

Nutzlos, Bien Pratique, or Muy Util?

Business Users Speak Out on the Value of Pure Machine Translation

D. Verne Morland
Global Learning Division
NCR Corporation
1700 S. Patterson Blvd., WHQ-3
Dayton, Ohio, U.S.A. 45479

Verne.Morland@ncr.com

Introduction

Few people are apathetic when it comes to the corruption of their native language. Although most of us do not expect business documents to be pure poetry, when we are presented with information that could plausibly have been written by a 5 year-old, we are apt to be vocal in our displeasure or snide in our laughter. On the other hand, if we find ourselves in a job in which much of what we are expected to do is communicated in a language that we do not understand well, we may be more tolerant of odd words, poor syntax, and awkward style in order to have access to useful information.

This paper addresses the steps global companies should take to ensure that their early forays into the world of pure machine translation (MT) are successful in the eyes of the most important audience - the users of the translations. The data presented below both supports and contradicts several points of conventional wisdom regarding the use of translated materials within global corporations.

Case Study

NCR Corporation, headquartered in Dayton, Ohio, is a global technology solutions and services company with 32,000 associates in 80 countries. Over 50% of NCR's workforce resides outside the United States.

Like most modern companies, NCR is very good at "getting the message out" - publishing reams of announcements, brochures, instructions, and other documents to our employees and customers. More challenging by far is "getting the message in" -

ensuring that these messages are read, understood, and acted upon by the recipients (Morland, 2002, 1).

Although English is the official language of the company, many associates are not fluent in English. This impairs not only their ability to read company documents and converse with their English-speaking colleagues; it also makes it more difficult for them to stay abreast of global company developments and even to take advantage of specific opportunities - for example, training programs - that would help them improve their performance.

NCR's Global Learning division identified two major contributors to effective multinational communication: 1) personalizing the message and 2) presenting the message in the recipient's preferred language.

In 1998 we introduced a personalization service on the company's online university web site - NCRU. This service, called "MyNCRU," allows users to submit a simple profile of their job responsibilities and interests and it then provides them with a personalized interface containing news and calendar information tailored to their profiles. In addition, in July 2001 the division began publishing a monthly email newsletter - the *MyNCRU Personal Learning News* - that is individually constructed for each of more than 6,000 recipients (Morland, 2002, 2).

One of the fields in the profile is "Preferred Language." At this time, of the 14,893 current employees who have registered for the MyNCRU service, 1,528 or about 10% have indicated that they would prefer a language other than English. The breakdown, based on the choices we provided, is:

English:	13,365	89.7%
German:	413	2.8%
French:	351	2.4%
Spanish:	346	2.3%
Japanese:	192	1.3%
Italian:	174	1.2%
Other:	52	0.4%

Table 1: Language Preferences of Registered Users of the MyNCRU Service

The Translation Challenge

NCR was founded in 1884 and the following year its founder, John H. Patterson, opened sales agencies in three countries outside the U.S. As a result of doing business internationally for over 100 years, NCR recognizes the importance of translating and localizing much of its sales collateral, solution documentation, and training. However, the combination of personalization and translation presented Global Learning with a special challenge: how to translate hundreds of copies of a monthly newsletter, each copy of which is unique to its recipient.

To address the needs of our non-English-speaking audience, in the spring of 2000 Global Learning sponsored a competitive MT "fly-off" evaluation in which two suppliers - SYSTRAN Software and Lernout & Hauspie - were invited to demonstrate the effectiveness of their systems in performing real-time, pure MT translations of sample pages from the NCRU web site. The structure and outcome of that evaluation is documented elsewhere (Morland, 2002, 1). Suffice it to say that we tested four languages (French, German, Spanish, and Japanese), selected the SYSTRAN Enterprise Server, and started a customized implementation project.

Of particular interest, considering the subject of this paper, the internal experts who evaluated the quality of the translations were not at all convinced that either MT system was ready for widespread use. We gave them four choices for what NCR should do with respect to the machine translation of web sites for their languages:

1. Use it now throughout the NCR intranet - it works.
2. Test it with a larger audience in selected areas of the NCRU web site only.
3. Do not apply it now, but continue to monitor progress in this technology closely.
4. Don't waste more time - this is still years away from a practical use.

None of the evaluators felt that either system was ready for a large-scale deployment, and only a couple thought it would be worthwhile to test them with a larger audience on one web site. Most suggested monitoring the technology further and a few believed that no use would be practical for several years. (Our Japanese evaluators unanimously vetoed any immediate implementation and so we deferred our acquisition of the translation engine for that language. We purchased an English to Italian engine instead.)

In a discussion of this negative result, one of the Spanish evaluators, an associate in Argentina, made an insightful observation (reproduced here verbatim):

"...I am fluent in English, and can read it effortlessly. (probably this is true with most of the evaluators). So, I surely prefer to read English than bad Spanish. But maybe it is not true for all the people that only reads English with great effort.

