Localization and the humanization of technical discourse.
Revising the Suppositions

Written for the Third International FEDER.CEN.TR.I. Conference, Rimini, Italy, October 2002

Intercultural Studies Group
Universitat Rovira i Virgili
Tarragona, Spain

Abstract: Discursive analysis of two fragments of IT discourse reveals features that allow one to be classified as "professional" and the other as relatively "non-professional". The question is then raised as to whether the movement from one discourse to the other might be characterized as "humanizing", or at least in tune with some concept of humanism. Discussion of this possibility in terms of Renaissance humanism, Bernstein's sociolinguistics and Buber's metaphysics concludes that, in many respects, the apparent humanization of technical discourse is in fact segmenting social readerships, separating production from end-consumption. Such segmentation, and the mediation that supports it, is seen as a general problem confronting the reigning ideologies of localization.

T he last thing localizers want to hear, especially in a climate of recession, is that they are being inhuman. So I won't tell you that; I will try not to be the typical academic, using the authority of an obscure past to preach irrelevance a busy present. Let's say, instead, that our language, our technical discourses, can be humanizing or dehumanizing. Or better: there is often something missing in technical communication, something we sense the lack of, something we would like to recapture, for our own lives if not for the sake of efficiency, something that is somehow human or humanizing. And only an academic, comfortably ensconced away from harm and finance, would have the time and trouble to ask what that something is. Bear with me.

Part of this is what everyone already knows. It is embedded in the core ideology of localization, surely dedicated to giving end-users the language that is most theirs, adapted to their social and cultural conventions, entering the intimacy of everyday lives. To localize should surely be to humanize, in some nebulous sense. On the other hand, to internationalize, to produce the generic codes proper only to the locale of production, could be to dehumanize, in an equally undefined sense. The computer screen full of source code, although very much part of a communication act, is somehow not what we would want to call humanizing discourse. It might be closer to God, nearer the upper reaches of our postmodern Towers of Babel, accessible only to those able to climb, but its fleeting universality is not quite what the customer is supposed to want. This simple observation might suggest a first idea of the difference: to humanize a discourse is to put a face on it, to bring it closer to the communication acts that belong to many overlapping domains, beneath the purely technical.

Humanization should thus involve the selection of quite local language varieties, the use of short sentences and decidedly non-technical terms, the eschewing of acronyms, the adaptation of terms of address, examples, place names, colours, date conventions, currencies, and so on, right through to cuteness conventions like that annoying little dog on Microsoft XP. Humanization might also work according to hierarchies of expertise, recognizing that within the one linguistic-cultural locale there

are degrees of professional competencies and needs. The marketing of XP in Professional and Home editions would thus ideally bring communication closer to the specific user. And this, ideally, would be humanizing. To localize, as we have said, should be to humanize. And that is precisely what is being done out there, perhaps.

So why talk about human qualities at all? Because, hopefully, they enable us to take the questioning further. Perhaps they will even enable me to move perversely from a very restricted example (found in the recondite crannies of my PC) to the ideologies of Renaissance humanism, to a dated Marxist sociolinguistics, and to the metaphysics of Martin Buber. Humanism might yet give you your money's worth of ideas (assuming this paper is free).

For example

Here are fragments of two texts, both about passwords (the first is from the manual for Dreamweaver; the second is from Microsoft Word). One text would seem more localized than the other:

password NN 2 IE 3 DOM 1
The password object is a form control generated with an INPUT element whose TYPE attribute is set to "password". This object is similar to the text object, except that the characters typed into the text box by the user are converted to asterisk or bullet symbols for privacy.

Word provides several security and protection features. You can do any of the following:
· "Protect" a document to restrict the types of changes users can make to it. (For example, protect an online form so users can fill in only the designated areas.) For extra security, you can assign a password to prevent unauthorized users from "unprotecting" the document. For more information, click ÿ .

