

Article development led by [acmqueue](http://queue.acm.org)  
queue.acm.org

## Old questions being answered with both AI and HCI.

BY SPENCE GREEN, JEFFREY HEER,  
AND CHRISTOPHER D. MANNING

# Natural Language Translation at the Intersection of AI and HCI

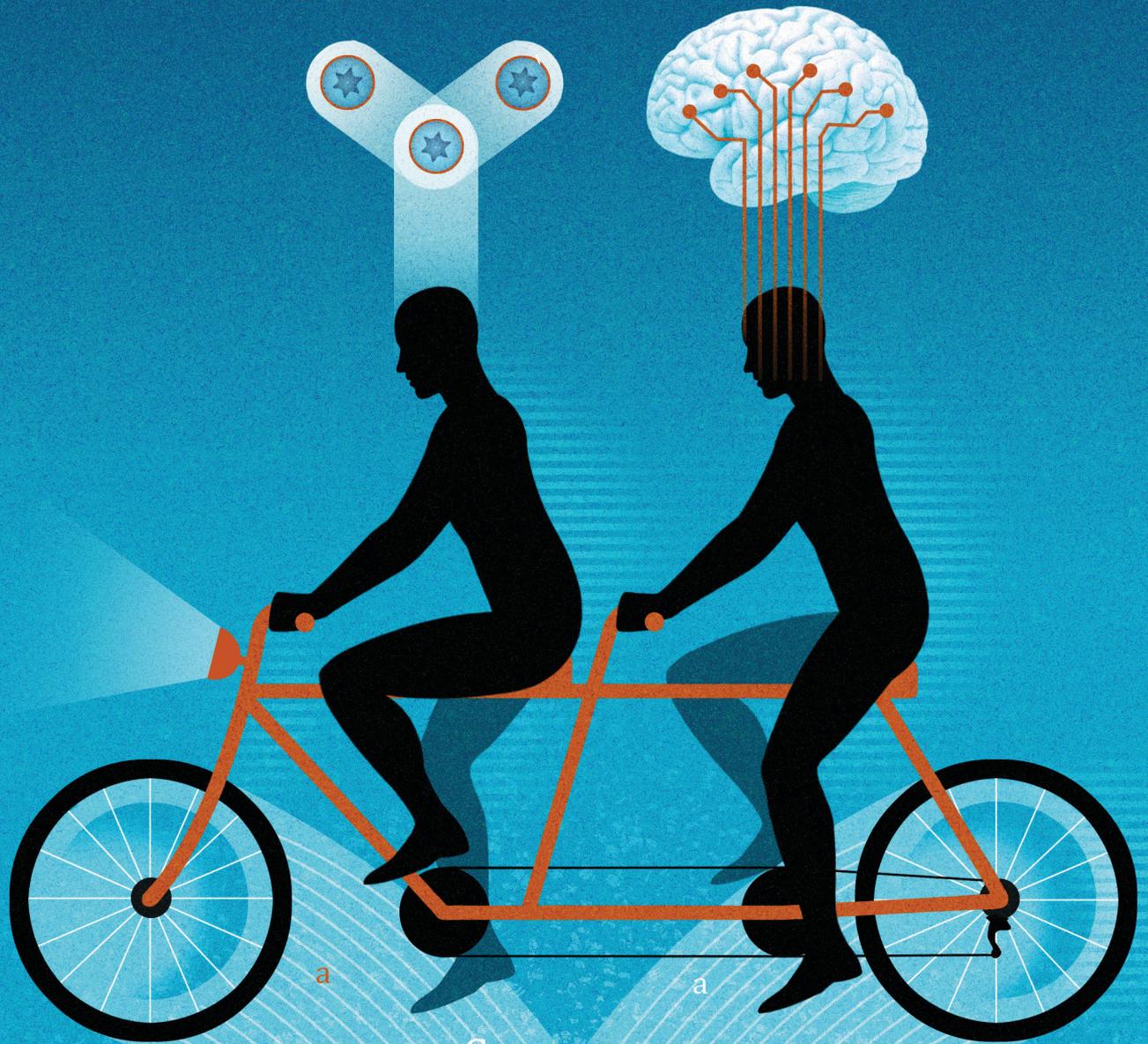
THE FIELDS OF artificial intelligence (AI) and human-computer interaction (HCI) are influencing each other like never before. Widely used systems such as Google Translate, Facebook Graph Search, and RelateIQ hide the complexity of large-scale AI systems behind intuitive interfaces. But relations were not always so auspicious. The two fields emerged at different points in the history of computer science, with different influences, ambitions, and attendant biases. AI aimed to construct

