Has machine translation improved? historical comparisons[1]

W. John Hutchins
(Email: WJHutchins@compuserve.com)
(Website: http://ourworld.compuserve.com/homepages/WJHutchins)

1. Introduction

A question frequently asked by those new to the field is whether machine translation has improved – in the last five years, or in the last ten years, or in the last twenty, etc. In many respects, the answer is quite easy. It is obvious that programs for automatic translation run much faster than in the past because computers are faster. Equally obvious is that text input is much easier than in the days of punched cards and paper tape, and that the output is much easier to read than in the days of almost illegible computer printouts all in upper case. MT systems are also becoming cheaper by the year, even by the month: what cost over two thousand dollars five years ago is now on sale for under five hundred dollars. Finally, it is plain that more and more large organisations can and do save translation costs by using MT systems as aids in the production of draft translations. Twenty years ago this would have seemed almost inconceivable.

However, these are not the improvements that are usually meant when the question is asked. The 'real' question is whether the quality of machine translation has improved. The answer is not at all obvious. Some will say that it *must* have. Others of an older generation insist that MT has the same problems and throws up the same errors as it did in the 1960s and 1970s. Many, however, will perhaps share my instinct that while progress in quality is not evident during say the last ten years, there are discernible advances since the early 1980s and definite improvements since the ALPAC report of the mid 1960s. But what is the evidence for this impression, strong as it may be. Indeed, perhaps we are being misled or dazzled by the genuine and widespread technical advances. Perhaps, MT has not improved all that much. And if it hasn't why should this be.

This paper will be an attempt to look at the output of MT systems from earlier periods of the field's history and to see whether today's systems would perform any better on the same texts.[2] It is an informal (i.e. wholly unsystematic) effort at historical comparison. It is unsystematic because it does not attempt to conform to any accepted evaluation methodology – being completely subjective. Indeed, there does not seem to be at present any available methodology for longitudinal evaluation. But it is unsystematic primarily because it is necessarily based on partial evidence. It is surprisingly difficult to find examples of actual translations by MT systems before the last decade. Most often, we find fragments (individual sentences) and not extended texts. If we do find more extensive translations, they are usually given without the original source texts – and so no comparisons with other systems (of the time or later) are possible. Furthermore, nearly all the examples are from research prototypes – and often, selected in order to give the most favourable impression of the system's capabilities. Some have been included in this paper out of historical interest – and to

[1] This article is a longer version of the one presented at the MT Summit IX conference in New Orleans (September 2003), published in the proceedings of the conference. It provides more and longer examples and includes a section on English-French translation.

[2] For general historical overviews of the development of MT since the 1940s see Hutchins 1986, 2003.

illustrate how untypical of actual performance they may be – but they have been disregarded in the overall assessments of progress in the concluding sections. A final limitation is that comparisons with modern systems is restricted to those readily available to the non-specialist user – this means commercial systems for personal computers and on-line translation services. No doubt, researchers will claim that more recent developments in corpus-based techniques will produce better results, but as far as the general public is concerned, these potential improvements are not yet real.

Comparisons are made on the basis of 'raw' (unedited) output. It is recognised that this should not be the sole criterion for evaluating the usefulness of systems (e.g. for producing publishable quality texts), but it is only by examining unedited output that we can judge whether MT engines themselves have improved – and not post-editing facilities. Whatever the context of use (dissemination, assimilation, etc.) it must surely be true that the better the quality of 'raw' translation the more "useful" the system. Among many other factors, evaluations should be able to say whether a new product (or a new version of an old product) is or is not better in terms of translation quality than an older product (or an older version of the same product).

For the sake of keeping the survey within bounds I have looked at Russian-to-English systems of the 1960s and 1970s, at English-to-French systems of the mid 1970s, and at French-to-English and German-to-English systems of the early 1980s and later. I have selected example translations which appear to be genuine and compared their output with those produced by some current systems from the same texts. Some, however, are suspiciously good, and may well have been doctored for publicity purposes.

Some of the source texts in the examples are difficult to translate even for expert translators. In some cases, it may be that they were originally chosen to illustrate how badly MT systems were performing. On the other hand, it can be argued that it is precisely the most difficult texts for which those with only weak knowledge of the language need most help and where assistance from MT would be most welcomed.

Comparisons with current systems have been confined to commercially available systems and/or to on-line MT services. The Russian-English examples have been compared with (a) the output by Systran, i.e. the latest package (Systran Personal 4.0) and two on-line services (Babelfish and Lycos) – in most cases all three produce identical results, but from time to time there are differences – and (b) the output from PROMT On-line and/or the latest package @promt Standard. For French-English, Systran is again used, but also output from the online services FreeTranslation and Reverso. For German-English comparisons are made not only with Systran output but also with results from Personal Translator PT, Reverso, and FreeTranslation.

This paper is an extended version of a paper given at the MT Summit IX conference in New Orleans, September 2003. Although it gives more and longer extracts, and also includes examples from non- or pre-operational systems of the past, it does not include complete texts and translations. These will be found on separate files on my website, which constitutes therefore the beginnings of a database or archive of older MT systems. From time to time it is hoped to augment this database with more examples as they are discovered in documents, company reports and research team archives.

2. Russian-English

2.1. IBM-USAF Translator, 1959

The first genuine output of a MT system is provided by descriptions of the Russian-English system developed by Gilbert King for the US Air Force Foreign Technology Division (Dayton, Ohio). This system was based on a special-purpose high-capacity dictionary storage device (the 'photoscopic store'). Essentially it was dictionary-driven word for word system, where dictionary entries included cues for rearrangement of target language output under specific conditions. These conditions were primarily lexical cooccurrences and concatenations of grammatical categories. King's system was based to a large extent on the lexicographic research of Erwin Reifler and his colleagues at the University of Washington (Seattle) on a Russian-English MT systems – many examples of the Washington output were published, but unfortunately without giving the original Russian texts (e.g. Reifler 1960). The following is an illustration:

(1) Problem about/against/with rolling/sliding riding (of)difficult/heavy/serious (of)solid/hard/firm (of)body(s) on/by/along/for/in/-- (of)(to/for)(by/with/as)horizontal *(of)(to/for)horizontal-surface in/to/at/on/of/like present/real time/tense (is)solved/allowed/absolved only/as-soon-as in/to/at/on/of/like (of)few (of)private/particular/quotient/partial cases/chances/occurrences

Even without having the original Russian, we can be confident that current systems are an improvement.

The same is true for the IBM-USAF system. An example of an original Russian text (2) and the English output (3) follows:.

(2) [original Russian]: В арсенале средств научно-атеистической пропаганды, материалистического воспитания трудящихся появилось новое оружие – журнал "наука и религия". Вышел в свет его первый номер. Это сборник содержательных, разнообразных, с интересом читающихся материалов, острие которых направлено против религиозных суеверий и предрассудков.

Журнал издается Всесоюзным обществом по распространению политических и научных знаний – массовой самодеятельной организацией передовой советской интеллигенции, призванной активно помогать партии в осуществлении решений XXI съезда КПСС, в формировании человека коммунистического общества. Выполнение этой благородной задачи предполагает полное преодоление религиозной идеологии, пропаганду материализма и атеизма.

(3) [IBM-USAF Translator, 1959]: In arsenal means scientific-ateistichesky propaganda, materialistic training worker appeared new weapon – magazine "Science and Religion". Issueed in light it first number. This collection containedkh, various, with/from interest reading matterevilv, sharp that directed against religious suerveriy and prejudice.

Magazine publish All-Union society by propagation political and scientific knowledge – mass organization advance Soviet intelligentsia, prizvannoy active help party in realization decision/solution XXI congress CPSU, in forming man communist society. Accomplishment this noble problem assume full surmounted religious ideology, propaganda materialism and ateizma.

The IBM-USAF lacked some items of vocabulary, failed to interpret some case endings *(-kh*), produced an inexplicable translation of *материалов* as *matter* and *evilv*, created some aberrant English forms (*issueed*), and in general failed to analyse the Russian syntax. (For another extract and discussion, see Hutchins 2000). For the same text, Systran and PROMT produce the following (4, Systran) and (5, PROMT):

(4) [Systran]: In the arsenal of the means of scientific atheistic propaganda, materialist training of workers did appear new weapon periodical is "Science and religion". Its first number appeared. This is the collector/tank of the meaningful, diverse, with the interest read materials, whose point is directed against the religious superstitions and the prejudices.

Periodical is published by All-Union society on the propagation of political and scientific knowledge by the mass amateurish organization of progressive-minded Soviet intelligentsia, called to actively help party/batch in the realization of the solutions OF KHKHI of the Congress OF CPSU ⇊ - COMMUNIST PARTY OF THE SOVIET UNION⇊, in molding of man of Communist society.The accomplishment of this noble/precious objective

assumes the complete overcoming of religious ideology, the propaganda of materialism and atheism.

(5) [Promt]: In an arsenal of means of scientific - atheistic propagation, materialistic education of working the new weapon - magazine " a science and religion " has appeared. His(its) first number was published. It is the collection substantial, various, with interest of read materials which edge is directed against religious superstitions and prejudices.

The magazine is issued by the All-Union society on distribution of political and scientific knowledge - the mass amateur organization of the advanced Soviet intelligency called actively to help to a party(set) in realization of decisions of XXI congress of the CPSU, in formation of the person of a communistic society. Performance of this noble problem(task) assumes full overcoming religious ideology, propagation of materialism and atheism.

Neither Systran nor PROMT are good translations, although the gist can be extracted. Some failings of the IBM-USAF system are still present. Interpretation of the pre-nominal construction *с интересом читающихся материалов (*meaning *'materials to be read with interest*) still eludes both Systran and PROMT. In some cases Systran is more successful than PROMT (*point* vs. *edge*, *intelligentsia* vs. *intelligency, to actively help* vs. *actively to help*), in others PROMT is better (*collection* vs. *collector/tank*, *of KHKH* vs. *XXI*).[3] Both, however, are clearly better than the IBM-USAF version.

A later version of the IBM-USAF systems was demonstrated in 1964 at the New York World's Fair. The quality of the translated examples (Hutchins 1986: 68) was suspiciously flawless:

(6) [IBM-USAF Translator, 1965]: All this page is machine translation of Russian text, which is printed on preceding page. This page of translation from Russian into English is not perfect due to unsolved problems of grammar.

Before machine can translate from one language into another, linguists have to introduce in memory unit of machine large quantity of grammatical rules, which increase intelligibility of given translation. But because in language exists significant variety and complexity, all grammatical rules of any language are not developed completely at present time for use by computers.

The best effort of today's Systran is:

(7) [Systran]: This entire page is the machine transfer/translation of the Russian text, which is printed on that preceding page. This page of translation from Russian into English is not perfected as a result of the unresolved questions of grammar.