The first text is ostensibly for a relative professional; the second is more for a non-professional end-user. We might assume that the logical progression is for concepts to travel from the first type of discourse to the second, since the language of intercultural production precedes the result of localization. Then again, the more technical text is almost certainly the result of a lost oral tradition, the kicking around of ideas, the strange pacts of informality and technicality that mark the intercultures of this field. Here we can only leave those wondrous developmental dynamics to the speakers involved (suspecting, nevertheless, that some process of dehumanization led to the first of the above texts). Let us study no more than the assumed movement from the professional to the non-professional. How might we recognize this as humanization?

Readability?

The linguistic differences between the above discourses are fairly obvious. It takes less than a minute to get numbers on a few of the features involved (here for the whole texts, not just the above fragments):

	Professional	Non-professional
Words per sentence	17.9	9.9
Characters per word	4.7	4.5
Passives	36%	0%
Flesch readability score	53.7	67.4
Flesch-Kincaid score	10.2	6.3

On the basis of this analysis, we might claim that humanization involves avoiding complex syntax and using short words, which are the most common. So humanization might be no more than writing a text that is easy to read. Here the Flesch-Kincaid score indicates that the non-professional text can be read by a sixth-grader (in the American school system); the professional one requires a tenth-grader. However, if that were all, what would keep humanization from being no more than the dumbing-down of discourse? Write as if you were speaking to a child (a sixth-grader, aged eleven or so), and the result will necessarily be human? And if you write in the most efficient way for a technician, as might have been done in the first of the texts above, is the result necessarily dehumanized? There must be something wrong here.

Like most attempts at corpus studies in the field of translation, the above approach errs by using the available tools, getting the numbers, and forgetting the kind of thought that might make the results meaningful.

The sociology of electronic communication, by whatever yardstick, suggests that discursive competence in this particular field tends to be *inversely* proportional to age. Beyond early adolescence, the older you are, the more difficult it tends to be for you to understand this kind of discourse. Yet the above analysis, based on tools for all kinds of discourse comprehension, assumes that the opposite is true, that age will bring complexity, if not insight. No, a youth spent with technical discourse might bring its own modes of comprehension, and the corresponding discursive features have no reason to be located on any kind of general age-graded scale. Our search for humanization should surely look elsewhere.

We might, for example, consider where the term itself has come from.

Rhetoric and social segmentation?

The term "humanism" has meant an enormous number of things to almost as many people (see, for example, Bödeker 1982). It comes to us from the Renaissance, from the Italian cities of the fifteenth century, where the *studia humanitatis* were associated with values like respecting the linguistic integrity of texts, of placing reason above religion. This often meant constructing civil society through secular education and by creating networks between professional intellectuals. In linguistic terms, that kind of humanism placed a high value on the manner of expression, over and above the content of what is expressed. Such were the calls for "eloquence", and the corresponding development of studies of rhetoric. It is very hard for us now to appreciate the connection between "eloquence" and "human values", but such things were certainly of a piece in an age just escaping from Christian fundamentalism. Pre-humanists read and translated Plato in terms of how well his pagan ideas fitted in with established doctrine; the humanist Leonardo Bruni read and translated him as a person using stylistic devices, conveying a personality as well as ideas. That might be what is meant by a "humanizing" discourse: Bruni did indeed use philology to put a human face on his texts, including his translations.

Let's go back to our two text fragments. Admittedly they seem a long way from anything like Renaissance eloquence. And yet, perhaps surprisingly, our very mode of analysis can now be said to be humanist. We have instinctively started to evaluate these fragments in terms of their forms of *expression*, quite independently of any consideration of content. Further, the consistency with which the less professional text has been constructed would suggest that considerable intellectual effort has been

invested in this kind of rhetorical analysis. As indeed would our attention to localization of all kinds: vast resources are being invested in the manner of expression, over and above the development of content. We are still within at least that part of the Renaissance, for better or for worse. This is not quite a matter of addressing everyone as if they were children.