Maybe you could find a group of evaluators that need the translations and ask them not if the translation is good (it is not), but whether they would prefer to read the translated version, however bad, rather than the original."

The project team decided to press on, and the results achieved thus far with larger, more randomized audiences support the observation that people who can read English only "with great effort" do in fact benefit from machine translations.

Producing a Global Newsletter with Pure MT

In July of 2001, Global Learning began production of the monthly newsletter and immediately began sending MT translations to those who requested French, German, Italian, or Spanish. The ground rules for the translated copies were the following.

1. Users would understand they were receiving pure MT
2. Users would have easy, one-click access to the original English text
3. Users would easily be able to change their language preference back to English if they found the MT translations unsatisfactory.

The process by which the translated newsletters are produced is illustrated in Figure 1 below. Each subscriber's personal profile is retrieved from the MyNCRU database (Step 1) and an English copy of his or her newsletter, constructed from news and calendar items stored in the NCRU News database, is created (Step 2). The English copy of the newsletter is stored on the NCRU web site and a copy is sent to the appropriate language engine for translation (Step 3). The translated copy is also stored on the NCRU web site and a copy is sent to the subscriber through the NCR intranet mail system (Step 4).

Note that the introduction to the translated newsletter provides the subscriber with a prominent link to the English copy. It also tells the reader that should he have any questions about the translation, the English original takes precedence and he can view the English copy in a separate browser window with a single click of his mouse.

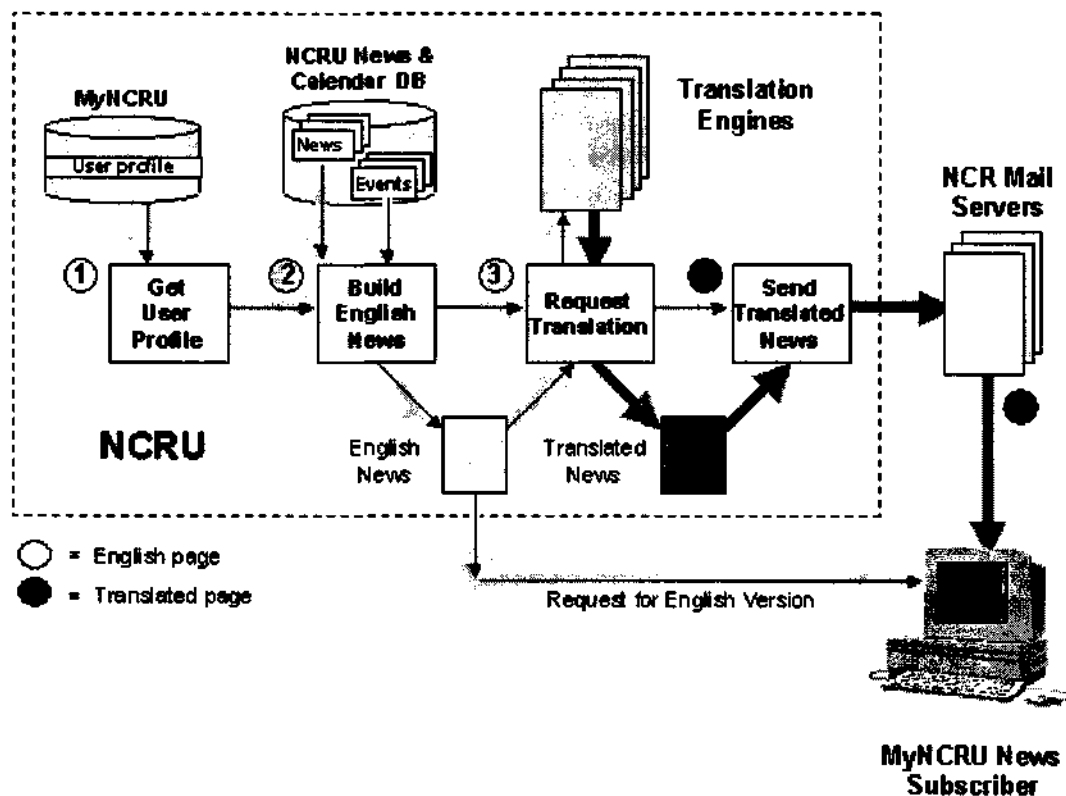


Figure 1: Process for building translated copies of *MyNCRU Personal Learning News*

Getting Input from Real Users

About one year after we introduced the MT translations (May 2002) we invited the 485 associates who had received one or more translated copies of the *MyNCRU Personal*

Learning News to participate in a survey on the value of that translation (see survey questions and results in Appendix A). We received responses from 280 associates (58%). This response rate implies that the quantitative results are accurate within a confidence interval of plus or minus 4 percentage points at a 95% confidence level.

The invitations to the survey and the survey itself were translated - by humans - into each of the four languages: French, German, Italian, and Spanish. Our use of human translators was not, in itself, a serious criticism of machine translation. We wanted to ensure that the invitations were correct and professional in appearance and that the survey questions could be easily read and correctly interpreted.