Before machine it can translating from one language to another, linguists must introduce into the storage unit a large quantity of grammatical rules, which increase the clearness of this transfer/translation. On because in the languages there is a significant variety and a complexity, all grammatical rules of any language are not developed completely at present for the use by computers.

It seems clear that for demonstrations to the public at the Fair, IBM decided to ensure good-quality results. The true capabilities of the Mark II version cannot now be assessed as no other example translations have been located.

2.2. Georgetown systems at EURATOM and Oak Ridge National Laboratory, 1963

Apart from the IBM system, the only systems of the first "generation" of MT to come into operation were those delivered by the Georgetown University team in 1963 and 1964 to EURATOM at Ispra in Italy, and the US Atomic Energy

[3] It should be noted that the Babelfish version of Systran has problems with quotation marks and apostrophes. The raw output for the first paragraph was:

In the arsenal of the means of scientific atheistic propaganda, materialist training of workers did appear new weapon?the periodical of?nauka and religion?.Its first number appeared.this is the collector/tank of the meaningful, diverse, with the interest read materials, whose point is directed against the religious superstitions and the prejudices.Periodical is published...

If, however, they are omitted, such problems are avoided.

Commission at Oak Ridge National Laboratory. The project at Georgetown, led by Léon Dostert, had begun in 1954 after the public demonstration of the small-scale Russian-English system in New York (Dostert 1955, Garvin 1967); during the late 1950s and early 1960s it was the largest in the US. The systems installed at EURATOM and the AEC were the GAT-SLC version developed mainly by Michael Zarechnak and Tony Brown (Zarechnak 2000). It was not the caricature of a 'direct translation' system that it is often portrayed. Analysis of Russian passed through three stages: morphemic (inflections of individual words), syntagmatic (relations of immediately adjacent words), syntactic (identification of basic (nuclear) structures, governing elements of noun and verb phrases).

First an example from EURATOM:

(8) [original Russian]: О возможности возникновения нестабильностей в плазме, захваченной поперечным магнитным полем.
Исследованы условия возникновения и развития нестабильности в плазме, захваченной поперечным магнитным полем и прошедшей через диафрагму. Показано, что при заземлении металлической диафрагмы нестабильности плазмы не развиваются.
Вопрос об устойчивости плазмы, захваченной поперечным магнитным полем и движущейся вдоль магнитного поля, представляется весьма существенным при рассмотрении возможностей использования этой плазмы для заполнения магнитных ловушек.

In 1967, this was translated by the Georgetown system at EURATOM (Perschke 1968):

(9) [Georgetown, 1967]: Concerning the possibility of emergence of instabilities in plasma, which was captured by cross magnetic field.
Studied the conditions of emergence and the developments of instability in plasma, which was captured by cross magnetic field and which passed through diaphragm. Showed, that upon the grounding of metallic diaphragm of instability plasmas are not developed.
Question concerning the stability of plasma, which was captured by cross magnetic field and moving along magnetic field, is very essential upon the consideration of possibilities of utilization of this plasma for filling magnetic trap.

Obvious problems are verb-initial constructions (*studied*, *showed*), the difficulties of selecting appropriate articles – Russian has none – and the treatment of prepositions (in particular, *of* in genitive constructions).

Translations of the same text (8) by the current Systran system (10) and by PROMT (11) are:

(10) [Systran]: On the possibility of the appearance of instabilities in the plasma, seized by transverse magnetic field.
Are investigated the conditions for appearance and development of instability in the plasma, seized by transverse magnetic field and which passed through the diaphragm. It is shown that during grounding of metallic diaphragm the instabilities of plasma are not developed.
A question about the stability of plasma, seized by transverse magnetic field and which moves along the magnetic field, is very essential in the examination of the possibilities of using this plasma for the filling of magnetic traps.
(11) [Promt]: About an opportunity of occurrence нестабильностей in the plasma seized by a cross magnetic field.
Conditions of occurrence and development of instability in the plasma seized by a cross magnetic field and past through диафрагму are investigated. It is shown, that at grounding metal диафрагмы instability of plasma do not develop.
The question on stability of the plasma seized by a cross magnetic field and moving along a magnetic field, is represented rather essential by consideration of opportunities of use of this plasma for filling magnetic traps.

Systran still occasionally fails to deal with verb-initial constructions (*Исследованы...* translated as *Are investigated...*); both Systran and PROMT translate *захваченной* as *seize* rather than the better EURATOM rendition (*captured*); PROMT does not

recognise the plural *instabilities* (*нестабильностей)* and lacks dictionary entries for *diaphragm.* Nevertheless*,* in general, both Systran and PROMT are improvements*.*

Similar comments apply to the AEC examples (Henisz-Dostert 1967).

(12) [original Russian]: Экспериментальные результаты и их обсуждение.
Конденсированные на холодную подложку слои теллура обладали сопротивлением, лежащим в пределах нескольких сот килоом. Измерение сопротивления в большинстве случаев производилось в воздухе. Прямой и обратный ходы температурной зависимости сопротивления при нагревании и охлаждении не совпадают, причем после каждого цикла нагревание – охлаждение сопротивление образца при данной температуре возрастает

The Georgetown system at the Atomic Energy Agency produced:

(13) [Georgetown, 1964]: Experimental results and their discussion.
Condensed on cold support the layers of tellurium possessed resistance, which lie within the limits several hundredth kiloom. The measurement of resistance in the majority of cases was carried out in air. The straight and back courses of temperature relation of resistance upon heating and cooling do not coincide, yet after each ring heating – cooling the resistance of form at the given temperature increases

The pre-nominal construction (*Condensed…tellurium)* eludes the USAEC system – as it did the older IBM-USAF system above (3) and the modern systems (4) and (5). The term *килоом* was obviously not in the dictionary. In the last sentence inappropriate selections of translations for прямой, обратный and ходы make understanding difficult.

The current outputs from Systran and PROMT are:

(14) [Systran]: Experimental results and their consideration.
The layers of tellurium condensed on the cold base layer possessed the resistance, which lies within the limits of several hundred kilohms.The measurement of resistance in the majority of the cases was made in air. The straight/direct and back strokes of the temperature dependence of resistance during the heating and the cooling do not coincide, moreover after each cycle heating - cooling the resistance of sample at this temperature grows/rises

(15) [Promt]: Experimental results and their discussion.
The layers of tellurium condensed on a cold substrate had the resistance laying within the limits of several сот килоом. Measurement of resistance in most cases was made in air. Direct and return courses of temperature dependence of resistance at heating and cooling do not coincide, and after each cycle heating - cooling resistance of a sample at the given temperature grows.

The pre-nominal constructions (*condensed...*) which failed the Georgetown system (13) are correctly dealt with by both Systran (14) and PROMT (15). However, Systran repeats the puzzling *straight/direct and back strokes*; while PROMT is much better (but still not accurate) with *direct and return courses.* On the other hand, PROMT fails to translate сот (hundred), possibly because it lacks a dictionary entry for *килоом* (*kilohm*). In general, it is not obvious that current systems are always significantly better the Georgetown systems of the mid 1960s.

2.3 Systran 1970 and 1976

The Russian-English version of Systran was the first to become operational. It was installed in 1970 at the USAF Foreign Technology Division, replacing therefore the IBM system (section 2.2 above) – because it gave better results. We should expect, therefore, to find that the quality of Systran output in the 1970s to be high.

The first example is from about 1972 (Bruderer 1978a: 283-285). The Russian original (16) is translated by Systran in 1972 as (17):

(16) [original Russian]: Швейцарская общественность встревожена, пишет в одном из последних номеров базельская газета «Национальцайтунг». Недавние американские заявления о том, что США могут применить на Ближнем Востоке силу, подчеркивает газета, вызывают тревогу ко всем мире. Что касается Швейцарии, то, если такой курс получит продолжение, она рассмотрит вопрос о выходе из недавно созданного по

настоянию Вашингтона международного энергетического агентства, объединяющего ряд капиталистических стран, крупнейших потребителей нефти.

(17) [Systran, 1972]: The Swiss public is worried, the Basel newspaper "Nationalzeitung" writes in one of the last issues. Recent American statements about the fact that the USA can use force in the Near East, the newspaper emphasizes, cause alarm all over the world. As concerns Switzerland, then, if this course continues, it will examine the question concerning an exit from the recently created on the insistence of Washington international energy agency, which unites a number of the capitalist countries, the greatest users of oil.

Although there is, not surprisingly, a clumsy translation of the complex pre-nominal clause construction (*из недавно созданного по настоянию Вашингтона международного энергетического агентства*) as *from the recently created on the insistence of Washington international energy agency,* the mistakes in Systran in 1972 are negligible.

The current Systran system produces:

(18) [Systran]: Swiss community is worried, writes in one of the last numbers the Basel newspaper of "Natsional'tsaytung". Recent American statements, that THE USA can use force in the Near East, newspaper emphasizes, is caused anxiety to all peace/world. As far as Switzerland is concerned, if this course continues, she will examine a question about the output/yield from recently created on the insistence Washington of the international energy agency, which unites a number of the capitalist countries, the most important users of oil.

Inexplicably, the result is not as good. Where the 1972 version (17) has correctly translated вызывают as *cause*, the current version (18) has *is caused*. Where it has translated о выходе correctly as *concerning an exit*, the current system has *about the output/yield*. As we might expect, the pre-nominal construction is still problematic; but the current output (18) is possibly more garbled than that of 1972: *from recently created on the insistence Washington of the international energy agency.*

The PROMT system translates the paragraph as follows:

(19) [Promt]: Swiss community is worried, writes in one of the last numbers the Basel newspaper of 'Natsional'tsaytung'. Recent American statements, that THE USA can use force in the Near East, newspaper emphasizes, is caused anxiety to all peace/world. As to Switzerland, if such rate will receive continuation, she(it) will consider the problem on an output(exit) from Washington recently created on the insisting of the international power agency uniting a number(line) of the capitalist countries, the largest consumers of petroleum.

Like the current Systran (18), the first sentence retains the placement of the *Basel newspaper* before *in one of its last numbers*. It also has *is caused* instead of *causes*. In the second sentence, *if such rate will receive continuation* is markedly inferior to *if this course continues* (19) – and may be misunderstood. Like Systran (both old and new) PROMT fails with the complex pre-nominal construction – there is little to choose between them in terms of confusion. Overall, neither (18) nor (19) are improvements on the 1972 version (17).