One of the main debates associated with humanism is the extent to which eloquence isolates readerships. For some (including, I believe, Leonardo Bruni), the full panoply of a language's expressive resources is ideally to be used to move people, to address a wide range of social groups, and to form the common bonds of civil society. Eloquence is in the mouth of the great politician (which Bruni was); it is able to convince many different kinds of people to cooperate with each other; it is most powerful when in a discourse of the general, open to technicalities from all sides, cross-fertilizing fields of knowledge in the form of metaphor. For others, however, eloquence can equally involve an excessive preoccupation with correctness, with complexity, and with abstruse language resources, all of which may lead to quite the opposite result. For Bödeker, "the *studia humanitatis* presuppose a high degree of education with elitist features, which thus compromise their wider effectiveness" (1982: 1070). You set out to move a lot of people, you develop techniques for doing so, but you soon finish up with so many technicalities or old words that you fail to get beyond a narrow professional coterie. In the age of globalized personal computing, technology is supposed to be available to everyone, yet vast amounts are locked up in discourses only accessible to the production sectors. The attempt to democratize might thus produce its own negation. You are either privy to the obscurities of production, or you are located as a child-like end-user, fit only to consume. Let us call this "social segmentation". It is a common enough problem.

In our examples, a trace of such segmentation might be found in the inverted commas. The non-professional text marks the metaphorical status of the terms "protect" and "unprotect", thus implicitly justifying the technical use of the second term as a verb (it does not appear as such in most general dictionaries, including those that come with Microsoft Word). The barrier operated on here thus stands between the technical and everyday use of language. The text is written in such a way that the reader may cross that barrier, effectively learning the function of the technical verb. You may be positioned as a child, but you *can* learn to advance, step by little step. In the professional text, on the other hand, the inverted commas mark off the one piece of language that does *not* assume any technical definition. Indeed, "password" is the only term that could be any other term, since it is an arbitrary setting. It is thus marked off from a text where every other element is implicitly defined by the concepts comprising professional competence.

The professional text might thus be said to protect the technical domain it assumes: each unexplained term is indeed a shibboleth. The non-professional text, suitably idealized, could present barriers only in order to teach across them, effectively unprotecting its origins. The latter quality would appear to be humanizing, in the most positive democratic sense of the Renaissance.

Then again, wait a minute: the verb "to unprotect" is hardly a major advance. Further, considered in functional terms, the professional text in this case is more consistently pedagogical. In fact, it is a running interlinear translation between HTML and natural language:

The password object is a form control generated with an INPUT element whose TYPE attribute is set to "password". This object is similar to the text object, except that the characters typed into the text box by the user are converted to asterisk or bullet symbols for privacy.

HTML Equivalent
<INPUT TYPE="password">

And so on. The non-professional text might teach the user a new verb; but the professional one is teaching a whole language. There is thus little reason to characterize one as more pedagogical than the other. And there are quite good sociological reasons to see the professional text as far more empowering than the timid exposure to a slightly new verb. One kind of discourse trains its reader to remain passive; the other gives the tools for the creation of texts. The true password is conveyed in the professional text.

Social segmentation is certainly what is at stake here. There is an extreme discursive gap between the two texts (I struggle to find discourses mixing these features, but perhaps I have not looked everywhere). This means that the receiver tends to be positioned in one locale or the other, either an end-user or a potential producer. Further, from this perspective, there is nothing essentially wrong with the production of texts that are difficult to read (difficult for whom?), as long as the effort invested in reception is proportional to the value of the knowledge conveyed. An easy text that conveys banalities is no better or worse than a difficult text giving the keys to power. Of course, the best text, by far the most humanizing, would be the one that makes the most valuable content available to widest cross-section of people. In the relative absence of such discourses, in the strangely binary division between reception and production positions, the Renaissance paradox of social segmentation remains part and parcel of our most humanizing efforts.

Three lessons here. First, beware "humanist" modes of analysis that would seek to evaluate means of expression independently of content. Second, there is no *a priori* reason why the discourse of Plain English should, through localization, force its neat Puritan criteria onto all other languages. Third, the more serious problem is the gap between discourses, not simplification for the sake of simplification. This gap could set up a second technological divide, no longer between the haves and the have-nots of material machines, but between the active and passive users of language. In effect, the more condescending forms of localization would divide the world into text producers who will always be producers, text consumers who can only remain consumers, and the excluded, who remain unlocalized.

