Automating Language

Sizing up the Japanese machine-translation market.

by Rose Lockwood

Rose Lockwood

The contrast between machine translation (MT) "East and West" has been highlighted recently by the published results of a Japanese MT market survey carried out earlier this year by the Asia-Pacific Association for Machine Translation (AAMT).[1] Using questionnaires and interviews, an audit company collected data from a target list of 44 suppliers of MT systems sold in Japan. The systems covered by the survey included both packaged products, which are bought standalone, and preinstalled (or "bundled") systems that are sold OEM, principally by computer manufacturers.

The study categorized products by four groups of language pairs: English to Japanese, Japanese to English, bidirectional J/E, and systems which translate to/from Japanese and languages other than English (principally, Chinese and Korean). Out of the 44 MT suppliers surveyed, 19 reported their sales.

The results of the survey highlight the strong impact the Internet has had on the shape and direction of the MT market in Japan, where the Internet boom was triggered by the release of Japanese Windows 95. At that time few Web sites were available in Japanese, creating the first real surge in the market for gisting translation for Web sites (i.e., translation purely for the purpose of understanding, rather than for publication), to enable Japanese users to access and read pages written in English. The strength of this market was confirmed in 1996 when the first vendor to release a Web-browsing translation product (Nova with NetSurfer) became the market leader.

Nova entered a market crowded with MT products. In a 1996 study by Equipe, we identified 56 different products available for English-to-Japanese alone. The proliferation of products came about partly because of massive investment in MT R&D, funded jointly by MITI and major Japanese corporations in the 1980s. By the early 1990s almost all the major Japanese electronics firms had developed MT engines, many of which were being launched as products. With few exceptions (such as Fujitsu's Atlas product, which was marketed in Europe because of Fujitsu's acquisition of British computer manufacturer ICL), the market for Japanese MT was more or less exclusively in Japan, and the products ran on various proprietary Japanese platforms.

With the success of Nova's product, most of these developers began to release PC and/or Web-oriented versions of their MT software. Toshiba's Astransac product was released in Windows and Web versions, and other major manufacturers followed suit by relaunching older engines: Hitachi's HiCats, NEC's Pivot E/J, Oki's Pensee, Sharp's Power E/J, etc. As these names imply, another significant trend was well

Revenue from Japanese MT Systems
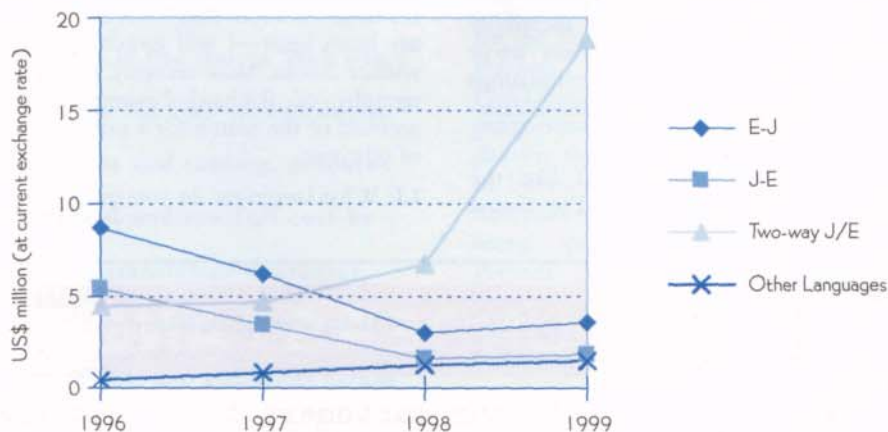
Source: AAMT Survey, 1999

Unit Sales of MT in Japan: Packaged vs. Bundled
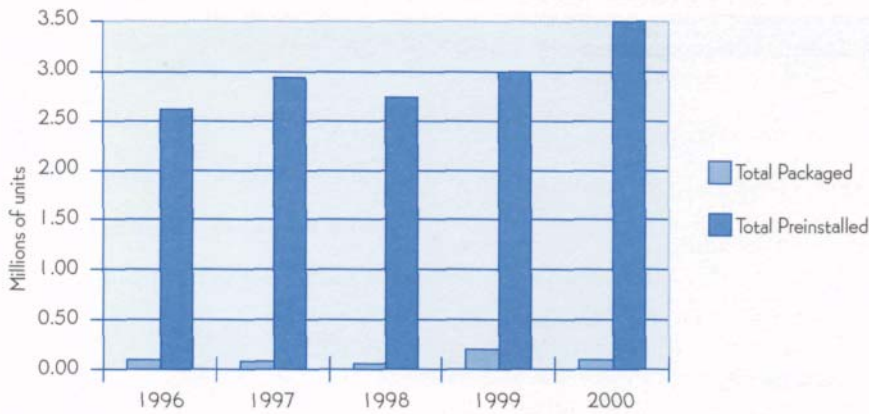
Source: AAMT Survey, 1999

under way by 1996—namely the decline in fortunes for single-direction language pairs and the rise of fully bidirectional products. At more or less the same time, a similar development was taking place at the low end of the US market for products such as Globalink, which generally abandoned single-direction products, and sold bidirectional systems for the major European languages paired with English.