a rival, and perhaps a successor, to the human intellect. Early AI researchers such as McCarthy, Minsky, and Shannon were mathematicians by training, so theorem-proving and formal models were attractive research directions. In contrast, HCI focused more on empirical approaches to usability and human factors, both of which generally aim to make machines more useful to humans. Many attendees at the first CHI conference in 1983 were psychologists and engineers. Presented papers had titles such as “Design Principles for Human-Computer Interfaces” and “Psychological Issues in the Use of Icons in Command Menus,” hardly appealing fare for mainstream AI researchers.

Since the 1960s, HCI has often been ascendant when setbacks in AI occurred, with successes and failures in the two fields redirecting mindshare and research funding.<sup>14</sup> Although early figures such as Allen Newell and Herbert Simon made fundamental contributions to both fields, the competition and relative lack of dialogue between AI and HCI are curious. Both fields are broadly concerned with the connection between machines and intelligent human agents. What has changed recently is the deployment and adoption of user-facing AI systems. These systems need interfaces, leading to natural meeting points between the two fields.

Nowhere is this intersection more apropos than in natural language processing (NLP). Language translation is a concrete example. In practice, professional translators use suggestions from machine aids to construct final, high-quality translations. Increasingly, human translators are incorporating the output of machine translation (MT) systems such as Google Translate into their work. But how do we go beyond simple correction of machine mistakes? Recently, research groups at Stanford, Carnegie Mellon, and the European CasmaCat consortium have been investigating a human-machine model like that shown in Figure 1.

For the English input “Fatima dipped the bread,” the baseline MT



a a

خ s wd g k h

سك ق ذ

X m क ङ

È अ t

P श B o

y ج Q

system proposes the Arabic translation غمس فاطمة الخبز, but the translation is incorrect because the main verb غمس (in red) has the masculine inflection. The user corrects the inflection by adding an affix ت, often arriving at a final translation faster than she would have on her own. The corrections also help the machine, which can update its model to produce higher-quality suggestions in future sessions. In this positive feedback loop, both humans and machines benefit, but in complementary ways. To realize this *interactive machine translation system*, both interfaces that follow HCI principles and powerful AI are required.

What is not widely known is this type of system was first envisioned in the early 1950s and developments in translation research figured significantly in the early dialogue between AI and HCI. The failed dreams of early MT researchers are not merely historical curiosities, but illustrations of how intellectual biases can marginalize pragmatic solutions, in this case a human-machine partnership for translation. As practicing AI and HCI researchers, we have found the conversation today has many of the same features, so the historical narrative can be instructive. In this article, we first recount that history. Then we summarize the recent breakthroughs in translation made possible by a healthy AI-HCI collaboration.