A typical example of the capability of Systran in the mid 1970s is one of the translations from 1976 used by Knowles (1979) in his 'error analysis' of the Systran system – a much longer extract will be found in the database. The translation was produced as part of a trial by the Gesellschaft für Mathematik und Datenverarbeitung (Bonn, Germany). From the original Russian text (20), the output is (21):

(20) [original Russian]: Вертолет, летательный аппарат тяжелее воздуха с вертикальными взлетом и посадкой, подъемная сила в котором создается одним или несколькими (чаще двумя) несущими винтами… Вертолет взлетает вертикально вверх без разбега и совершает вертикальную посадку без пробега, неподвижно висит над одним местом, допуская поворот вокруг вертикальной оси в любую сторону, производит полет в любом направлении со скоростями от нуля до максимальной. При вынужденной остановке двигателей в полете вертолет может совершить планирующий спуск и посадку, используя самовращение (авторотацию) несущих

винтов. Во избежание срыва потока с лопастей и для увеличения скорости полета некоторые вертолеты имеют небольшое крыло, разгружающее несущие винты.

(21) [Systran, 1976]: A helicopter, a flight vehicle heavier than air with vertical by takeoff and landing, lift in which is created one or by several (more frequent than two) rotors. .. A helicopter takes off upward vertically without a takeoff and it accomplishes vertical fitting without a path, motionlessly 'will hang' above one place, allowing rotation around a vertical axis to any side, flight in any direction at speeds is produced from zero to the maximum. With the forced engine shutdown in flight, the helicopter can complete gliding descent and fitting, using autorotation (autorotation) of rotors. To avoid flow separation from blades and for an increase in the velocity of flight some helicopters have the small wing, which unloads rotors.

In the first sentence, a copula *is* should have been inserted – not, however, an easy problem when translating from Russian, where often a dash or comma indicates equation. The presence of *by* should have ideally been eliminated, and *lift in which* would be more comprehensible as *where the lift...* The parenthesised *more frequent than two* should be *more often two*. More serious are the translations in the second sentence of посадку as *fitting* (instead of *landing*) and of без разбега as *without a path (*instead of *without a runway).* The repeated *fitting* in the third sentence might still be confusing, but in general the translation serves well as a rough version.

When commenting upon the quality of the Systran translations, Knowles (1979) identified areas where improvements should be possible without major additional research. In his view, the failures of Systran were due not to inadequate system architecture and data processing capacities but to deficiencies in its language data – and these could be readily improved in time. He was, therefore, convinced that Systran output could be improved greatly. His 'error analysis' identified the main faults as: missing words, incorrect article use, incorrect prepositions, and incorrect word order in English. How then does the current Systran system compare?

(22) [Systran]: Helicopter, aerodyne with the vertical by the takeoff and landing, lift in which is created by one or several (more frequent than two) rotors… Helicopter takes off vertically upward without the takeoff/run-up and accomplishes vertical fitting/landing without the path/range, motionlessly it will hang over one place, allowing rotation around the vertical axis to any side, flight in any direction with speeds of from zero to maximum is produced. With the emergency stopping of engines in flight the helicopter can complete gliding descent and fitting/landing, using autorotation (autorotation) of rotors. To avoid flow separation from the blades/vanes and for an increase in the velocity of flight some helicopters have the small wing, which unloads the rotors.

The use of the uncommon *aerodyne* in place of the earlier *flight vehicle* is a surprise. As in 1976, the first sentence lacks a copula *is*, it retains the intrusive *by* and it still has *lift in which* and *more frequent than two.* There are however two improvements in the second sentence where alternative translations are offered: *run-up* as well as *take-off* and *landing* as well as *fitting* (in this case *landing* correct). Unfortunately neither of the two alternatives for без разбега (*path*/*range*) are correct. A step backward is the placing of *is produced* at the end of this sentence where in the earlier version it was placed correctly. However, with the rest of the paragraph there is a definite improvement.

PROMT produces the following:

(23) [Promt]: The helicopter, the flying device is heavier than air with vertical rise and landing(planting), elevating force in which is created by one or several (is more often to two) bearing(carrying) screws … the Helicopter flies up vertically upwards without start and makes vertical landing(planting) without run, motionlessly hangs above one place, supposing turn around of a vertical axis in any party(side), makes flight in any direction with speeds from zero up to maximal. At the compelled stop of engines in flight the helicopter can make planning descent(release) and landing(planting), using self-rotation (autorotation) of bearing(carrying) screws. In order to prevent failure of a stream from blades and for

increase of speed of flight some helicopters have the small wing unloading bearing(carrying) screws.

In the first sentence, *flying device* is better than *aerodyne* (but not than the earlier *flying vehicle*), however the lacking (or misplaced) *is* gives a confusing structure. The use of *screws* instead of *rotors* is an unhappy choice; likewise, of *rise* instead of *takeoff*; and of *without start* instead of *without take-off*. On the other hand *landing without run* in (23) is better than the puzzling *landing without path* (retained essentially in the current Systran (22).). But, in the same sentence, PROMT is worse with *supposing turn around of a vertical axis* instead of *allowing rotation around vertical axis* in both old and current Systran versions (21) and (22). Likewise, *at the compelled stop of engines* in PROMT is worse than Ssytran's *with the emergency stopping of engines* (21) and *with the forced engine shutdown* (22).

In general, the current Systran is marginally better than the 1976 Systran version, but PROMT is perhaps not. What is clear is that we cannot point to the substantial improvements in Russian-to-English MT which Knowles would probably have expected to have occurred in the last 25 years.

3. English to French

After the Russian-to-English and English-to-Russian systems of the 1950s and 1960s, the earliest working MT systems were for translation from English into French. These were the domain-specific Météo systems developed by the TAUM team (University of Montréal) for the Canadian weather service and put into public operation first in 1977, and the Systran system acquired by the European Commission in 1976.

3.1. Météo 1976

The Météo system should probably not be compared directly with current general-purpose MT systems for English-to-French, as it was designed specifically for the 'sublanguage' of meteorological forecasts. However, texts of weather reports and forecasts are obvious candidates for Internet translation services. It is, therefore, interesting to see how well systems such as Systran and other online MT services can deal with texts in this domain.

An early (1976) example of Météo translation (Bruderer 1978b: 128) is:

(24) [original English] Halifax and vicinity. South Shore. Valley. Mostly sunny becoming cloudy in the afternoon. Strong gusty northeasterly winds 30 to 35. Clear overnight with a few fog patches near dawn. Thursday increasing cloudiness with a chance of rain in the evening.

(25) [Météo, 1976] Halifax et banlieue. Côte sud. Vallée d'Annapolis.
Généralement ensoleillé devenant nuageux dans l'après midi. Strong gusty northeasterly winds 30 to 35. Clair durant la nuit avec quelques bancs de brouillard vers l'aube. Jeudi nebulosité croissante avec possibilité de pluie dans la soirée.

The example illustrates the practice of Météo to retain the original sentence if any word in it cannot be found in the lexicon.

The output from current Systran in its Babelfish version is as follows:

(26) [Systran]: Halifax et proximité. Rivage Du sud. Vallée. La plupart du temps devenir ensoleillé nuageux l'après-midi. Venteux fort enroule northeasterly 30 à 35. Dégagez durant la nuit avec quelques pièces rapportées de brouillard près de l'aube. Opacité croissante de jeudi avec une chance de pluie en soirée.

For the fifth sentence/phrase the Systran at Lycos has a slight difference:

(26a) Le northeasterly venteux fort enroule 30 à 35.

Since it is clear that Systran does not have an entry for *northeasterly* (as apparently neither did Météo in 1976) an amended version was submitted with hyphenated *north-easterly*, for which Babelfish and Lycos produced, respectively:

(26b) Vents nord-est venteux forts 30 à 35.

(26c) Le nord venteux fort enroule vers l'est 30 à 35.

It is not clear why there should be a difference and why Lycos should fail when Babelfish succeeded. In both cases, however, the translation of *clear* as a verb and of patches as non-meterological *pièces* produced confusing output (26, sixth sentence)

However, both Systran versions were better than the FreeTranslation, a frequently used online service:

(27) [FreeTranslation]: Halifax et le voisinage. Rive du sud. Vallée. Devenir surtout ensoleillé nuageux de l'après-midi. Fort de grand vent nord pâques remonte 30 à 35. Clair du jour au lendemain avec les pièces quelques brouillard approche de l'aube. Jeudi augmenter cloudiness avec un hasard de pluie dans le soir.

Failure to identify *easterly* as an adjective for a wind (rather than a calendar date: *Easter - pâques*), and the absence of *cloudiness* from the dictionary contribute further confusion (despite the success with *clear*)

.

3.2. Systran 1976

In 1976, the European Commission's experience with Systran was still at an early stage. From the text of an English abstract (28), it produced the French version (29).

(28) [original English]: The biochemistry and reaction kinetics of foods, and their use in practice with the help of the Chamber of Technology. The kinetics of changes in various foods during processing are briefly discussed, with references to chemical, organoleptic, enzymic, microbiological and physicochemical changes. Applications of the results of studies on reaction kinetics to selection and modification of processing conditions is briefly discussed.

(29) [Systran, 1976]: Le kinetics biochemistry et réaction des produits alimentaires, et leur utilisation dans la pratique avec l'aide du salle de la technologie. Le kinetics des changements dans les différents produits alimentaires pendant la transformation est brièvement discuté, au sujet du produit chimique, et et changements microbiologiques et physicochemical enzymatiques organoleptic. L'application des résultats des études sur le kinetics de la réaction au choix et la modification des conditions de transformation est brièvement discutée.

At this early developmental stage, the dictionary lacked entries for *kinetics*, *biochemistry*, *physicochemical* and *organoleptic*, leading to obvious failures of translation. It failed to translate *chamber* as *chambre* in this context. Current versions from Systran are definitely improvements (30), although there is still a failure to recognise *reaction cinetics* as a noun phrase in the first phrase – with consequent distortions in the placement of *cinétique*. (Although it gets it right in the last sentence.)

(30) [Babelfish, Lycos, and Systran Personal] De biochimie et de réaction des nourritures, et leur utilisation la cinétique dans la pratique avec l'aide de la chambre de la technologie. La cinétique des changements de diverses nourritures pendant le traitement est brièvement discutée, avec des références au produit chimique, les changements organoleptiques, enzymatiques, microbiologiques et physico-chimiques. Des applications des résultats des études sur la cinétique de réaction au choix et à la modification des états de traitement est brièvement discutées.

FreeTranslation (31) has similar problems with the noun phrase *reaction kinetics*, but more seriously the absence of some words (although readily understood in isolation by French readers) causes mistranslations of sentences. The translation of *processing conditions* as *traitement de conditions* is apparently the result of a wrong analysis of the phrase out of context.