The survey results provided interesting and provocative information in two forms: quantitative data and verbatim comments and recommendations.

On the quantitative level, the responses to most of the questions followed the traditional bell curve with a majority gravitating to the central tendency and two minorities in the positive and negative extremes. For example, when asked, "How would you rate the overall quality of the newsletter translation?" most (68%) held to the middle saying "Fair" or "Good." Similarly, when asked, "How would you rate the usefulness of the newsletter translation?" the majority (79%), again, fell near the median. However, when we compare these responses on a single graph (Figure 2) we see an important shift.

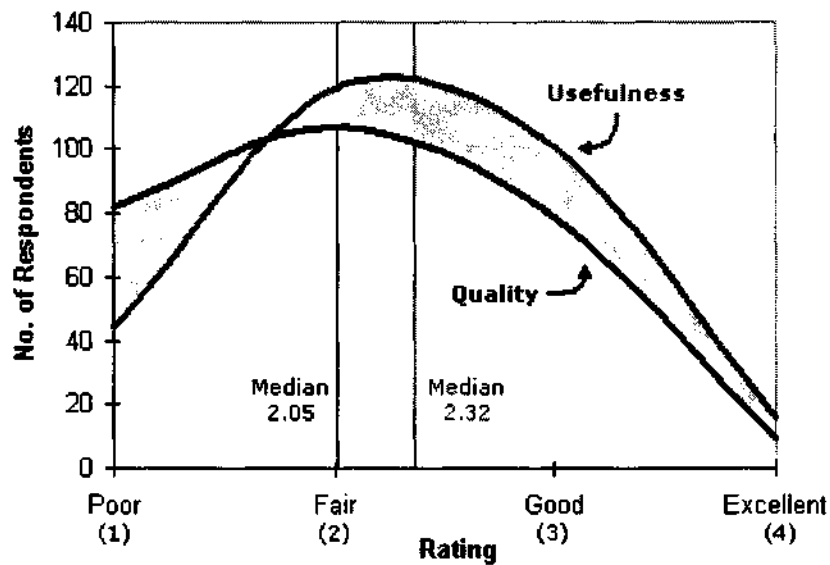


Figure 2: Comparison of Translation Quality and Usefulness Ratings

Although most respondents recognize that the quality of the pure machine translations is fair to poor, a significant number of *these same respondents* rate the translations higher when asked about the translations' usefulness.

This result confirms what many MT proponents have themselves experienced. Pure MT is rough - often obscure, frequently humorous - but it can be useful. If one really has little facility in the source language, pure MT translations, however clumsy, can be a boon to understanding and, by extension, to productivity.

Another key finding of our survey is that nearly two-thirds of the respondents said that they would recommend pure MT to their colleagues (Figure 3).

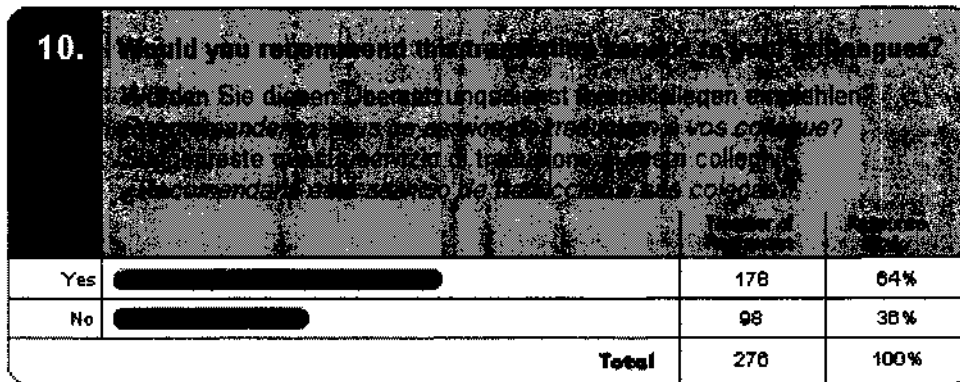


Figure 3: Responses to the "Would You Recommend" Question

We used cross-tabulations to study several relationships in which anecdotal feedback had piqued our interest. These provided insights into such questions as:

1. Are some language audiences more receptive to MT than others?
2. Do people with limited English abilities find the MT to be higher quality?
3. Do people with limited English abilities find the MT more useful?

But before getting to those questions, here is one useful finding that will assist in correctly interpreting the data on subsequent correlations. In our survey we asked respondents to self-assess their ability to read and understand English. As Figure 4 shows, there was no significant difference in this factor across the four language groups.