The AAMT survey reports sales figures for 19 of the 44 identified suppliers, showing clearly that bidirectional Japanese/English products are the most widely sold; those vendors will gain nearly US$20 million (at current US$/yen exchange rates) in revenue from two-way J/E this year.

We know that the actual market for Japanese MT products is greater than reported in this survey, since fewer than half the vendors responded. Even a conservative guess, however, would yield a market size of US$40 to US$50 million for Japan.

The growth in demand for bidirectional J/E products has coincided with another powerful trend in the Japanese market: the move toward bundling MT systems with computer hardware—especially PCs, laptops, and notebooks—as preinstalled applications. Here the dynamics have been quite different from those in the US and Europe, where widespread "retail" access to MT has been recently driven by Internet access to MT, most notably the Systran system. In spite of a slight faltering during the 1998 economic crisis, unit sales of MT systems have steadily risen since 1995, with over 3.5 million unit sales projected for next year.

With so many products in the market, it is not surprising that price competition has driven the cost of these retail Japanese MT systems to a minimum, especially when they are bundled with hardware products. As in North America and Europe, most

MT vendors in Japan entered the market with high-end, extremely expensive systems. This makes it difficult to track falling prices over the decade, since market-dominant systems today are so different from their predecessors. Although Japanese MT prices were already falling before the impact of Web products (having declined, on average, from around US$6,500 in 1990 to less than US$3,000 in 1992), once the low-end preinstalled systems came on stream, prices dropped off the horizon. Today the average revenue earned from a preinstalled MT system in Japan is under US$5.

The fortunes of packaged low-end MT products have been quite similar in Japan, North America, and Europe, with prices driven down under US$100 and products now being sold in mainstream retail software channels. But MT markets in North America and Europe on one hand, and Japan on the other, have diverged sharply

in response to the Web. MT packaged with personal-productivity tools such as Web browsers has been widely adopted in Japan. The unique linguistic position of the Japanese as a leading economic power sharing no linguistic traditions with its principal partners/competitors in North America and Europe has created an unusually robust market. In a sense, markets needing European-language MT have had a tougher R&D challenge—to provide translation between many different language communities, and a very large number of potential language pairs and directions.

The differences between the regions are stark. In NA/Europe there have been many individual products available for years that handle single-language pairs (generally, but not exclusively paired with English), but only one or two products capable of handling the multilanguage complex which characterizes the NA/Eu-

Japanese MT: What Will Drive Future Sales

Opinion of Japanese MT Manufacturers (percent of respondents)
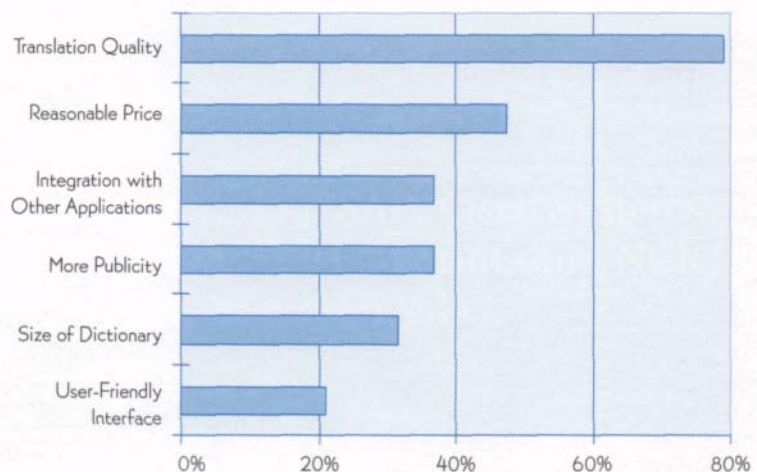
Source: AAMT Survey, 1999

ropean market. Until recently, the only significant requirement was for Japanese/English in Japan and this need was met by a large number of products doing basically the same job. For European-language markets, there are many jobs to be done, and fewer products to do them. Lernout and Hauspie have bought a significant portion of the commercial MT industry. This acquisition campaign is specifically aimed at overcoming the "language-poor" position of most MT vendors, by consolidating a number of products covering a wide spectrum of languages, and using that platform, over time, as a development base for vastly expanded linguistic coverage.[2]

Curiously, the market impact has been similar even though the forces driving developments in different regions vary. The "market value" of Web gisting MT is being redefined as we speak, and the business model for MT is radically transformed. End users increasingly expect gisting MT to be essentially free of charge. So who pays? In Japan where the hardware vendor may well own the MT software, he absorbs the cost of MT as a value-add to drive sales of computers. Vendors of European-language MT are now looking for revenue from service models—such as portal advertising, MT integrated with human translation services, etc.