(31) [FreeTranslation]: La biochimie et la réaction cinétiques de nourritures, et leur usage dans la pratique à l'aide de la Chambre de Technologie. Le cinétique de changements dans les diverses nourritures pendant le traitement est brièvement discuté, avec les références à chimique, organoleptic, enzymic, les changements de microbiological et physicochemical.

Les applications des résultats d'études sur la réaction cinétique à la sélection et la modification de traitement de conditions sont brièvement discutées.

However, Systran is not always better than Free Translation, as illustrated by another example of a translation from 1978 (33), a current Systran translation (34) and a current Free Translation (35).

(32) [original English] Water information for Europe – the art and technology of communication. About half a million people are employed in the water industry in the EEC countries. The water industry is thus a major employer with wide-ranging, multidisciplinary information requirements.

(33) [Systran, 1978]: Informations d'eau pour l'Europe – l'art et la technologie de la communication. Environ des personnes de un demi-million sont employées dans l'industrie d'eau dans des pays DE LA CEE. L'industrie d'eau est ainsi un employeur important avec des besoins compréhensifs et multidisciplinaires de l'information.

(34) [Systran Personal, Lycos]: Arroser l'information pour l'Europe - l'art et la technologie de communication. Au sujet de un demi-million les gens sont employés dans l'industrie de l'eau dans les pays de la CEE. L'industrie de l'eau est ainsi un employeur important avec des conditions étendues et multidisciplinaires de l'information.

(35) [FreeTranslation]: L'information d'eau pour l'Europe – l'art et la technologie de communication. Des gens d'un demi million sont employé dans l'industrie d'eau dans les pays de EEC. L'industrie d'eau est ainsi un employeur majeur avec large étendre, les conditions d'information de multidisciplinary.

The interpretations of *water information* as a command (*arroser*) and of *about a million people* as a topic introduction are surprising in the current Systran version (34). Overall, it is worse than the 1978 result (33). By contrast, in both cases FreeTranslation is more successful (35). It also managed to generate *d'un* instead of incorrect *de un* (34). Unfortunately the absence of *multidisciplinary* in the FreeTranslation dictionary let it down in the last sentence.

4. French to English: Systran

For French-to-English translation the only operational system before the 1990s was the one installed at the Commission of the European Communities in the late 1970s.

4.1. Systran 1982

Examples of Systran in operation in the early 1980s are given by Pigott (1982) and Wagner (1985). Pigott was one of those responsible for the implementation and development of Systran at the Translation Service during the 1980s. His examples come from 1982.

(36) [original French]: Application au central de télésurveillance d'engins sur pneus. Il s'agit d'un matériel à microprocesseur qui est chargé de traiter les informations fournies par un appareillage de détection de passage d'engins sur pneus (D.P.E.) Il édite des rapports de poste et de journée indiquant la durée et l'importance relative des périodes consacrées par chaque engin surveillé aux diverses activités possibles: évacuation des produits, transport de matériel, entretien, station service … ainsi que le nombre de godets de charbon évacués.

(37) [Systran 1982]: Application to the exchange of télésurveillance of equipment on tyres. It is a matter of an equipment with microprocessor which is charged to treat the information provided by an equipment of equipment passage detection on tyres (D.P.E.) It publishes station and day reports indicating the duration and the importance relative of the periods devoted by each instrument supervised to the various possible activities: evacuation of the products, transport of equipment, maintenance, station service … as well as the number of evacuated coal cups.

Problems at this time related primarily to gaps in the lexicon. In this example (37), we find that each of *appareillage*, *engins* and *matérial* are rendered as *equipment*. Nearly all these problems have been solved in the current version of Systran (38). Even so, there is no improvement in the translation of *chargé de traiter (*which should be *charged with processing)*

(38) [Babelfish] Application to the remote monitoring station of tyred vehicles. It is about a microprocessor-based setup which is charged to process the data provided by a tyred vehicle traffic recorder (D.P.E.) It publishes reports/ratios of shift and day indicating the duration and the importance relating of the periods devoted by each machine monitored to the various possible activities: evacuation of the products, transport of material, maintenance, service station... as well as the number of evacuated buckets.

Reverso (39) is generally no worse and no better than Systran in 1982 (37). It avoids the multiple use of *equipment,* but fails to disentangle *un appareillage de détection de passage d'engins sur pneus* producing a nearly unintelligible *equipment of detection of passage of machines on tires.* However, more serious is the mistaken interpretation of *le nombre* as a imperative verb.

(39) [Reverso]: Application to the exchange of electronic surveillance of machines on tires. It is about a material with microprocessor which is in charge of treating(handling) the information supplied by an equipment of detection of passage of machines on tires (D.P.E). He(it) publishes(edits) reports(connections) of post(post office) and of day indicating the duration and the relative importance of periods dedicated by every machine watched in the different possible activities: evacuation of products, transport of material, interview(maintenance), station service as well as counts him(it) of evacuated jars(gores) of coal.

The version from Free Translation (40) is not much better. Even apart from the missing vocabulary items (*télésurveillance*, *godets*), it fails to identify *de journée* as a modifier of *rapports*, misplacing its translation of *surveillé*, with consequent garbling of sense.

(40) [Free Translation]: Application to the exchange of télésurveillance of vehicles on tires. It is a matter of an equipment to microprocesseur that is loaded to treat information furnished by a detection casting off of passage of vehicles on tires (D.P.E.) It edits post reports and of day indicating the length and relative importance of the periods consecrated by overseen every vehicle to the various possible activities: discharge of the products, equipment transportation, maintenance, station service... as well as the number of godets of evacuated coals.

4.2. Systran 1983

Further examples of Systran in the early 1980s are given by Wagner (1985). From an original (41), the EC Systran system of the time (1983) produced (42):

(41) [original French]: Où en est-on, environ un an après la réunion du Comité permanent de l'emploi? A l'époque, les organisations syndicales avaient estimé que la Commission européenne prenait le problème par le petit bout de la lorgnette, sans le situer dans le contexte d'une stratégie globale et d'une politique industrielle plus dynamique. Il faut partager le travail, réduire le temps de travail, mais ce n'est là qu'un palliatif.

(42) [Systran 1983]: How far was, approximately a year after the meeting of the Standing Committee on Employment? At the time, the trade-union organizations had felt that the Commission had taken the problem by the small end of the lorgnette, without locating it in the context of a total strategy and of a politique industrial more dynamic. It was necessary to share work, to reduce the working time, but it was there only one palliative.

Apart from one lexicon omission (*politique*), the system at this time could not deal with the common French expression *où en est...;* nor with the complexities of the less common idiom *ce n'est là que...* The current Systran systems (44) are improvements on the whole – the main exception being a mistranslation of the last sentence due to a failure to treat the comma before *réduire* as a sign of coordination. The treatment of *où en est...* is still clumsy, but it is now intelligible; however, the phrase *ce n'est là que...* is still a problem. It is, of course, not surprising that it can give no English equivalent for the metaphor *prener... par le petit bout de la lorgnette* – either now (43) or in the past (42).

(43) [Babelfish, Lycos, Systran Personal 4.0]: Where is one, approximately a year after the meeting of the Standing Committee on Employment? At the time, the trade-union organizations had estimated that the European Commission took the problem by the small

end of the spyglass, without locating it in the context of a total strategy and a more dynamic industrial policy. Work should be shared, to reduce the working time, but it is only one palliative there.

The output from Reverso On-line (44) is also better on the whole than the 1983 Systran (42). There is idiomatic rendition of *où en est...* although spoilt by the intrusion *on* before *approximately*. Like the current Systran systems it fails with the metaphor and with the phrase *ce n'est là que...*

(44) [Reverso]: Where are we, on approximately one year after the meeting of the permanent Committee of the employment(use)? In the time, unions had considered that the European Commission took the problem by the small end of the opera glasses, without placing him(it) in the context of a global strategy and a more dynamic industrial politics(policy). It is necessary to share the work, to reduce the working time, but it is there only a palliative.

However, FreeTranslation (45) is markedly inferior – a less acceptable version for *où en est...*, a mistranslation of *à l'époque,* and of *une politique industrielle plus dynamique* (presumably because it does not recognise *politique* as a noun*),* as well as (just like Systran and Reverso) a failure with *ce n'est là que:*

(45) [FreeTranslation]: Where some is one, about a year after the meeting of the permanent Committee of the employment? To the era, the union organizations had considered that the European Commission took the problem by the small end of the lorgnette, without situating it in the context of a global strategy and of a political industrial one more dynamic. It is necessary to divide the work, reduce the work time, but this is not there that a palliative one.

Another example from Wagner illustrates other problem areas:

(46) [original French]: M. le Président rappelle que le problème de la réduction du temps de travail a été étudié à la réunion de Munich. Différentes thèses s'affrontent: pour les syndicats, la réduction du temps de travail contribuera à supprimer le chômage, mais les employeurs pensent qu'elle supprimera des emplois soit en renchérissant les coûts, soit en accroissant la productivité. Il serait souhaitable de poursuivre aujourd'hui cette discussion en laissant de côté tous les présupposés idéologiques. Pour commencer, il convient de demander au représentant de la Commission européenne, qui a réussi à rejoindre Luxembourg malgré un mouvement de grève à Bruxelles, s'il souhaite compléter l'exposé qu'il avait présenté à Munich.

In 1983, the EC version of Systran produced:

(47) [Systran 1983]: The Chairman reminded the meeting that the problem of the reduction of the working time had been studied to the meeting of Munich. Various theses faced themselves: for the trade unions, the reduction of the working time would contribute to suppress unemployment, but the employers thought that she would suppress an employment either by increasing the costs, or by increasing the productivity. It would be desirable to continue today this discussion by leaving some on the side of all ideological presupposed them. To start, it was advisable to require from the European Commission representative, which had succeeded in joining Luxembourg despite a strike movement in Brussels, if it wished to supplement the statement that it had presented in Munich.

There is an awkward translation of *s'affrontent*, an understandable mistranslation of *elle* as *she*, a failure to treat the idioms *laisser de côté* and *il convient de*, to deal with *les présupposés* (*presuppositions*) and the (less tolerable) mistranslations of *qui* as *which* and *il* as *it* (twice).

Current Systran systems (48) produce a better rendition of *s'affronter*, manage to correctly translate *elle*, but still fail with *laisser de côté*, *les présupposés*, and *qui* and *il* in the final sentence. They also introduce an unnecessary *to* before *ask*. Overall, however, there is a definite improvement:

(48) [Babelfish, Lycos, Systran Personal]: Mr. the President points out that the problem of the reduction of the working time was studied with the meeting of Munich. Various theses clash: for the trade unions, the reduction of the working time will contribute to remove unemployment, but the employers think that it will remove employment either by increasing the costs, or by increasing the productivity. It would be desirable to continue this discussion today by leaving side all the presupposed ideological ones. To start, it is advisable to ask to the European Commission representative, which succeeded in joining Luxembourg in spite

of a movement of strike in Brussels, if it wishes to supplement the talk that it had presented in Munich.