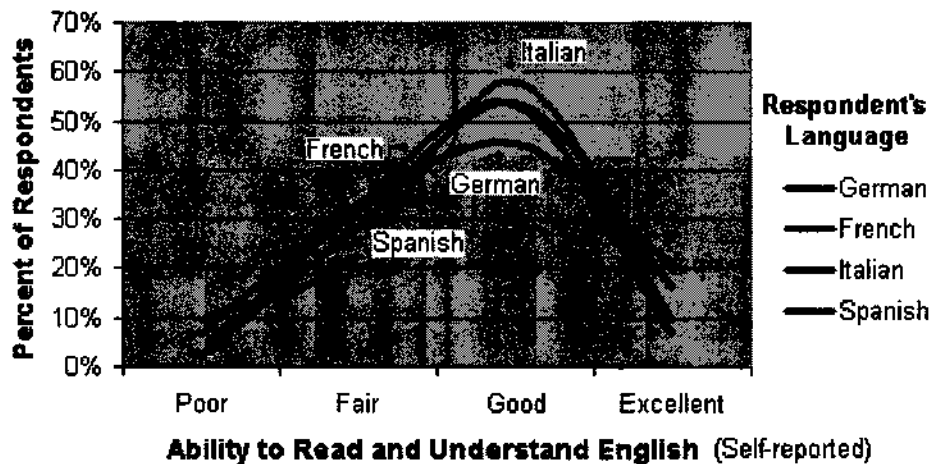


Figure 4: English Ability by Respondent's Native Language

The majority of respondents in each language report that they have a good understanding of English. (This is not necessarily true for all NCR employees. Based on our response rate, the sample is representative of all NCR employees *who have received one or more copies of the MT translated learning newsletter.*)

Figure 5 illustrates the breakdown of responses to the question, "How would you rate the overall quality of the newsletter translation?" by language group.

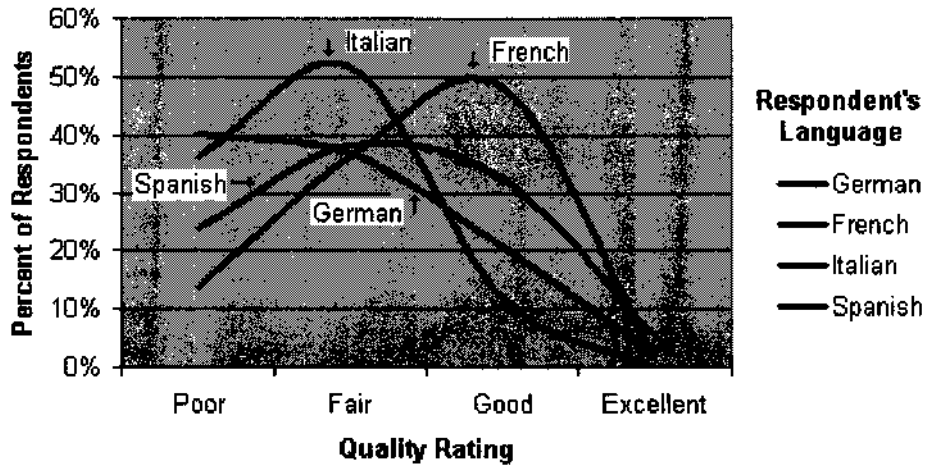


Figure 5: Translation Quality Ratings by Respondent's Native Language

Not surprisingly, very few respondents rate the quality as excellent. The 88 German respondents take the dimmest view of translation quality and the 58 French respondents seem the most impressed. This may be due to the French translations simply being better than the German, or it may be that the French respondents were more tolerant than their German counterparts.

When we look at the quality ratings broken down by the respondents' (self-reported) ability to read and understand English (Figure 6), we see a fairly predictable pattern.

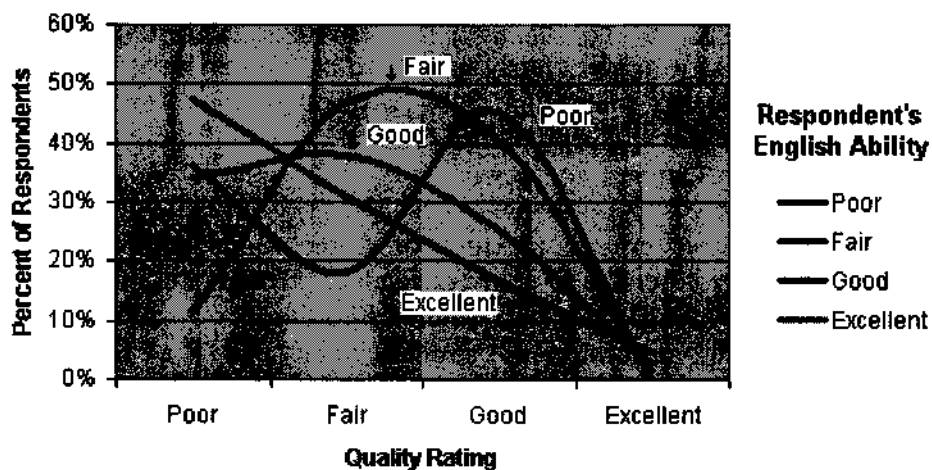


Figure 6: Translation Quality Ratings by Respondent's English Ability

As before, very few respondents rate the quality as excellent, yet we can now see that those who say they understand English well, rate the quality of pure MT translations the lowest. Those who say they have a "good" understanding of English are also negative on MT, but less so. And those who report that they are only "fair" or "poor" in English are progressively more forgiving of MT's mistakes.