Explicitness

One of the most obvious linguistic differences between our two discourses is the relative explicitness of the professional text. The text uses no proforms and has no marked instances of syntactic deletion. The non-professional text, on the other hand, uses the pronouns "you" and "it", contracts to the imperative ("click"), and allows quite colloquial deletion such as the connector "that" ("the types of changes [that] users can make to it"). This tendency toward greater explicitness is also present in the use of passives in the professional text (there are none in the non-professional text), since their purpose is not agent-deletion (the agent is in fact named extensively in "by the user") but rather syntactic expansion in the name clear attribution. A further

consequence of this explicitness is the transgression of stylistic norms outlawing repeated terms ("object... object... object", for example). All in all, the professional text is explicit, whereas the non-professional text uses many of the implicitness strategies of everyday or spoken discourse.

If the human face of discourse smiles in the use of everyday forms, then the non-professional text must surely be more humanizing, on this count at least?

Perhaps the sociolinguistic categories most readily applicable here are those formulated by Basil Bernstein in the 1960s (generally known through Bernstein 1971). Everyday language is, for Bernstein, to be seen in terms of "restricted codes", full of elision and implicit information. Such codes are used in situations that are often repeated, where the roles of the participants are highly determined by contextual factors. This would archetypically be the discourse of the family. "Elaborate" codes, on the other hand, would make relations explicit, expanding syntax. This would certainly be the language of science, the professional discourse we have seen above, but also the language of Western schooling. Bernstein used these concepts as a way of explaining why working-class children under-perform at school: they struggle to make the abrupt jump from restricted to elaborate codes. They are thus kept from the language of social power. Just as some users fail to jump to the discourses of computers.

Bernstein's categories were severely criticized in the United States, notably by Labov, because they were seen as partaking of a general "deficit linguistics" (see, for example, Labov 1972). The notion of restricted codes could be used to claim that American Black Vernacular, for example, was an inferior language and should not be used in schools. Labov thus undertook to demonstrate the narrative and creative virtues of that particular variety, adding his voice to the various campaigns to have it used as a bridging language in the early years of schooling. Those debates, and the ensuing "ebonics", were strangely reminiscent of the medieval concerns with the ability of the vernaculars to express the eloquence of Greek and Latin, until civic participation finally overcame the ideological hierarchy of sacred languages. But our interests now lie elsewhere.

Both Bernstein and Labov would agree that schooling requires transitions from one kind of language to another. The difference, for them, lies more in the values to be attached to the codes or varieties concerned. For Bernstein, restricted codes firstly restrict social mobility, locking speakers within the roles of class. Elaborate codes, on the other hand, are both the discourse of power and open to a wide range of subject positions, potentially allowing many parts of society to interact (as indeed was the Renaissance ideal). Labov simply refused that distribution of values, highlighting the more affective values of stories, playfulness, and belonging. Yet for us, in view of the problem of social segmentation, Bernstein's critique should retain its pertinence.

Why insist on "explicitness" (or "elaborate codes") as a feature of technical discourse? Why is it so much in evidence in our examples, and virtually anywhere else in computers and on the Internet? Part of the answer must surely lie in the way these texts are used. No one would dream of reading a Help file in a linear fashion, from beginning to end. No one reads miles of WWW on the computer screen, from top to bottom. In both cases, we look for information using a Find function or a search engine of some kind. All we want is the sentence or paragraph containing the answer to our question. And once we get to that little text segment, it must make sense in its own right, without discursive deictics to what came immediately before and what comes immediately afterwards (the only deictics now are the hyperlinks). Hence the

need for those small units to be as explicit as possible; hence the eschewing of implicitness on all levels.

Explicitness is thus not particularly dehumanizing in itself. Electronic means of communication simply make our technical discourse non-linear, non-narrative, and as open as possible to different reading positions. Nor is this particularly new. The technological advance from the papyrus scroll to the book was already a major step toward non-linearity, since it enabled browsing and indexing (O'Donnell 1998). Electronic media just extend that tendency. And Renaissance humanism, let it be remembered, was very much a culture of the book. Our critique of dehumanization should perhaps not insist too much that people write like they speak.