By contrast, FreeTranslation (49) and Reverso (50) are no better than Systran in 1983 (47). FreeTranslation has the same mistakes, but in addition it also fails to recognise *productivité* and *idéologiques* and has a more awkward rendition of *il convient de*.

(49)[FreeTranslation]:Mr. the President recalls that the problem of the reduction of the work time was studied at the meeting of Munich. Different theses confront themselves: for the unions, the reduction of the work time will contribute to eliminate the unemployment, but the employers think that she will eliminate from employments be while increasing the costs, be while increasing the productivité. It would be desirable to follow today this discussion while leaving out all them presupposed idéologiques. To begin, it suits to request the representative of the European Commission, that succeeded rejoining Luxembourg despite a strike movement to Brussels, if it wishes to complete the exposition that it had presented to Munich.

The version from Reverso is worse in parts than either the 1983 Systran (47) or the current Systran and FreeTranslation outputs (48, 49): e.g. *reminds* instead of *recalls* or *points out*, *syndicates* instead of *unions*, *kill(abolish)* instead of *remove* or *eliminate*. Elsewhere there are much better translations, e.g. *are in confrontation*, *presuppositions*, and *who*, *he* in last sentence. Overall the result is not significantly worse than the 1983 Systran version.

(50) [Reverso]: Mr. President reminds that the problem of the reduced working week was studied in the meeting of Munich. Various theses are in confrontation: for syndicates, the reduced working week will contribute to kill(abolish) the unemployment, but the employers think that it will kill(abolish) employments(uses) either by making more expensive(by adding) the costs, or by increasing the productivity. It would be desirable to pursue this discussion today by leaving aside all the ideological presuppositions. To begin, it is advisable to ask to the representative of the European Commission, who managed to join Luxemburg in spite of a strike action in Brussels, if he wishes to complete the statement which he had presented in Munich.

The general impression is that the French-to-English translation by the Systran, FreeTranslation and Reverso systems are improvements on the translations of the early 1980s.

3.3. Systran, 1987

Over the years, the Translation Service of the European Commission (European Union, EU) has worked on various improvements of Systran, focussing specifically on the treatment of EU administrative language. Many of these changes are not reflected in the publicly available versions of the system.

In 1998 John Beaven (1998) illustrated the progress which had been made by the EU Systran in the previous ten years (not just for French-English, but also for English-to-French, English-to-German, English-to-Spanish, French-to-German, and French-to-Spanish).

As one example, the system produced from the French (51) the English output (52) in 1987 and the English output (53) in 1997.

(51) [original French]: La Commission a également adopté la directive 95/44/CE prévoyant des dérogations au régime phytosanitaire en ce qui concerne des travaux à des fins d'essais ou à des fins scientifiques ou des travaux sur les sections végétales.

(52) [1987 output] The Commission adopted also the directive 95/44/CE anticipating exemptions from the plant health rule with regard to work at ends of tests or at scientific ends or work on the vegetable selections.

(53) [1997 output] The Commission also adopted the directive 95/44/CE making exemptions from the plant health arrangement with regard to work for the purpose of tests or for scientific purposes or work on the plant selections.

The most obvious improvement is the correct translation of *fins* in the last sentence, and it is not perhaps surprising that the current commercial Systran program

(Babelfish and Systran Personal 4.0) still has difficulties with this usage – in the second case interpreting it as an adjective. It has also unexpected difficulty with the gerundive, inserting an unneeded *of*:

> (54) [current Systran]: The Commission also adopted directive 95/44/CE envisaging of the exemptions from the plant health mode with regard to work at ends of tests or fine scientists or work on the vegetable sections.

Another short example from Beaven's article illustrates the treatment of the French subjunctive.

> (55) [original French]: Aucun progrès ne peut cependant plus être accompli avant que le Parlement ne se soit prononcé.
> (56) [1987 output]: No progress can however no longer be accomplished before Parliament decided.
> (57) [1997 output]: No progress can however be made before parliament has decided.
> (58) [current Systran]: No progress can be accomplished however more before the Parliament did not decide.

There is a marked improvement for the EU's own Systran system, but it is equally clear that the commercial Systran today (58) has not incorporated the change and consequently is worse than the EU version in 1987 (55).

5. German to English

Although MT research on German began during the first decades of activity (e.g. at Massachusetts Institue of Technology and the University of Texas), operational systems for the language did not appear until the 1980s, and commercial systems not until the early 1990s. Some of the earliest examples of German-to-English MT are therefore primarily from pre-prototype systems.

5.1. Systran 1976

Bruderer (1978: 298-299) includes an early example of translation by Systran in about 1976, evidently distributed to demonstrate the system's capabilities. The results (60) were remarkably good:

> (59) [original German]: Das vielleicht bedeutendste Problem betrifft die Geschwindigkeit, für die die Concorde ausgelegt ist. Eng damit verbunden ist die Wahl der aerodynamischen Formgebung und der Werkstoffe. Die grundlegenden Unterschiede zwischen der Concorde und anderen Verkehrsflugzeugen sind in vielfacher Hinsicht nur gering. Der Kern des Problems ist vielmehr, ob dieses Flugzeug hinsichtlich Geschwindigkeit, Konstruktionsmethoden und Gewinnpotential einen logischen Schritt in der Entwicklung darstellt. Obwohl die Entwurfsmannschaften zu wiederholten Malen um ausführliche Angaben über Nutzlastkapazität, Reichweite und andere Leistungsmerkmale ersucht wurden, haben sie sich bisher nur zu recht allgemeinen Auskünften verstehen können.
> (60) [Systran, 1976]: The perhaps most important problem concerns the speed, for which the Concorde is designed. The choice of the airframe structure and of the materials is closely connected with that. The basic differences between the Concorde and other commercial planes are only small in many respects. The heart of the problem is rather, if this airplane represents with regard to speed, design methods and profit potential a logical step in the development. Although the design teams were asked for explicit information on payload capacity, range and other performance features repeatedly, they were able to agree up to now only to quite general information.

Almost all that could be criticised in this translation is the placement of the direct object noun phrase (*a logical step in the development*). By contrast, the current Systran German-English system displays a number of inaccuracies:

> (61) [Babelfish, Lycos] The problem perhaps most meaning concerns the speed, for which the Concorde is appropriate. Connected with it the choice of the aerodynamic shaping and the materials is close. The fundamental differences between the Concorde and other airliners are only small in multiple regard. The core of the problem is rather whether this airplane represents a logical step in the development regarding speed, structural design methods and profit potential. Although the draft crews were asked to repeated marks for

detailed data concerning payload capacity, range and other capability characteristics, them so far only rightfully general information to understand to have been able itself.

The adjectival *vielleich bedeutendste* is unaccountably placed after its noun (*problem*). In the next sentence, the phrase (*eng damit verbunden – closely connected with it*) is incorrectly divided: *damit verbunden* (*connected with it)* and *eng ... ist* (*is close)* so that the adverb *eng* is rendered as a verb form*,* and, compounding the error, placed at the end of the sentence. Other mistakes are the rendition of *Entwurfsmannschaften* (*design teams*) as *draft teams*, and most seriously of all the completely unintelligible output of the last clause. In fact, this clause defeats all current systems.

(62) [FreeTranslation]: The perhaps most important problem concerns the speed for which the Concorde is displayed. Closely therewith the choice of the aerodynamic Formgebung and the materials connected. The basic differences between the Concorde and other wrong airplanes are only slight in multiple regard. The kernel of the problem is rather, whether this airplane with respect to speed, construction methods and profit potential a logical stepped under development represents. Although the design teams were requested to repeated paintes around detailed details about payload capacity, range and other service features, they were able to understand itself previously only to quite general information.

(63) [Reverso]: The perhaps most important problem concerns the speed for which Concord is laid out. Is closely connected with it the choice (election) of the aerodynamic design and the materials. The fundamental differences between Concord and other airliners are only low in multiple respect. The core of the problem is rather whether this aeroplane represents a logical step in the development with regard to speed, construction methods and profit potential. Although the draft teams were asked repeatedly of detailed information about pay load capacity, range and other performance characteristic features, up to now they could get on only according to law to general information.

However, both FreeTranslation and Reverso are better than Systran (61) in certain respects: both interpret *vielleicht bedeutendste* correctly. Reverso, in particular, is superior to Systran in the third and fourth sentences (while still misinterpreting *zu recht* – here as *to law*), and in the second its performance is marred only by the sentence-initial placing of *is closely connected*.

5.2 METAL 1975

Bruderer gives also an example of translation by the METAL system, still under development at the University of Texas (Bruderer 1978: 260ff). The output is clearly at an early phase of the project (65), but nevertheless still indicative of relatively good-quality performance.

(64) [original German]: Als Rückstossmotor kann jede Maschine bezeichnet werden, bei welcher ein Gas oder eine Flüssigkeit unter Druck mit hoher Geschwindigkeit aus einer Öffnung ausströmt und dadurch eine Reaktions- oder Rückstosskraft hervorruft. Zur Veranschaulichung dieses Vorganges diene folgendes Beispiel. In einem allseitig geschlossenen Zylinder befindet sich ein Gas unter Druck. Das Gas wird sich in dem Behälter gleichmässig verteilen und der Druck ist an allen Teilen der Innenfläche gleich gross. Die Druckwirkung lässt sich so definieren, dass das Gas aus einzelnen Molekülen besteht, welche in Wirklichkeit so klein sind, dass wir sie mit blossem Auge nicht sehen können. Wir wollen sie uns aber als Kugeln vorstellen.

(65) [METAL, 1976]: Any/each/every machine/machinery where a gas or a liquid flows out from an opening under pressure with high velocity and thereby causes a reaction or recoil force/power can be called a jet propulsion motor. The following example may serve for the illustration of this process. A gas is under pressure in a cylinder closed on all sides. The gas will spread/distribute in the container uniformly and the pressure is equally great/high/extensive at all parts of the inner surface. The pressure effect can be defined in such a way that the gas consists of individual molecules which are so small in reality that we cannot see them with the naked eye. But we want to imagine them as balls/globes/spheres.

Despite the presence of alternative translations of certain words (a sign that disambiguation rules were not fully developed), the output (65) can be readily understood. By contrast, the current Systran German-English system (66) has problems of word order in the first sentence (*recoil engine* should come after *called*, *causes* should come after *thus, an opening* should come after *leaks out* and be preceded by *from*), an incorrect ver form (*serve*) in the second, a failure to interpret *gleich* as an adverbial modifier of gross in the third sentence (i.e. *equally large*), and the mistranslation of the reflexive *vorstellen* as *present... to us* rather than *imagine*. Nevertheless, although overall poorer than METAL in 1976 (65), particularly the first sentence, the results are probably intelligible by experts in the field.