The market for MT in Japan continues to be closely linked to the fortunes of the Web, though not in the same way as in NA/Europe. In the last two years, as Asian-language sites have proliferated, there has been increasing interest in MT between Japanese and languages such as Chinese and Korean. By contrast, most work on Japanese paired with European languages other than English seems to be driven from Europe—presumably because English works as an effective "pivot" language, since so many European sites are bilingual with English. Growth in the take-up of systems for languages other than English is extremely strong—well in excess of 50 percent per year—though from a very low base. Since translation quality is considered the most significant market driver, at least by the manufacturers surveyed, this trend may also indicate that the quality of Japanese/Chinese and Japanese/Korean products is generally high, given the linguistic similarity of source and target languages in these systems.

In the future, Japanese Internet users will need a broader portfolio of language pairs in order to access the increasingly multilingual Web. As this need evolves, and as the linguistic diversity of the Web matures,

Japanese customers will move toward a service model for MT, where publishers or intermediaries provide a wide range of translation-automation options in which MT to and from many languages is free.
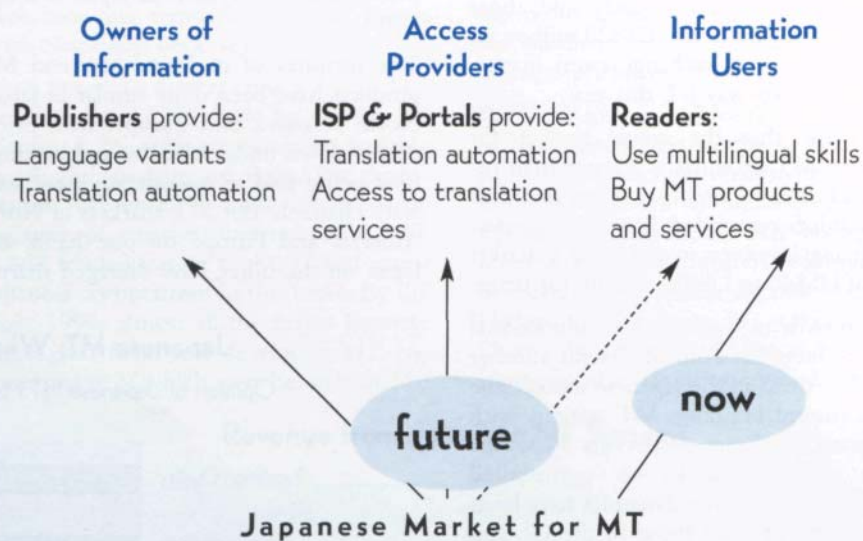
Notes

[1] Results of the study are published in the *Proceedings of the MT Summit VII*, Asia-Pacific Association for Machine Translation (AAMT), Singapore, 1999.

[2] "A Compendium of Translation Software" has been compiled by John Hutchins, president of the IAMT, and is available in draft (PDF format) at www.eamt.org/archive.

Equipe is a consortium of analysts and consultants specializing in the language industries. Based in Cambridge (UK), Equipe has associates throughout Europe, and regularly tracks markets, applications, and opportunities for language technology and information engineering. Contact Rose Lockwood at rose.lockwood@equipe.co.uk

End users increasingly expect gisting MT to be essentially free of charge. So who pays? In Japan where the hardware vendor may well own the MT software, he absorbs the cost of MT as a value-add to drive sales of computers. Vendors of European-language MT are now looking for revenue from service models—such as portal advertising, and MT integrated with human translation services.

Sizing Up the Japanese MT Market

Owners of Information	Access Providers	Information Users
Publishers provide: Language variants Translation automation	**ISP & Portals** provide: Translation automation Access to translation services	**Readers**: Use multilingual skills Buy MT products and services

future

now

Japanese Market for MT

Unique circumstances in the Japanese market—high penetration of PCs at Internet take-off, combined with early dominance of English on the Web and weak English-language skills—have created an unusually robust market for Web-oriented retail MT systems. As the Internet becomes more multilingual, Japanese users will need access to more languages. Will they continue to buy products aimed at individual target languages such as English? Probably not. It is likely that multilingual services will gradually replace the market for "client" products such as MT embedded in Web browsers. These services may be delivered by the intermediaries of the Web—ISPs, and especially the portals which will emerge serving specific communities of interest, such as e-procurement or other forms of e-commerce. Or the services may be offered by publishers of information, as an alternative to full translation into large numbers of language variants. Either way, the future for MT manufacturers serving the Japanese market will be very different from the past.