Persons

I borrow from Arnaud Laygues (2001) the following passage from Martin Bubner's *Ich und Du*, published in 1923 and translated here into unfortunately sexist English (2000):

> The attitude of the man is twofold, in accordance with the twofold nature of the primary words which he speaks.
> The primary words are not isolated words, but combined words.
> The one primary word is the combination *I-Thou*.
> The other primary word is the combination *I-It*; wherein, without a change in the primary word, one of the words *He* or *She* can replace *It*.
> Hence the *I* of man is also twofold.
> For the *I* of the primary word *I-Thou* is a different *I* from that of the primary word *I-It*. (2000: 19)

When I first saw this passage, tenuously related to the training of translators (in Laygues), my editorial instinct was to delete it as irrelevant and excessively abstruse. I now force it on any student able to handle it. Buber here is talking about interpersonal relations, clearly about the way subjectivity is positioned by those relationships. The theologian's primary frame is the act of prayer, where the communication positions the *Thou*, this intimate second person, as God. We are not the same person when (or if) we pray (to a *Thou*) as when we work on a computer (with an *It*), or so posits Buber. Nor are these two kinds of relation are ethically equal. The only stable relationship is the content of the primary word *I-Thou*. The relation with the third person, with things, is a constant search for content, as when one definition requires another, one possession needs the protection of another, in the process otherwise known as semiosis.

We may choose to ignore the mysticism of dialogue with the divine. Yet many would sense there is a substantial difference between communicating with an intimate "you" (a close friend, a spouse, a parent, a sibling) and communicating with a "he" or a "she", people relegated to the status of third-person things. This may be no more than the Kantian ethics of treating people as ends rather than means, as second persons rather than third persons. Yet here is it framed in terms of communication, and specifically of pronouns. I give this piece of Buber to my students in the hope that they will seek out the *human relations* behind the texts, the person behind the client, the interpersonal behind the objective. And if that kind of search can be made, I would hope that the things of this world, the countless tasks of *I-It* relations, can be seen as exchanges between people able to help each other. Such might be the ethics of

cooperation as mutual benefit, which we leave for another day. One might more modestly see the exercise as a plea for human context.

Consider the pronouns in our examples. Not surprisingly, the more professional text is entirely in the third person, as a world of relations between things. The non-professional text, on the other hand, consistently uses the second person "you"; it allows a discursive position for the user who should be able to carry out the actions described. To that extent, we might be tempted to consider the second text to be more humanized, if and when we can consider the use of the second person a truly humanizing trait.

Such discursive positioning of the second person is nevertheless not unambiguous. It is often not followed in translation into Spanish, for example, where devises such as the reflexive are felt to be less intrusive, less "pushy". Further, when the second person is discursively necessary, as in the many imperatives, it is more often than not the *formal* second person in Spanish (*usted*). Too many "tú"s would be felt to be condescending, too much like talking to a child.

Even in English, surely, this discursive relation is rather less than satisfying. For a start, Buber was calling for an *I-Thou* bond, and here there is no explicit *first* person anywhere. The person speaking is the program, the machine, the company, the legally responsible party, a thing, never a person. Do we willingly allow objects to command our intimate subjectivity so abruptly? This is essentially the problem of profiling, as when a search engine thinks it is has picked up my preferred language because of the country I live in (wrong!), or Yahoo! puts me in the UK because of some false detail I might have revealed somewhere (very wrong!), or when, even more mysteriously, the same automatized profiling places me in Australia-NZ (correct?), when I thought that piece of identity had been bleached away long ago. Even more frustrating is the hierarchical profiling that positions the second person in the limited categories of Home and Professional, as if real lives were lived in such places.

What is upsetting here? Great efforts have been made to have machines adjust communication to the profile of the user. This, in discursive terms, is perhaps the closest that technical communication can come to the intimate *I-Thou* relationship, as if our soul had been seen by God. It should be extremely humanizing: the machine of the primary word *Machine-Thou* is a different machine from that of the primary word *Machine-It*. But the result, of course, is that the second person is felt to be equally mechanized; the discursive "you" stays at the level of convenient convention; the discourse is humanized, but only on the surface. Indeed, for many readers, the result is felt to be overly restricting, telling the user what to do, whereas the third-person discourse of the professional text allows greater liberty: one can use the information for whatever one wants to do. That kind of freedom might ultimately be more humanizing. The *deus ex machina* is ultimately an unwanted intrusion into private scenes.