(66) [Systran]: Recoil engine each machine can be called, with which a gas or a liquid leaks out under pressure with high speed an opening and thus a reaction or a recoil force causes. For the illustration of this procedure serve the following example. In an generally closed cylinder is a gas under pressure. The gas will distribute itself evenly in the container and the pressure is equivalent large at all parts of the inner surface. The pressure effect can be defined in such a way that the gas consists of individual molecules, which are so small in reality that we cannot see them with the naked eye. We want to present it to us however as balls.

5.3. Logos 1983/84

Probably more typical of the state of German-English MT in the 1980s is the output from the Logos system; an example from the installation at SAP in 1983 is given by Tschira (1985):

(67) [original German]: Im Geschäftsleben werden die Geschäftsvorfälle regelmässig durch Belege dokumentiert, wie etwa Buchungsbelege, Aufträge, Bestellungen. Die in ihnen enthaltenen Daten werden – in Gegensatz zu den Stammdaten – als "Bewegungsdaten" bezeichnet.

In herkömmlichen EDV-Anwendungs-Systemen wurden oft die Daten eines Geschäftsvorfalls getrennt bearbeitet und gespeichert. In den SAP-Systemen vermeiden wir diese unnatürliche Aufteilung und speichern einen Geschäftsvorfall statt dessen einheitlich als sog. "Dokument", um dem Sachbearbeiter jederzeit einen Überblick über den gesamten Vorfall zu ermöglichen.

(68) [Logos 1983]: In the business-life, the business transactions are documented regularly by document/vouchers, as for example posting documents, orders, orders. The data included in them are designated as movement data – in contrast to the master data -.

The data of a business transaction were often processed and stored separately in conventional EDP-application-S-systems. In the SAP-systems, we avoid this unnatural breakdown and store a business transaction instead of which than sucked uniform. At any time, to make possible document, about the person in charge an overview of the entire occurrence.

It may be noticed in the first paragraph that Logos could not distinguish between *Aufträge* and *Bestellungen*, and that it misplaced the embedded phrase. In the second paragraph, Logos fails to recognise the genitive link -s-, the standard phrase *statt dessen*, but most seriously of all it fails to recognise *sog.* as an abbreviation of *sogenannt* (*so-called*), with a consequential nonsensical translation as *sucked* and a mis-parsing of the whole sentence.

However, current Systran systems (69) are no better at dealing with *sog.*, and in addition they badly distort the second sentence.

(69) [Babelfish, Lycos, Systran Personal 4.0] In the business life the business transactions are documented regularly by vouchers, as for instance vouchers, orders, orders. The data contained in them become - in contrast to master data - when "designates transaction data".

In conventional EDP application systems the data of a business transaction were often worked on and stored separately. In the SAP systems we avoid this unnatural allocation and store a business transaction instead of its uniformly as if sucked. "document", in order to at any time make for the specialist possible an overview of the entire incident.

Other current systems do succeed in dealing with *sog* (*sogennant*), but they fail elsewhere. Reverso (70) does not recognise EDV-Anwendung and like Logos (68) fails to make the link to systems. More seriously *getrennt* is interpreted as a full verb and not a past participle adverb (*separately*), and the persistent translation of *Vorfall* as *incident* rather than *transaction* would probably mislead. However, it does recognise *statt dessen* correctly as adverbial *instead*.

(70) [Reverso]: In the business life the commercial incidents are documented regularly by vouchers, approximately like vouchers, orders, orders. The data contained in them are called - in contrast to the master data - "movement data".

In conventional EDV Anwendungs to systems the data of a commercial incident were treated often separated and were stored. In the SAP systems we avoid this unnatural subdivision and store a commercial incident instead uniformly as so-called. "Document" to make possible a view about the whole incident for the specialist at any time.

With the Personal Translator (71), succeeds with sog., but this success is greatly outweighed by other major errors, of which perhaps the morphologically odd *finishedly* and *storedly* are the most immediately obvious – Systran (69) provides the correct interpretation.

(71) [Personal Translator PT]: In the business life the business transactions of regelmässig are documented by pieces of evidence such as booking pieces of evidence, orders, orders. The data contained in them are described as "movement data" in contrast for the master data.

The data of a business transaction were finishedly and storedly often differentiated in conventional EDP application systems. In the SAP systems we avoid this unnatural division and store a business transaction instead of this one uniformly as a so-called "document" to make a summary of the complete occurrence possible for the clerk any time.

The FreeTranslation version (72) also succeeds with sog., but unfortunately (as Logos and Systran) treats the point as an end of sentence marker, with consequential misparsing of the remainder. In addition, there is again a misleading interpretations of *Vorfall*, *getrennt*, and *statt dessen*, and now also a mistaken translation of *Stammdaten*.

(72) [FreeTranslation]: In the business life, the business occurrences are documented regularly through records, as well as about entry record, orders, orders. The data contained in them are designated – in contrast to the tribe data – as “movement data”.

In conventional EDP-uses-systems, the data of a business occurrence were separated worked and stored often. In the SAP systems, we avoid this unnatural distribution and store a business occurrence instead of its uniformly as so-called. “Document”in order to enable the expert always an overview over the total occurrence.

Overall, the current systems do not show an advance on the Logos system of the mid 1980s. In places, there are improvements, but elsewhere there are steps backward.

Another example from Logos of the same date (73, 74) illustrates further problems of German-English translation (Ordish 1984).

(73) [original Geman]: Lassen Sie mich an dieser Stelle sehr dankbar die Beteiligung spanischsprechender Länder an unseren internationalen klinischen Prüfungsprogramm vermerken.

Mexico, oben bereits zweimal erwähnt, ist seit Jahren erfolgreich und aus dem Kreis der Länder für frühe klinische Prüfungen nicht wegzudenken. Die Zuwendung der Medizin und die verständnisvolle Unterstützung durch die zuständige Geschäftsführung und Landesleitung verdient besonders hervorgehoben zu werden.

(74) [Logos, 1983] Please allow me to be very gratefully recorded at our international clinical trial programme at countries Spanish-speaking to/? this place the participation.

Mexico, already twice mentioned above, is successful and unimaginable from of the circle of the countries for early clinical checks for years. Allowance of medicine and the understanding support through the appropriate business management and country management earned to be emphasized especially.

Most of the translation of the first sentence is an incoherent muddle, primarily through misinterpretation of the phrase starting with *Beteiligung* (*participation*). The current Systran is a definite improvement (75):

(75) [Babelfish, Lycos, Systran Personal] Let me here very gratefully the participation of Spanish-speaking countries to our international clinical test program note.

Mexico, above already twice mentioned, is not to be excluded for years successful and from the circle of the countries for early clinical examinations. The allowance of the medicine and the understanding support by the responsible management and national line to be earned particularly emphasized.

Nevertheless, Systran has problems with word order (placing verbs at the end of sentences: *note*), failing to interpret *ist...erfolgreich* as an independent clause (*has been successful*) and therefore confusingly misplacing (*not to be excluded)* before rather than after *for years*. The translation of *Zuwendung* as *allowance* (as Logos did also) rather than *allocation* or *grant* might lead to misunderstanding.

FreeTranslation (76) makes much the same mistakes, except that fortunately, even though it has no entry for *wegzudenken,* it succeeds with *ist erfolgreich*. Better overall, perhaps, for this example passage is the output from Reverso (77), the main faults being: *let note me*, *at the top* (instead of *above*), *not to imagine*, *allowance* (instead of *grant*), and *deserving* (rather than *deserve*).

(76) [FreeTranslation] Let me here very gratefully the participation Spanish language of countries in our international clinical test programs note.

Mexico, above already twice mentions, has been for years successful and out of the circle of the countries for early clinical tests not wegzudenken. The grant of the medicine and the sympathetic support through the responsible management and country direction earned especially emphasized to become.

(77) [Reverso] Let note me at this point very gratefully the participation of Spanish-speaking countries(lands) in our international clinical test program.

Mexico, at the top already two times mentioned, is successful for years and not to imagine from the circle of the countries(lands) for early clinical tests. The allowance of the medicine and the understanding support by the responsible management and regional administration deserving particularly to be emphasized.

In general, however, none of the current German-English translation systems produce better results for this example than Logos in the mid 1980s, indeed in some cases Logos was superior.

5.4. METAL 1992

One of the first installations of the METAL system developed by Siemens from initial research at the Univesity of Texas was at the German software company SAP. Examples of translations at the time are found in Grasmick and Elliott (1992). In the one reproduced (79), the first paragraph and the beginning of the second are acceptable; but the system has clearly problems with the second sentence of the second paragraph. The coordinate construction is confused (*faced with the costs of testing* inserted into *to keep as small as possible... losses*), the *um...zu* construction is incorrected interpreted as *around*, and there is a failure to place *is... advisable* after *an agreement.*

(78) [original German] Wenn in diesen Richtlinien nicht anders angegeben, sind die Prüfungen an einem einzigen Prüfling, der alle zutreffenden Prüfungen bestehen muss, vorzunehmen. Der Prüfling muss einem verkaufsfähigen Gerät entsprechen oder ein für den Kunden versandfertiges Gerät sein.

Wenn eine Prüfung nach diesen Richtlinien zu Zerstörungen führen kann, darf eine Nachbildung verwendet werden, um die zu prüfende Bedingung darzustellen. Angesichts des Prüfaufwandes und um Verluste möglichst gering zu halten, ist eine Abstimmung zwischen allen Beteiligten hinsichtlich Prüfprogramm, Prüflingen und Prüffolge ratsam.

(79) [METAL 1992] If not specified differently in these guidelines, the tests are to be done on an single test item which must pass all valid tests. The test item must correspond to a salable device or be a device ready for shipment to the customer.
If a test according to these guidelines can lead to destructions, a copy may be used in order to represent the condition to be tested. To keep as small as possible faced with the costs of testing and around losses, is an agreement between all participants with regard to test program, test items and test sequence advisable.

The output from Systran (80) is less successful in the first sentence. There is misplacement of *not* from the *if*-clause to the main clause (with consequential misunderstandingby readers), there is a failure to move the verb (*make*) from sentence-final position and to change it into a passive form (i.e. *the examinations are to be made only on one test specimen*), and there is a mistaken translation of *bestehen* as *exist* rather than *pass*. By contrast the second paragraph is more successful, although there is still the failure to interpret the *um...zu* construction (*in order to*) and there is the unsatisfactory translation of *Abstimmung* as *tuning* instead of *agreement*.

