- Busa, Roberto. "The Annals of Humanities Computing: The Index Thomisticus". *Computers and the Humanities* 14 (1980): 83-90.
- Bush, Vannevar. "As We May Think", *Atlantic Monthly*, July 1945; an online version is at the time of writing available at the following address: <http://www.csi.uottawa.ca/~dduchier/misc/vbush/as-we-may-think.html>.
- Eliade, Mircea. *Patterns in Comparative Religion*. Trans. Rosemary Sheed. London: Sheed and Ward, 1958.
- Eliot, T. S. *Four Quartets*. New York: Harcourt, Brace and Company, 1943.
- Freud, Sigmund. "Die Libidotheorie und der Narzißmus". In *Vorlesungen zur Einführung in die Psychoanalyse*. Vol. 11 of *Gesammelte Werke*. London: Imago, 1940.
- Johnson, Samuel. *The Plays of William Shakespeare*. In *Johnson on Shakespeare*. The Yale Edition of the Works of Samuel Johnson. Vols. 7 and 8. Ed. Arthur Sherbo with an Introduction by Bertrand H. Bronson. New Haven: Yale Univ. Press, 1968.
- Kafka, Franz. "Josefine, die Sängerin oder Das Volk der Mäuse". *Das Franz Kafka Buch*. Comp. Knut Beck. Frankfurt am Main: S. Fischer, 1983.
- Lancashire, Ian. "Uttering and editing: Computational text analysis and cognitive studies in authorship". *Texte et informatique. Texte* 13/14 (1993): 173-218.
- Linguaggi nella società e nella tecnica*. Convegno promosso dalla Ing. C. Olivetti & C., S. p. A. per il centenario della nascita di Camillo Olivetti. Milano: Edizioni di Comunità, 1970.

... Process of Sp...

James J. Y. *Language*... Princeton Unive...

McCarty, Willard. "Enco... Part 2: The Metal... 15.10.1994: 261-3

"Language, Learning... Investigations". In *CAI... CCALL2 CCELA02: The... Little Victoria BC Lang...*

"The Shape of the Mir... *Archiv* 22 (1989): 161-8

Merleau-Ponty, Maurice. *S...*

Milton, John. *Paradise Lo...* Y. Hughes. New York: Od...

New Grove Dictionary of... Macmillan, 1980.

Pinker, Steven. *The Language...* New York: Harper, 19...

Potter, Rosanne. "Statist... *Computers and the Human...* 25 (1991).

Saumjan, Sebastian Konst... simulating natural languag...

Schnelle, Helmut. "Lingu...

Steiner, George. *After Bal...* Oxford: Oxford University...

Stevens, John K. "Rever... (1985): 287-99.

Valéry, Paul. *Cahiers*. Ed. Gallimard, 1974.

Liu, James J. Y. *Language, Paradox, Poetics*. Ed. Richard John Lynn. Princeton: Princeton University Press, 1988.

McCarty, Willard. "Encoding Persons and Places in the *Metamorphoses* of Ovid. Part 2: The Metatextual Translation". *Texte, métatext, métalangage. Texte* 15/16 (1994): 261-305.

—. "Language, Learning, and the Computer: desultory postprandial investigations". In *CALL: Theory and Application. Proceedings of CCALL2/CCELAO2, The Second Canadian CALL Conference*. Ed. Peter Liddell. Victoria, BC: Language Centre, University of Victoria, 1993: 37-55.

— "The Shape of the Mirror: Metaphorical Catoptrics in Classical Literature". *Arethusa* 22 (1989): 161-95.

Merleau-Ponty, Maurice. *Signes*. Paris: Gallimard, 1960.

Milton, John. *Paradise Lost*. In *Complete Poetry and Major Prose*. Ed. Merritt Y. Hughes. New York: Odyssey Press, 1957.

New Grove Dictionary of Music and Musicians. Ed. Stanley Sadie. London: Macmillan, 1980.

Pinker, Steven. *The Language Instinct: How the Mind Creates Language*. 1994; rpt. New York: Harper, 1995.

Potter, Rosanne. "Statistical Analysis of Literature: A Retrospective on *Computers and the Humanities*, 1966-1990". *Computers and the Humanities* 25 (1991).

Saumjan, Sebastian Konstantinovic. "Linguistic models as artificial languages simulating natural languages". In *Linguaggi* 285-302.

Schnelle, Helmut. "Linguistics and automata-theory". In *Linguaggi* 325-40.

Steiner, George. *After Babel: Aspects of Language and Translation*. 2nd edn. Oxford: Oxford University Press, 1992.

Stevens, John K. "Reverse Engineering the Brain". *Byte Magazine* (April 1985): 287-99.

Valéry, Paul. *Cahiers*. Ed. Judith Robinson. Bibliothèque de la Pléiade. Paris: Gallimard, 1974.

Voegelin, Eric. *Israel and Revelation*. Vol. 1 of *Order and History*. Baton Rouge: Louisiana State University Press, 1956.

Wooldridge, T. Russon, ed. *A TACT Exemplar*. CCH Working Papers 1. Toronto: Centre for Computing in the Humanities, 1991.

350 Jahre Rechenmaschinen. Vorträge eines Festkolloquiums veranstaltet vom Zentrum für Datenverarbeitung der Universität Tübingen. Ed. Martin Graef. München: Carl Hanser Verlag 1973.