### A Short History of Interactive Machine Translation

Machine translation as an application for digital computers predates computational linguistics and artificial intelligence, fields of computer science within which it is now

classified. The term artificial intelligence first appeared in a call for participation for a 1956 conference at Dartmouth College organized by McCarthy, Minsky, Rochester, and Shannon. But by 1956, MT was a very active research area, with the 1954 Georgetown MT demonstration receiving widespread media coverage. The field of computational linguistics grew out of early research on machine translation. MT research was oriented toward cross-language models of linguistic structure, with parallel theoretical developments by Noam Chomsky in generative linguistics exerting some influence.<sup>21</sup>

The stimuli for MT research were the invention of the general-purpose computer during World War II and the advent of the Cold War. In an oft-cited March 1947 letter, Warren Weaver—a former mathematics professor, then director of the Natural Sciences division at the Rockefeller Foundation—asked Norbert Wiener of the Massachusetts Institute of Technology (MIT) about the possibility of computer-based translation:

*Recognizing fully...the semantic difficulties because of multiple meanings, etc., I have wondered if it were unthinkable to a computer which would translate...one naturally wonders if the problem of translation could conceivably be treated as a problem in cryptography. When I look at an article in Russian, I say "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."*

Wiener's response was skeptical and unenthusiastic, ascribing difficulty to the extensive "connotations" of language. What is seldom quoted is

Weaver's response on May 9<sup>th</sup> of that year. He suggested a distinction between the many combinatorial possibilities with a language and the smaller number that are actually used:

*It is, of course, true that Basic [English] puts multiple use on an action verb such as get. But even so, the two-word combinations such as get up, get over, get back, etc., are, in Basic, not really very numerous. Suppose we take a vocabulary of 2,000 words, and admit for good measure all the two-word combinations as if they were single words. The vocabulary is still only four million: and that is not so formidable a number to a modern computer, is it?*

("Basic English" was a controlled language, created by Charles Kay Ogden as a medium for international exchange that was in vogue at the time.)

Weaver was suggesting a distinction between theory and use that would eventually take root in the empirical revolution of the 1990s: an imperfect linguistic model could suffice given enough data. The statistical MT techniques described later are in this empirical tradition.

### Use Cases for Machine Translation

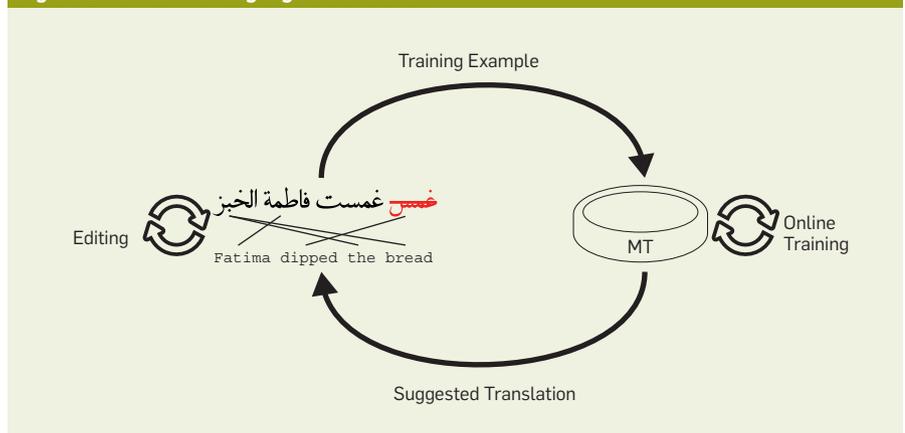
By 1951 MT research was under way, and Weaver had become a director of the National Science Foundation (NSF). An NSF grant—possibly under the influence of Weaver—funded the appointment of the Israeli philosopher Yehoshua Bar-Hillel to the MIT Research Laboratory of Electronics.<sup>19</sup> That fall Bar-Hillel toured the major American MT research sites at the University of California–Los Angeles, the RAND Corporation, U.C. Berkeley, the University of Washington, and the University of Michigan–Ann Arbor. He prepared a survey report<sup>1</sup> for presentation at the first MT conference, which he convened the following June.