In Figure 7 we take this insight and move beyond the notion of translation quality to the more salient issue of usefulness. When we segment answers to the question, "How would you rate the usefulness of the newsletter translation?" by the respondents' English ability, we see an even stronger vote in favor of MT by the two lower groups.

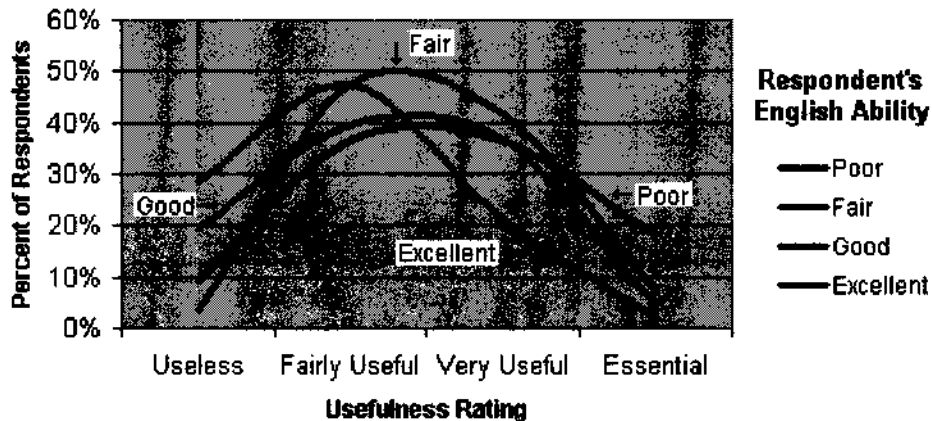


Figure 7: Translation Usefulness Ratings by Respondent's English Ability

Note that 18% of those who rate their English ability as "poor" actually consider the MT translation "essential" (although in absolute terms this represents only 2 of 11 in that ability category). Perhaps more important, of the larger group of 84 respondents who said that their English was only "fair," a full 90% said that the translation of the newsletter was "fairly useful" (48%) or "very useful" (42%). Even among the largest group - the 155 who said their English is "good," 76% (118) also found the translation of the newsletter "fairly useful" (39%) or "very useful" (37%).

On an overall basis, if we consider "fairly useful" to be the minimum rating that would justify the use of a pure MT translation, 84% (246) of the total sample, regardless of language ability, rated it at that level or better. (See the overall summary of responses to Question 6 in Appendix A.)

Next we look at another critical measure of a program's success: "Would you recommend this translation service to your colleagues?" As Figure 8 illustrates, the answers to this question, by language group, parallel the attitudes revealed in Figure 5. The French respondents, being most impressed with the quality of MT are most likely to recommend it and the German respondents, being least impressed, are least likely to want to inflict it on their compatriots.

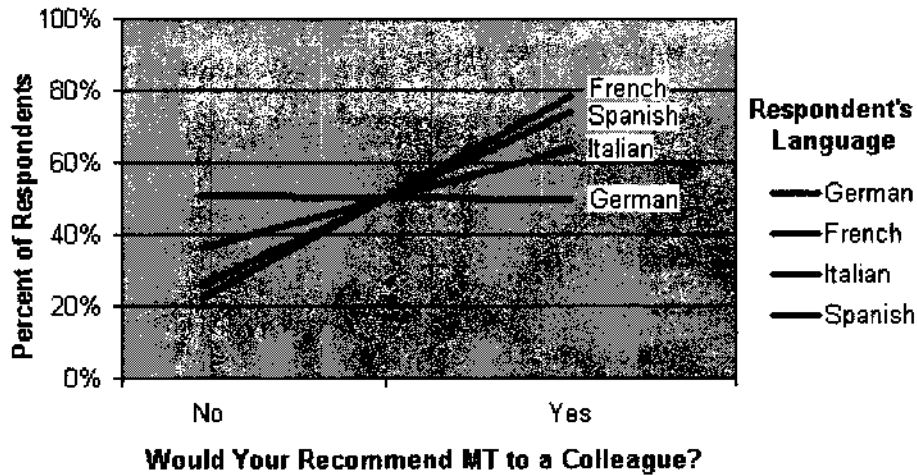


Figure 8: Recommendation to Colleagues by Respondent's Native Language

An important question that is always on the mind of effective communicators is: "How many recipients of my materials simply discard them without reading them?" In marketing terms this might be phrased, "What is the leaked market?"

Overall, when asked directly, "Would you read this newsletter if it were not translated into your language?" 16% (44 of 280) said, "No." (See the responses to Question 7 in Appendix A.) Figure 9 gives some additional insight into this issue by language group.