(80) [Babelfish, Systran Personal] If indicated in these guidelines differently, are not the examinations of only one test specimen, which must exist all applicable examinations to make. The test specimen must correspond to saleable equipment or be equipment ready for dispatch for the customer.
If an examination can lead after these guidelines to destruction, a reproduction may be used, in order to represent the condition which can be examined. In view of the inspection expenditure and around losses to keep as small as possible, a tuning between all involved one is advisable regarding test program, test specimens and testing sequence.

Overall, the current Systran is not inferior to METAL in 1992. However, FreeTranslation (81) and Reverso (82) are definitely inferior. Even so, Reverso does better than Systran (80) by getting the correct word order in the last sentence (*is advisable*).

(81) [FreeTranslation] If in these guidelines differently indicated, are not to be undertaken the tests at an only examinee, that must exist all correct tests. The examinee must correspond to a sale capable device or be a device delivery finished for the customer.
If a test can lead after these guidelines to destructions, a copy may be used in order to represent the condition to be tested. In view of the test expense and in order to hold losses if possible slightly, is a voting between all participant with respect to test program, examinee and test sequence advisably.

(82) [Reverso] If in these directives in a different way given(indicated), the tests are not in a single test specimen who must insist all appropriate tests on deciding. The test specimen must correspond to a device capable of sales or be a device ready for dispatch for the customer.
If a test can lead by these directives to devastations, an imitation may be applied to represent the condition to be checked. In view of the test expenditure and in order to hold(regard) losses as low as possible, a vote(coordination) is advisable between all partners with regard to test program, to test specimens and test consequence.

The general impression from these examples is that there has been no progress in German-English systems since the early 1990s.

5.5. Globalink 1990

Finally there is the question whether there are any discernible improvements in systems intended for the general public ('home use' systems), running on persoanl computers. In 1990, an MT evaluation group at Essex University made an asessment of the translation quality and operational performance of the commercial Globalink GTS system (University of Essex 1991). Included were evaluations of German to English translation (the group also evaluated English-to-German, English-to-Spanish and Spanish-to-English). Two extracts of results will be illustrated. The first is paassage from the introductory guide to the system itself:

(83) [original German]: Computerprogramme, die Übersetzungen erzeugen, sind ein nützliches Werkzeug und Hilfsmittel, stellen jedoch kein völlig automatisiertes Verfahren dar, das die Rolle des Übersetzers übernehmen könnte. Das Globalink Translation System kann also den hochqualifizierten Übersetzer nicht ersetzen. Allerdings steigert GTS die Produktivität eines Übersetzers, der an langen, fachsprachlichen Texten arbeitet, verkürzt seine Arbeitszeit und verbessert die Genauigkeit seiner Übersetzung. Das System dient ebenfalls der zeitsparenden Anfertigung von Übersetzungsentwürfen, wenn es lediglich darum geht, den allgemeinen Sinn eines Schriftstückes zu erfassen.

(84) [Globalink GTS, 1990] Computer programs, the translations produce, an useful tool and resource are, present however no fully automated process, which could take over the role of the translator. The Globalink @@Translation system can not replace therefore the highly qualified translator. Indeed GTS raises the productivity of a translator, who shortens at long, @@fachsprachlichen texts works, his work time and improves the precision of his translation. The system serves likewise the time-saving @@Anfertigung of draft translations, if it goes solely about it, to record the general sense of a document.

There are many mistakes in the Globalink translation (84); most obvious are missing lexical items, the retention of German verb-final word orders (*produce, are, works*), failure to identify a relative pronoun (…*programs, the translations produce*… instead of …*which produce translations*…), and the displacement of *shortens* from its rightful position before *his work time*.

Current systems are improvements. Overall, however, the best results are achieved with FreeTranslation (85) and Reverso.

(85) [FreeTranslation] The computer programs which produce translations are useful tools and aid, however, represent no completely automatical procedure which could take over the roll(role) of the translator. The Globalink translation system cannot replace for the highly-qualified translator. But GTS increases the productivity of a translator who works on long, technical texts, his(its) working hours shorten and improve the exactness of his(its) translation. The system also serves the timesaving preparation of translation drafts if it is a matter only of grasping(registering) the general sense of a document.

There is an awkward rendition in the second sentence (the intrusive *for* after *replace*), and in the third sentence (garbled completely by Globalink) there is only a misplacement of *shorten* after *working hours.* The result from Reverso is identical, both lexically and grammatically.

By contrast, the current Systran translation (86) is less successful. In particular, it has still problems with the analysis of the relative clause in the third sentence (*der an langen, fachsprachlichen Texten arbeitet*, i.e. *who works on technical texts*), including a mysterious translation of *an langen* as *enough on*, and (as in Globalink) verb-final placement. The same sentence illustrates also the familiar and long-standing problem for all MT systems in the interpretation of anaphors (*seine*, *seiner* translated as *its* instead of *his*).

(86) [Babelfish, Systran Personal 4.0] Computer programs, which produce translations, are a useful tool and aid, represent however no completely automated procedure, which could take over the role of the translator. The Globalink translation system cannot replace thus the highly-qualified translator. However GTS increases the productivity of a translator, that is enough on, specializedlinguistic texts works, shortens its work time and improves the accuracy of its translation. The system serves likewise the time-saving preparation of translation drafts, if it concerns only to seize the general sense of a document.

Overall, there are definite improvements in FreeTranslation and Reverso (and to less extent in Systran) over the output from Globalink in 1990 (84).

A second example text from the Globalink GTS system is an extract from an official letter to a student:

(87) [original German]: Wir haben Ihren Antrag auf einem Platz für den oben genannten Kurs bearbeitet und ich freue mich, Ihnen mitteilen zu dürfen, dass Sie für das Studienkolleg-Programm des am 1. Oktober 1990 beginnenden Studienjahres zugelassen sind. Sind Ihre Leistungen während des Studienkollegs zufriedenstellend, können Sie im

> Oktober 1991 ein dreijähriges Studium mit dem Abschluss “Bachelor’s degree” im Bereich der Wirtschaftswissenschaft beginnen.

For this Globalink (88) produced almost complete nonsense:

> (88) [Globalink GTS, 1990] We have processed Your proposition on a place for the above named course and I please me, You communicate to be allowed to, that You are admitted for the @@Studienkolleg – program the at the 1. October 1990 beginning study year. If Your performances Are during the study course of lectures satisfactory, can begin You in the October 1991 a three-year study with the conclusion “Bachelor’s @@degree” in the area of the economy sciences.

The current Systran version (89) is an improvement, even despite the surprising corruption *are pleased I* (instead of *I am pleased*…), the inappropriate *certified* (instead of *admitted*) and the misplaced *beginning* in the first sentence.

> (89) [Babelfish, Systran Personal 4.0] We worked on your request on a place for the course specified above and are pleased I to be allowed to communicate to you that you are certified for the preparatory course program on 1 October 1990 of the academic year beginning. If your achievements are satisfying during the preparatory course, you can begin a three-year study in October 1991 with the conclusion "Bachelor's degree" within the field of the economic science.

The FreeTranslation and Reverso outputs (90, 91) are also better than Globalink, and on the whole better than Systran. Although in the first sentences there are similar (but different) confused sentence structures *(pleased me I*…), both systems (unlike Systran) succeed in correctly interpreting the pre-nominal construction *des am 1. Oktober 1990 beginnenden Studienjahres,* placing *beginning* immediately after *studies year*. In addition*,* Reverso also produces the correct *admitted* (where Systran has *certified*, and FreeTranslation *allowed*). However, FreeTranslation fails in the second sentence to recognise the German verb-intial construction (*Sind Ihre Leistungen*…) as an *if*-clause, which must lead to misinterpreation by readers. But while Reverso (like Systran) gets this right, it manages to produce an adverb (*satisfyingly)* instead of an adjective (*satisfying*, or better *satisfactory*) and the mistranslation *three-year-old* instead of *three-year*.

> (90) [FreeTranslation] we worked your proposition on a place for the course named above and please me I to be allowed to announce to you that you are allowed for that studies lecture program of the studies year beginning on the 1 October 1990. Your achievements are during the studies lecture satisfactorily, you can begin in October 1991 a three-year study with the termination "Bachelor's it degree" in the area of the economy science.
> (91) [Reverso] We have treated your application(motion) for a place(square) for the course named at the top and I am glad to be allowed to inform you that you are admitted for the study lecture-program of the academic year beginning on the 1st October, 1990. If your performances(achievements) are during the study lecture satisfyingly, you can begin in October, 1991 a three-year-old study with the conclusion " Bachelor's degree " in the area of the economic science.

While Systran, FreeTranslation and particularly Reverso are definite improvements on the Globalink output of the early 1990s, Personal Translator PT (92) is not, the first sentence is the worst of all.

> (92) [Personal Translator PT] Being allowed to we have edited your application on a place for the course mentioned above and inform to you I am happy that you are allowed for the studies lecture program of the academic year starting on October 1st, 1990. If your performances are satisfactory during the studies lecture, you can start with three-year studies with the degree "Bachelor" of degree in the area of economics in October 1991.

Although all these systems do represent advances on systems for ‘home use’ in the early 1990s, the persistent failures are surely damaging to MT as a whole, since it is very likely that these systems have been purchased and/or used (in the case of online systems) for the translation of precisely this kind of text for individuals with little or no knowledge of the source languages.

6. Summary and discussion of results

Table 1 summarises the observations given above, and adds assessments of quality based on additional longer examples in the database elsewhere on this website. Results from systems which were prototypes (pre-installation), 'demonstration' models, or customised, are *italicised*, and should generally be disregarded when making overall judgements about progress.

Judgement is complicated by two major factors: (a) the absence of examples of translations (with their source texts) for systems throughout the period, e.g. there are no Russian examples from the 1980s; and (b) the variable quality of current systems, not just between systems but also from one text type to another within the same system. To a large degree, individuals will look at these examples and come to different conclusions. Nevertheless, it does seem to me that in all these language pairs there has been definite progress since the mid 1960s and probably since the early 1970s. What is more uncertain is whether and in which languages there have been improvements since the early 1980s. The general quality of Russian-to-English translation seems to have scarcely improved at all. For French-to-English there has been a gradual improvement in translation quality since the mid 1980s, but perhaps not during the last decade. For German, however, it does seem that overall there have been some very gradual improvements over since the 1980s and to a limited extent in the 1990s (for 'home use' systems), although it must be said that the general quality is still some way behind that found in translations from French.