That report contains two foundational ideas. First, Bar-Hillel anticipated two use cases for "mechanical translation." The first is *dissemination*:

*One of these is the urgency of having foreign language publications, mainly in the fields of science, finance, and diplomacy, translated with high accuracy and reasonable speed...<sup>1</sup>*

The dissemination case is distinguished by a desired quality threshold. The other use case is *assimilation*:

Figure 1. Interactive language translation.



*Another is the need of high-speed, though perhaps low-accuracy, scanning through the huge printed output.*<sup>1</sup>

Bar-Hillel observed the near-term achievement of “pure MT” was either unlikely or “achievable only at the price of inaccuracy.” He then argued in favor of *mixed MT*, “a translation process in which a human brain intervenes.” As for where in the pipeline this intervention should occur, Bar-Hillel recommended:

*...the human partner will have to be placed either at the beginning of the translation process or the end, perhaps at both, but preferably not somewhere in the midst of it ...*<sup>1</sup>

He then went on to define the now familiar terms *pre-editor*, for intervention prior to MT, and *post-editor* for intervention after MT. The remainder of the survey deals primarily with this pre- and post-editing, showing a pragmatic predisposition that would be fully revealed a decade later. Having established terms and distinctions still in use today, Bar-Hillel returned to Israel in 1953 and took a hiatus from MT.<sup>21</sup>

In 1958 the U.S. Office of Naval Research commissioned Bar-Hillel to conduct another survey of MT research. That October he visited research sites in the U.S. and Britain, and collected what information was publicly available on developments in the Soviet Union. A version of his subsequent report circulated in 1959, but the revision published in 1960 attracted greater attention.

Bar-Hillel’s central argument in 1960 was that preoccupation with “pure MT”—his label for what was then called fully automatic high quality translation (FAHQT)—was “unreasonable” and that despite claims of imminent success, he “could not be persuaded of their validity.” He provided an appendix with a purported proof of the impossibility of FAHQT. The proof was a sentence with multiple senses (in italics) in a simple passage that is difficult to translate without extra-linguistic knowledge (“Little John was looking for his toy box. Finally he found it. *The box was in the pen*”). Some 54 years later, Google Translate cannot translate this sentence correctly for many language pairs.

Bar-Hillel outlined two paths forward: carrying on as before, or favoring



**Increasingly, human translators are incorporating the output of machine translation systems such as Google Translate into their work. But how do we go beyond simple correction of machine mistakes?**



some “less ambitious aim.” That less ambitious aim was mixed MT:

*As soon as the aim of MT is lowered to that of high quality translation by a machine-post-editor partnership, the decisive problem becomes to determine the region of optimality in the continuum of possible divisions of labor.*<sup>2</sup>

Bar-Hillel lamented “the intention of reducing the post-editor’s part has absorbed so much of the time and energy of most workers in MT” that his 1951 proposal for mixed MT had been all but ignored. No research group escaped criticism. His conclusion presaged the verdict of the U.S. government later in the decade:

*Fully automatic, high quality translation is not a reasonable goal, not even for scientific texts. A human translator, in order to arrive at his high quality output, is often obliged to make intelligent use of extra-linguistic knowledge which sometimes has to be of considerable breadth and depth.*<sup>2</sup>

By 1966 Bar-Hillel’s pessimism was widely shared, at least among research backers in the U.S. government, which drastically reduced funding for MT research as recommended by the ALPAC report. Two passages concern post-editing, and presage the struggles that researchers in decades to come would face when supplying humans with machine suggestions. First:

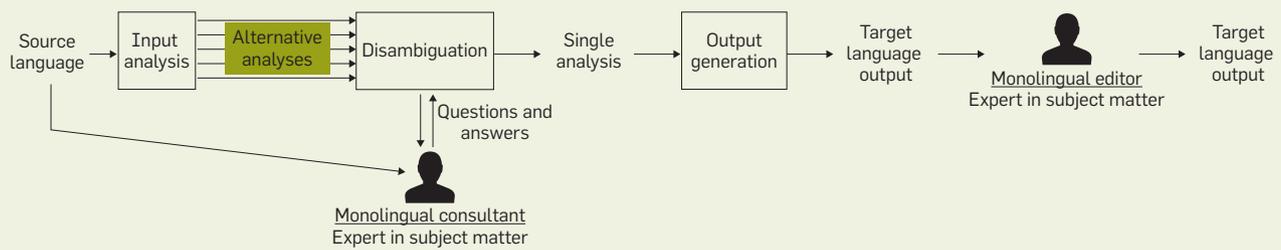
*...after 8 years of work, the Georgetown University MT project tried to produce useful output in 1962, they had to resort to post-editing. The post-edited translation took slightly longer to do and was more expensive than conventional human translation.*<sup>27</sup>

Also cited was an article by Robert Beyer of the Brown University physics department, who recounted his experience post-editing Russian-English machine translation. He said:

*I must confess that the results were most unhappy. I found that I spent at least as much time in editing as if I had carried out the entire translation from the start. Even at that, I doubt if the edited translation reads as smoothly as one which I would have started from scratch.*<sup>3</sup>

The ALPAC report concluded that two decades of research had produced systems of little practical value that did not justify the government’s level of financial commitment. Contrary to the popular belief the report ended MT research,

Figure 2. The MIND system.



it suggested constructive refocusing on “means for speeding up the human translation process” and “evaluation of the relative speed and cost of various sorts of machine-aided translation.”<sup>27</sup> These two recommendations were in line with Bar-Hillel’s earlier agenda for machine-assisted translation.

### The Proper Role of Machines

The fixation on FAHQT at the expense of mixed translation indicated a broader philosophical undercurrent in the first decade of AI research. Those promoting FAHQT were advocates—either implicitly or explicitly—of the vision that computers would eventually rival and supplant human capabilities. Nobel Laureate Herbert Simon famously wrote in 1960 that “Machines will be capable, within twenty years, of doing any work that a man can do.”<sup>29</sup> Bar-Hillel’s proposals were in the spirit of the more skeptical faction, which believed machine augmentation of existing human facilities was a more reasonable and achievable goal.

J.C.R. Licklider, who exerted considerable influence on early HCI and AI research,<sup>15</sup> laid out this position in his 1960 paper “Man-Computer Symbiosis,”<sup>24</sup> which is now recognized as a milestone in the introduction of human factors in computing. In the abstract he wrote that “in the anticipated symbiotic partnership, men will set the goals, formulate the hypotheses, determine the criteria, and perform the evaluations.” Computers would do the “routinizable work.” Citing a U.S. Air Force report that concluded it would be 20 years before AI made it possible “for machines alone to do much thinking or problem solving of military significance,” Licklider suggested that human-computer interaction research could be useful in the interim, although that interim

might be “10 [years] or 500.” Licklider and Bar-Hillel knew each other. Both participated in meetings coincident with the 1961 MIT Centennial (also present were McCarthy, Shannon, and Wiener, among others), where Bar-Hillel directly posed the question, “Do we want computers that will compete with human beings and achieve intelligent behavior autonomously, or do we want what has been called man-machine symbiosis?”<sup>16</sup> He went on to criticize the “enormous waste during the last few years” on the first course, arguing it was unwise to hope for computers that “autonomously work as well as the human brain with its billion years of evolution.” Bar-Hillel and Licklider also attended a cybernetics symposium in 1967<sup>17</sup> and a NATO workshop on information science in 1973.<sup>9</sup> The question of how much to expect from AI remained central throughout this period.

Licklider’s name does appear in the 1966 ALPAC report that advocated reduction of research funding for FAHQT. After narrating the disappointing 1962 Georgetown post-editing results, the report says two groups nonetheless intended to develop post-editing “services.” But “Dr. J.C.R. Licklider of IBM and Dr. Paul Garvin of Bunker-Ramo said they would not advise their companies to establish such a [post-editing] service.”<sup>27</sup>

The finding that post-editing translation takes as long as manual translation is evidence of an interface problem. Surely even early MT systems generated some words and phrases correctly, especially for scientific text, which is often written in a formulaic and repetitive style. The question then becomes one of human-computer interaction: how best to show suggestions to the human user.