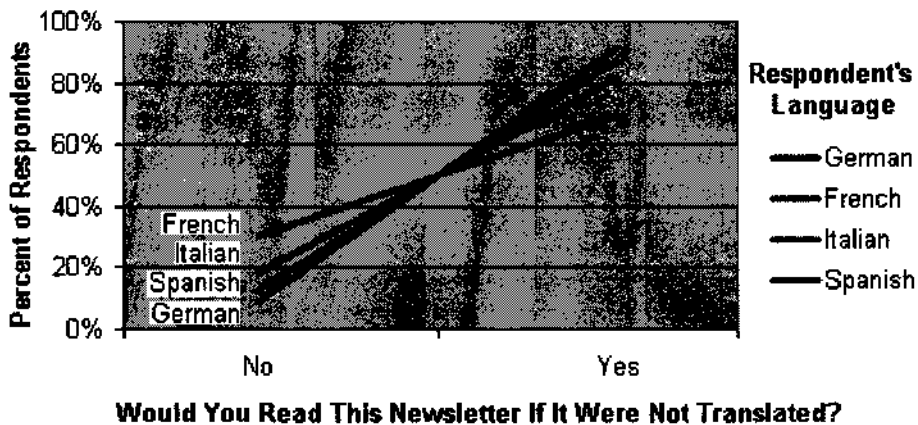


Figure 9: Willingness to Read English Newsletter by Respondent's Native Language

Out of 59 French respondents, 18 (31%) would be unwilling to read the newsletter if it were not translated. Six of 33 Italians (18%) and 13 of 105 Spaniards (12%) also said, "No." While this "leakage" might be acceptable in a mass market promotion, it would certainly be unacceptable to most managers of internal communications.

Finally, let's address a common myth in many U.S.-based multinational corporations, namely, that their managers are more fluent in English than the general workforce. As

Figure 10 shows, for this survey the self-reported ability to read and understand English did not vary significantly with organizational level.

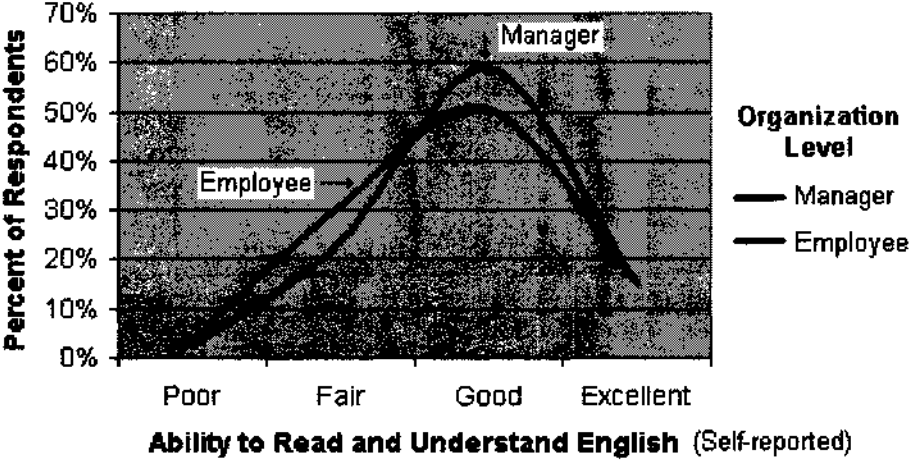


Figure 10: English Ability by Respondent's Level in the Organization

Qualitative Comments - the Voice of the Customer

As noted in the introduction to the previous section, on the quantitative level, the responses to most of the questions graphed as traditional bell curves with varying means, degrees of skew, and dispersal. The qualitative responses to our open-ended questions, however, were less moderate.

I have grouped a sampling of this user feedback into three categories: positive, negative, and forward-looking. To get the flavor of each consider the following. When asked, "Do you have any general comments or recommendations about using machine translation at NCR?" those who were impressed with MT's usefulness said things like:

- It is OK! Die Übersetzungen sind sehr gut zu verstehen.
(The translations are very easy to understand.)
- C'est bien pratique.
(It is very practical.)
- Ceci ajoute a ma comprehension.
(This adds to my comprehension.)

Table 2: Positive Verbatim Comments about Pure Machine Translation at NCR

On the negative extreme, respondents gave us the following pieces of their minds.

- Maschinelle Übersetzung ist nutzlos.
(*Machine translation is useless.*)
- This poor German hurts in my eyes.
- La traducción automática sólo sirve para hacer reír
(*Automatic translation only serves to make me laugh.*)
- Elimínenla! Destrozar un idioma es lastimoso...
(*Eliminate it! To destroy a language is pitiful...*)

Table 3: Negative Verbatim Comments about Pure Machine Translation at NCR

And we also received some interesting suggestions from respondents who were impressed with MT's potential for the future.