Our reflections here actually come from a very common teaching experience. Let us say we want students to learn to use a new software program. We give a demonstration, explain the principles, then leave them to get the thing going for themselves. We bequeath them our written notes; they have an online tutorial; they have a very complete Help file; all the information they could ever need is there, in written form, with all the written second and third persons imaginable. So what do they do? Within seconds they are screaming for personal individual attention from a human, hopefully a fellow student, more often the overworked teacher. Why should that be so?

The same thing happens, of course, in discussions between professionals or experts. The information is all in written documents somewhere, but what we want are the insiders' tips, the personal opinions, the individual comparisons, the chatting exchanged between first persons.

At both these levels, the second person looks for a first person. The second-person text goes part of the way there, but it requires completion in another mode, just as a Bible is completed in the communication of prayer.

Mediation and its vices

Is there a message here for localizers? Start praying, perhaps? Or better: stop praying to the marketing experts ever ready to profile "what the customer wants". It might be more important to think about the social effects of our divisions.

The more immediate problem, though, is that the people localizing technical discourse often have little vision of any second person to be addressed, let alone any possible social effects. Large teams mean that individuals work on fragments, necessarily leading to explicitness, mimicking the non-linearity with which their texts will be read. Bert Esselink (2000: 478-480) even envisages a near future where translators work only on the minor updates, produced only every so often so that localization companies can spread their workloads more evenly. Such fragmentation means, of course, that translators will never see the entirety of any project, and will never know to whom their work might be addressed. The result is work that is not just boring, but dehumanizing in virtually whatever sense you want to give the term. There lies the rub.

True, the bulk of the localization workload comes later, with revision processes, product testing, updates and more updates, along with the manifold coordination problems of project management. This is usually in-house professional stuff, where we might hope to find a first person prepared to think about a second person. In the real world, however, the factors dominating these processes tend to be the deadline, the legal regulation, and the market, rarely anything like human values. We remain within the world of third-person things.

The general underlying problem could be mediation, as in religion. The more stages and layers a message has to pass through, the more it is adapted to specific user profiles (localization, indeed), the more restrictive tends to be the resulting knowledge. Think of the medieval Vulgate, written in Latin to keep it from the masses, who thus had a religion of spoken narratives. Thus might accessible Help files keep people in their place. Far better, I suggest, to do the work of a Luther or a Wycliff, and bring all readers to the complexities, ideally placing all in the position of open-code producers. Of necessity, if not faith, the readers will find their way through.

At the end of the day, as if by paradox, it is in excessive mediation that our dehumanization occurs.

References

Bernstein, Basil. 1971. *Class, Codes and Control Vol. 1: Applied Studies towards a Sociology of Language*. London/ Boston: Routledge & Kegan Paul.
Bödeker, Hans Erich. 1982. 'Menschheit, Humanität, Humanismus'. Otto Brunner, Werner Conze, Reinhart Koselleck eds. *Geschichtliche Grundbegriffe.*

Historisches Lexikon zur politisch-sozialen Sprache in Deutschland. Vol. 3. Stuttgart: Klett-Cotta. 1063-1128.

Buber, Martin. 2000. *I and Thou*. Trans. Ronald Gregor Smith. New York: Scribner. First published in German as *Ich und Du*, 1923.

Esselink, Bert. 2000. *A Practical Guide to Localization*. Amsterdam & Philadelphia: Benjamins.

Labov, William. 1972. "The Logic of Nonstandard English". Pier Paolo Giglioli ed. *Language and Social Context*. Harmondsworth: Penguin 179-125.

Laygues, Arnaud. 2001. Review article of Buber, Marcel, and Levinas. Anthony Pym, ed. *The Return to Ethics*. Special issue of *The Translator* 7(2): 315-319.

O'Donnell, James. 1998. *Avatars of the Word. From Papyrus to Cyberspace*. Cambridge Mass, London: Harvard University Press.