Table 1

	Systran	Promt	Reverso	FreeT	PersT
Russian-English					
IBM-USAF 1959	B	B			
IBM-USAF 1964	*W*				
GU 1967	B	B			
Systran 1972	*W*	*W*			
Systran 1976	B	E			
English-French					
Meteo 1976	*W*		*W*		
Systran 1976	E		E		
French-English					
Systran 1982	B		E	B	
Systran 1985	B		B	B	
Systran 1987	W				
German-English					
Systran 1976	*W*		*W**	*W**	
METAL 1975	*W*		*W**	*W**	*WW**
Logos 1983	E		E	E	W
Logos 1984	E		E	E	W
METAL 1992	E		W	W	W*
Globalink 1990	B		B	B	E

Key:

Systran = Babelfish, Lycos, or Systran Personal 4.0; Promt = PROMT-Online or @promt Standard; Reverso = Reverso-Online; FreeT = FreeTranslation; PersT = Personal Translator PT. [For details of systems see *Compendium of translation software* on www.eamt.org]

B = better; E = equally good/bad; W = worse; WW = much worse

Italic = pre-operational systems; * = not illustrated in the paper

The impediments to the improvement of translation quality are the same now that they have been from the beginning. Apart from obvious failures of missing

words, disambiguation and incorrect selection of target language words, the difficulties are:

a) anaphora:
    pronouns (*it* vs. *she*/*he*),
    and definite articles (absence vs. presence – e.g. when translating from Russian and French)
b) inappropriate retention of source language structures:
    verb-initial constructions (from Russian (e.g. (10) *Are investigated*…) and from German (e.g. (63) *Is closely connected*…)),
    verb-final placements (translations from German),
    non-English pre-nominal participle constructions in both Russian (e.g. (4, 5) *with interest to be read materials)* and German (e.g. (89) *on 1 October 1990 of the academic year beginning*)
c) mistaken coordination (e.g. (70-72))
d) mistakes in morphology:
    single subject nouns with plural (or un-conjugated) verb forms,
    wrong verb formations (e.g. (7) *it can translating*…),
    adjectives not agreeing with nouns,
    wrongly formed adverbs, e.g. (71) *finishedly* and *storedly*)
e) word order mistakes (e.g. (61) *the problem perhaps most meaning*…)
f) prepositions:
    wrong choice, or absent,
    or inappropriately inserted (e.g. (22) *with the vertical by the take-off*),
    or doubled (e.g. (7) … *on because*…; (22) … *of from*…)
g) verb forms: lacking copula (from Russian), active instead of passive
h) general difficulties with any multi-clause sentence
i) embedded parentheses (e.g. (68-70))

Here is not the place to suggest how these problems might be solved, or even how the worse cases might be avoided, but it does seem strange after nearly 50 years of MT research that some systems still produce incorrect morphological formations, that some fail with noun-verb agreement, that some fail to adhere to English adjective-noun order, and that verbs are placed incorrectly at the beginnings or at the ends of sentences and clauses. Such apparently easily resolvable errors must undoubtedly contribute to the dissatisfaction of many users of MT systems.

It would seem plausible that output errors could be reduced by the use of a 'language model' such as those proposed in statistics-based and example-based approaches to MT (e.g. Knight & Chander 1994). A language model reflecting the most frequent two-word and three-word collocations in the target language could well deal with some of the problems with prepositions, with articles, with some of the word order mistakes, and probably with most of the morphological errors.

7. Conclusions

This paper has not been concerned with the improvability of any particular system – or any particular type of system architecture. Whether a system is inherenetly unimprovable or whether it is capable of improvement is a complex question involving system architecture, rule bases, dictionary entries, etc. In the 1980s it was often claimed that systems of the "first generation" with 'direct translation' architecture were inherently limited in potential translation quality, and that the later "second generation" systems of the 'transfer' type were more likely to produce good

quality output. This investigation can throw no light on this question – which is now, in any case, overtaken by the reorientation of MT research to data-driven and corpus-based mrthods. It should be noted, however, that advocates of some of these new methods now make similar claims about the inherent limitations of rule-based methods and the superiority of statistics-based and example-based approaches. Time will tell whether systems based on these newer approaches will perform better than those which have been investigated here.

These comments have been informal assessments of what has or has not been improved, they have not been well-founded comparative evaluations, although the examples may provide others with the data for making more rigorous assessments. Such exercises are needed if the MT community is to have some basis for claiming (or implying) that real progress has been achieved in recent years. Without them what measures can we provide the general public and what reasons have we for continuing to work in the dark?

What should always be stressed is that linguistic quality alone should not be the sole criterion for using MT systems. What is often more important is usability and usefulness. If MT is being used as the basis for translations of good quality (i.e. in dissemination of information, publicity, etc.) then what matters is whether improvements through post-editing and/or controlled language input can be achieved cost-effectively. If MT output is being used for assimilation purposes then what matters is whether the user can extract more information from a poor translation than from the original text – if he/she does not know the source language at all or only poorly, then even a very poor translation can be useful. Those who know the source language well may not need a translation at all – and should certainly not judge MT as if it were a human translation.

It should also be remembered that poor quality software is not unique to MT – while there is a market for low grade products there will be manufacturers to sell them. In order to have more informaed purchasers there is a need, long recognised but not yet achieved, for readily available and easily understood benchmarks and for authoritative and publicised ‘consumer evaluations’. This paper has not been concerned with the issue of poor quality current systems, but with the question whether there are now commercial (or operational) and/or online systems capable of producing ‘raw’ output which is of better quality overall than older operational systems. The conclusions have been tentative and uncertain, and so more investigation is desirable.

The scarcity of translation examples from older systems is an issue which ought to be addressed by the MT community if it is to have reliable records of its past achievements and failures. But it is not only the output that should be preserved, the systems and programs themselves should be kept – for two main reasons: a) to evaluate their performance as translating systems, and b) to find out how their programs worked in detail. As MT has become more commercial, the need for archiving has become more pressing. One message from this paper is, therefore, that steps should be taken now, before it is too late, to preserve old superseded MT software in an archive for future researchers.

8. References

ALPAC (1966): *Languages and machines: computers in translation and linguistics*. A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council. (Publication 1416.) Washington, D.C.: National Academy of Sciences, National Research Council.

Beaven, John (1988): 'MT: 10 years of development', *Terminologie et Traduction* (1988), 1: 242-256.

Bruderer. Herbert E. (1978a): *Handbuch der maschinellen und maschinenunter-stützten Sprachübersetzung: automatische Übersetzung natürlicher Sprachen und mehrsprachige Terminologiedatenbanken*. München: Vlg. Dokumentation Saur, 1978.

Bruderer, Herbert E. (1978b): *Sprache, Technik, Kybernetik: Aufsätze zur Sprachwissenschaft, maschinellen Sprachverarbeitung, künstlichen Intelligenz und Computerkunst.* Münsingen bei Bern: Vlg. Linguistik, 1978.

Dostert, Léon E. (1955): 'The Georgetown-I.B.M. experiment', in: Locke, William N. and A. Donald Booth (eds.) Machine translation of languages: fourteen essays (Cambridge, Mass.: Technology Press of the Massachusetts Institute of Technology, 1955), 124-135.

Grasmick, Daniel & Elliott, Keith H. (1992) 'Working with machine translation', *Society for Technical Communication, [Proceedings of] 39th Annual Conference, May 10-13, 1992, Atlanta, Georgia. 1992*; p.100-103. [Appendices of METAL translations.]

Garvin, Paul L. (1967): 'The Georgetown-IBM Experiment of 1954: an evaluation in retrospect', in: *Papers in linguistics in honor of Léon Dostert* (The Hague: Mouton), 46-67. Repr. in: Garvin, P.L. *On machine translation: selected papers* (The Hague: Mouton, 1972), 51-65.

Henisz-Dostert, Bozena 'Experimental machine translation', in: Austin, William A. (ed.) *Papers in honor of Léon Dostert* (The Hague: Mouton, 1967), 57-91.

Hutchins, W.John (1986): *Machine translation:past, present, future*. Chichester: Ellis Horwood; New York: Halsted Press.

Hutchins, W.John (2001): 'Gilbert W. King and the IBM-USAF Translator', in: Hutchins, W.John (ed.) *Early years in machine translation: memoirs and biographies of pioneers* (Amsterdam: Benjamins), 171-176.

Hutchins, W.John (2003): 'The development and use of machine translation systems and computer-based translsation tools'. *International Journal of Translation* 15 (1), 5-26

Knight, Kevin & Chander, Ishwar (1994): 'Automated post-editing of documents', *Proceedings of the Twelveth National Conference on Artificial Intelligence* July 31 – August 4, 1994, Seattle, Washington. (Menlo Park: AAAI Press), 779-784.

Knowles, F. (1979): 'Error analysis of Systran output – a suggested criterion for the 'internal' evaluation of translation quality and a possible corrective for system design', in: Snell, Barbara M. (ed.) *Translating and the Computer: proceedings of a seminar, london, 14th November 1978* (Amsterdam: North-Holland), 109-133.

Ordish, Robert (1984): 'Personal impressions of a machine translation system', *The Incorporated Linguist* 23 (3), 136-145.

Perschke, Sergei (1968): 'Machine translation – the second phase of development', *Endeavour* 27, 97-100.

Pigott, Ian M. (1982): 'The importance of feedback from translators in the development of high-quality machine translation', in: Lawson, Veronica (ed.) *Practical experience of machine translation: Proceedings of a conference, London, 5-6 November 1981*. (Amsterdam: North-Holland, 1982), 61-73.

Reifler, Erwin, ed. (1960): *Linguistic and engineering studies in the automatic translation of scientific Russian into English*. Technical report, phase II (contract AF 30(602)-1827). Seattle: University of Washington Press.

Richens, R.H. & Booth, A.D. (1955): 'Some methods of mechanized translation', in: Locke, William N. & Booth, A.Donald (eds.) *Machine translation of languages: fourteen essays*. (Cambridge, Mass.: The Technology Press of the Massachusetts Institute of Technology), 24-46.

Tschira, Klaus E. (1985): 'Looking back at a year of German-English MT with Logos', in: Lawson, Veronica (ed.) *Tools of the trade: Translating and the computer 5*... Proceedings of a conference… 10-11 November 1983, The London Press Centre (London: Aslib, 1985), 215-235.

University of Essex (1991): *Assessing a PC-based commercial MT system*. Essex MT Evaluation Group. Colchester: University of Essex Department of Language and Linguistics, January 30, 1991. (Working Papers in Language Processing, no.19)

Wagner, Elizabeth (1985): 'Rapid post-editing of Systran', in: Lawson, Veronica (ed.) *Tools for the trade: Translating and the computer 5.* Proceedings of a conference… 10-11 November 1983, The London Press Centre (London: Aslib, 1985), 199-213.

Zarechnak, Michael (2000): 'The early days of GAT-SLC', in: Hutchins, W.John (ed.) *Early years in machine translation: memoirs and biographies of pioneers* (Amsterdam: Benjamins), 111-128.