- Je pense qu'elle devrait être généralisée dans tous les pays et pour chaque employé. (*I think that it should be generally available in all countries and for each employee.*)
- Me gustaría disponer del traductor automático, instalado en mi PC.
(*It would please me to have the automatic translator installed on my PC.*)
- Es wäre gut, diesen Service auch für eigene Dokumente nutzen zu können... (*It would be good to be able to use this service also for our own documents...*)
- ¿Existe algún tipo de servicio de traducción de voz para conversación telefónica? (*Is there some type of translation service for voice during telephone conversations?*)

Table 4: Forward-looking Verbatim Comments about Pure Machine Translation at NCR

We asked one other open-ended question: "Are there other NCR publications and documents you would like to have translated by machine?" Some respondents were undecided ("not really") or felt they didn't need it ("not necessary"). Others made specific suggestions - "Rapport annuel NCR" (*NCR Annual Report*), "Manuales de Clientes y Planes Corporativos" (*client manuals and corporate plans*), "informations d'utilité quotidienne ou opérationnelle" (*useful daily or operational information*).

Some were eager to apply MT across the board - "tutte le comunicazioni in lingua inglese" (*all English communications*) - or more simply: "tout," "todos," "tutte," and even, despite the general attitude of the German respondents, "alle."

Of course, on the other end of the spectrum we had a fair number like this: "NO! It's senseless in the actual quality!", "No and never," and "NO!!!"

From our perspective, the voice of reason was this: "La informacion es siempre mas comprensible y de facil asimilacion si es enviada en las lenguas madres." (*Information is always more comprehensible and easy to assimilate if it is sent in the mother tongue.*)

It is interesting that most of the negative opinions were expressed in English, whereas the positive suggestions were generally made in the respondent's native tongue. This confirmed again our thesis that those who feel comfortable with English do not want pure MT translations, but those who are not as strong in English find it useful.

Conclusions

Overall the responses to our survey generally supported Global Learning's basic hypothesis about the use of pure MT, namely, that associates who are fluent in English rate machine translation to be of lower quality and less useful than associates for whom English is difficult. While that in itself may not seem surprising, our exploration of a less-frequently-heard-from audience - those who have difficulty with English - has led us to some critical insights regarding the successful deployment of this new technology in a global business.

Typically, translations of business documents are reviewed by associates who are fluent in both the source and target languages. Based both on their expertise and on the fact that they are more interested in languages than the average employee, these reviewers are often quite critical of the pervasive clumsiness and occasional inaccuracy of translations performed entirely by machine. Were we to listen only to the recommendations of this group, unreviewed machine translations would never see the light of day.

In any large, multinational group, however, there are certainly others who are not as proficient in reading and understanding English. These people tend to be underrepresented in policy decisions about the importance of translation. At NCR, as in most U.S.-based multinational corporations, the official language is English and while this may be quite a satisfactory arrangement for the English-speakers who established the policy, it is a problem for many others.

As the findings of our survey clearly illustrate, many NCR associates find the pure MT quite useful and a significant number (16% overall) told us, point blank, that they would not read our newsletter if it were not translated. Losing this portion of our audience entirely and compromising the utility of our newsletter for an even larger group is simply not acceptable.

The data, interpretations, and user comments presented above can be summarized as follows.

1. Ratings of MT usefulness are higher than ratings of MT quality.

2. A majority of those who have used pure MT in the *MyNCRU Personal Learning News* would recommend it to their colleagues.
3. A significant majority (84%) of all users, including those who report that they possess a fair to good understanding of English, find value in pure MT, rating it "fairly useful" or better.
4. A comparatively large minority (16%) said they will not read the subject publication unless it is translated.
5. Managers are not significantly better than employees in the general workforce in their (self-reported) ability to read and understand English,.
6. Some people who have been exposed to pure MT strongly dislike it and may even feel offended by it.
7. Some people who have been exposed to pure MT like it very well and would like to see its use extended to many other communications.
8. A large majority of respondents (71%) said they would be willing to spend a few minutes each month to improve the quality of MT output by identifying errors and suggesting corrections for their languages.

Recommendations

Reflecting on the results of this survey and on NCR's three years of experience with several facets of MT, a few key recommendations stand out. Here is the list; brief explanations follow.

1. Don't let internal language experts have exclusive input to MT decisions.
2. Guide users' expectations, give them easy access to the source text, and always give them the choice to opt-out of the MT program.
3. Make sure that a significant part of your MT budget is devoted to planned, ongoing refinements of the MT system and its control files.
4. Use a rapid prototyping approach to test new ideas on small, receptive audiences.
5. Develop a partnership relationship with your MT supplier.
6. Persevere.

NCR's survey clearly shows that the target audience is not multilingual internationalists. Every global corporation has people who are fluent in several languages and who appreciate and enjoy the nuances of transnational communication. Although they are not trained linguists, these are the people who are usually called upon to assist with internal

translation projects and, in general, they will not be satisfied with the output of pure MT running against uncontrolled source text. The primary beneficiaries of this technology are those employees who are not fluent in the source language and who will be made more productive by using the translations. While this distinction seems obvious, the lesson we learned is that the vocal disdain of the internationalists we consulted at the beginning of the project almost drowned out the growing chorus of encouragement that we received from true end-users as the project progressed. Remember the words of Samuel Johnson, "Nothing will ever be accomplished if all objections must first be overcome."

Make it clear to users that they are using MT. Give them quick and easy access to the source text from which their translation was produced. And always give them a choice to change their language preference back to the source language at any time. We preface each copy of the translated issues of the *Personal Learning News* with these statements.

"Machine translation is a new technology that produces output that can be incorrect or hard to understand. If you speak English well you may prefer to [change your preference to English](#). If you have difficulty with English, we think you may find this translation useful. If you have any questions about the translation, you can easily [view the original English version](#). If there are any discrepancies, the English original takes precedence."

The phrase "change your preference to English" is a hyperlink that takes users directly to a web page on which they can update their personal profiles. The phrase "view the original English version" is another hyperlink that opens a pop-up window containing the original English text.

Successful MT implementations require resources for continuous updates and improvement. The cost of these resources typically exceeds what is normally budgeted for software upgrades and maintenance as a percent of purchase price. Most MT systems have control or "tuning" files that specify preferred translations of selected words and phrases. They may also have lists of words that should not be translated due to their widespread use in their original forms. Many of the problems identified in our initial evaluation and many of the poor translations cited by our current users can be corrected by these control files.

The maintenance of these files requires the continuous intervention of people, in some cases linguists. The cooperation of rank and file users is also helpful to identify areas that need improvement. In our survey, we were delighted to see that 71% of our MT users said they would be "willing to spend a few minutes each month to provide this input" - the identification of errors and suggested improvements for their languages. Our challenge now is to find ways to channel this volunteer energy into direct and constructive inputs to our control files.

Our experience suggests that big projects involving new technologies are amenable to implementation in stages using a series of rapid prototypes with ever-larger user groups. We also found that our time scales were longer than we anticipated and we were grateful

for the fact that we had developed a long-term partnership relationship with our supplier. This allowed us to weather a number of setbacks and delays without jeopardizing the overall success of the project. (Given the technical nature of MT projects, it is also very desirable to get your MT supplier to assign a good liaison person to your account - preferably one of the designers on the technical staff, not a help desk technician.)

Although MT has been a part of the computing universe almost since its inception in the late 1950's, it is only now beginning to make the transition from research and government projects to practical applications in business. In this sense it is still a leading-edge technology with all of the challenges that phrase implies. Persevere.

Bibliography

Morland, D. Verne 2002 (1). "Getting the Message In: A Global Company's Experience with the New Generation of Low-Cost, High Performance Machine Translation Systems", Proceedings of the AMTA-2002, October 2002.

Morland, D. Verne 2002 (2). "Promoting Learning with Personalized Newsletters", The Technology Source, November/December 2002, <http://ts.mivu.org/>.

Appendix A:

Basic Results of the Survey on the Value of Machine Translation at NCR

1. In what language do you receive the NCRU newsletter?

- French	58	21%
- German	77	28%
- Italian	33	12%
- Spanish	103	38%

2. Are you still receiving the translated version of the NCRU newsletter or have you changed your preference back to English?

- Still receiving translation	186	68%
- Reverted back to English	89	32%

3. How would you rate your ability to read and understand English?

- Excellent	41	15%
- Good	141	54%
- Fair	76	28%
- Poor	11	4%

4. How many translated copies of newsletter have you received?

- Five or fewer	192	69%
- More than five	86	31%

4. How would you rate the overall quality of the newsletter translation?

- Excellent	9	3%
- Good	79	29%
- Fair	107	39%
- Poor	82	30%

6. How would you rate the usefulness of the newsletter translation?

- Essential	16	6%
- Very useful	101	36%
- Fairly useful	119	43%
- Not useful at all	44	16%

7. Would you read this newsletter if it were not translated into your language?

- Yes	236	84%
- No	44	16%

7. How often do you have to refer back to the English copy?

- Very often	61	22%
- Sometimes	102	37%
- Rarely	79	29%
- Never	35	13%

9. Are there other NCR publications and documents you would like to have translated by machine? (Please list the documents or types of information.)

- Open: 94 responses

10. Would you recommend this translation service to your colleagues?

- Yes	178	64%
- No	98	36%

11. Do you have any general comments or recommendations about using machine translation at NCR?

- Open: 91 responses

Demographics:

- 12. Division
- 13. Job Role
- 14. Level (manager or individual contributor)
- 15. Country
- 16. Name (optional)

Request for User Assistance:

17. Machine translation will constantly improve if readers are willing to identify errors and suggest corrections. Would you be willing to spend a few minutes each month to provide this input for your language?

- Yes	186	71%
- No	77	29%