Later, the human-machine scheme

would be most closely associated with Douglas Engelbart, who wrote a lengthy research proposal—he called it a “conceptual framework”—in 1962.<sup>11</sup> The proposal was submitted to Licklider, who was at that time director of the U.S. Advanced Research Projects Agency (ARPA). By early 1963, Licklider had funded Engelbart’s research at the Stanford Research Institute (SRI), having told a few acquaintances, “Well, he’s [Engelbart] out there in Palo Alto, so we probably can’t expect much. But he’s using the right words, so we’re sort of honor-bound to fund him.”<sup>32</sup>

“By augmenting the human intellect,” Engelbart wrote, “we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems.” Those enhanced capabilities included “more-rapid comprehension, better comprehension, ... speedier solutions, [and] better solutions.”<sup>11</sup> Later on, he described problem solving as abstract symbol manipulation, and gave an example that presaged large-scale text indexing like that done in Web crawling and statistical machine translation:

*What we found ourselves doing, when having to do any extensive digesting of journal articles, was to type large batches of the text verbatim into computer store. It is so nice to be able to tear it apart, establish our own definitions, and substitute, restructure, append notes, and so forth, in pursuit of comprehension.*<sup>11</sup>

He noted that many colleagues were already using augmented text manipulation systems, and that once a text was entered, the original reference was rarely needed. “It sits in the archives like an orange rind, with most of the real juice squeezed out.”<sup>11</sup>

## Martin Kay and the First Interactive MT System

By the late 1960s, Martin Kay and colleagues at the RAND Corporation began designing a human-machine translation system, the first incarnation of which was called MIND.<sup>5</sup> Their system (Figure 2), which was never built, included human intervention by monolingual editors during both source (syntactic) analysis and target generation (personal communication with Martin Kay, Nov. 7, 2014).

Figure 2 shows the MIND system.<sup>5</sup> Monolingual pre-editors disambiguate source analyses prior to transfer. Monolingual post-editors ensure target fluency after generation.

MIND was consistent with Bar-Hillel's 1951 plan for pre-editors and post-editors. Kay went further with a 1980 proposal for a "translator's amanuensis," which would be a "word processor [with] some simple facilities peculiar to translation."<sup>22</sup> Kay's agenda was similar in spirit to Bar-Hillel's "mixed MT" and Engelbart's human augmentation:

*I want to advocate a view of the problem in which machines are gradually, almost imperceptibly, allowed to take over...First they will take over functions not essentially related to translation. Then, little by little, they will approach translation itself.*

Kay saw three benefits of user-directed MT. First, the system—now having the user's attention—would be better able to point out uncertain translations. Second, cascading errors could be prevented since the machine would be invoked incrementally at specific points in the translation process. Third, the machine could record and learn from the interaction history. Kay advocated collaborative refinement of results: "the man and the machine are collaborating to produce not only a translation of a text but also a device whose contribution to that translation is being constantly enhanced."<sup>22</sup> These three benefits would now be recognized as core characteristics of an effective mixed-initiative system.<sup>6,18</sup>

Kay's proposal had little effect on the commercial "translator workbenches" developed and evaluated during the 1980s,<sup>20</sup> perhaps due to limited circulation of his 1980 memo (which would not be published until 1998<sup>23</sup>). However, similar ideas were being investigat-

ed at Brigham Young University as part of the Automated Language Processing (ALP) project. Started in 1971 to translate Mormon texts from English to other languages, ALP shifted emphasis in 1973 to machine-assisted translation.<sup>30</sup> The philosophy of the project was articulated by Alan Melby, who wrote that "rather than replacing human translators, computers will serve human translators."<sup>26</sup> ALP produced the Interactive Translation System (ITS), which allowed human interaction at both the source analysis and semantic transfer phases.<sup>26</sup> But Melby found that in experiments, the time spent on human interaction was "a major disappointment," because a 250-word document required about 30 minutes of interaction, which is "roughly equivalent to a first draft translation by a human translator." He drew several conclusions that were to apply to most interactive systems evaluated over the following two decades:

1. ITS did not yet aid the human translator enough to justify the engineering overhead.
2. Online interaction requires specially trained operators, further increasing overhead.
3. Most translators do not enjoy post-editing.

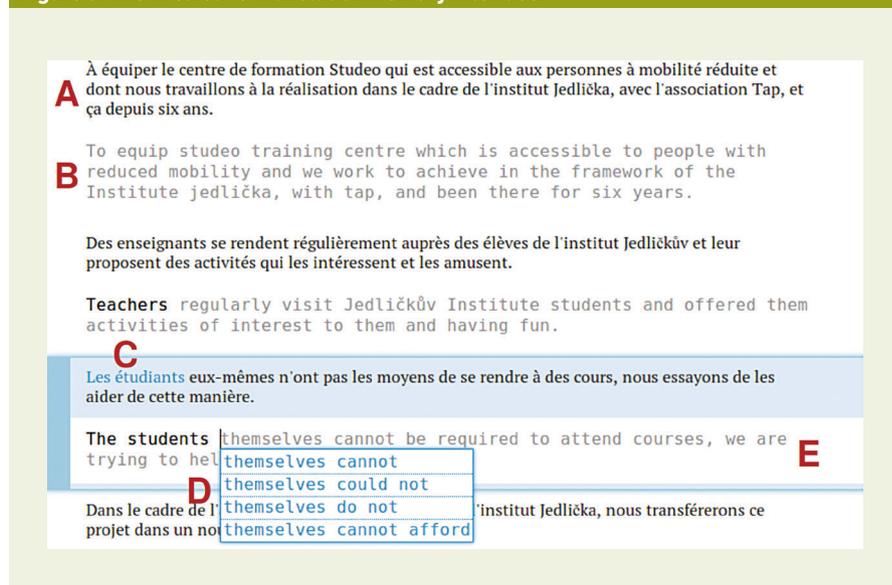
ALP never produced a production system due to "hardware costs and the amount and difficulty of human interaction."<sup>30</sup>

Kay and Melby intentionally limited the coupling between the MT system

and the user; MT was too unreliable to be a constant companion. Church and Hovy in 1993 were the first to see an application of tighter coupling,<sup>8</sup> even when MT output was "crummy." Summarizing user studies dating back to 1966, they described post-editing as an "extremely boring, tedious and unrewarding chore." Then they proposed a "superfast typewriter" with an autocomplete text prediction feature that would "fill in the rest of a partially typed word/phrase from context." A separate though related aid would be a "Cliff-note" mode in which the system would annotate source text spans with translation glosses. Both of these features were consistent with their belief that a good application of MT should "exploit the strengths of the machine and not compete with the strengths of the human." The autocomplete idea, in particular, directly influenced the TransType project,<sup>12</sup> the first interactive statistical MT system.

A conspicuous absence in the published record of interactive MT research since the 1980s is reference to the HCI literature. HCI as an organized field came about with the establishment of ACM SIGCHI in 1982 and the convening of the first CHI conference in 1983.<sup>14</sup> *The Psychology of Human-Computer Interaction*, by Card, Moran, and Newell, was also published that year.<sup>7</sup> It is now recognized as a seminal work in the field which did much to popularize the term HCI. Several chapters analyze text editing interactions, drawing con-

Figure 3. The Predictive Translation Memory interface.



clusions that apply directly to bilingual text editing, that is, translation. But we are aware of only two MT papers<sup>4,31</sup> among the thousands in the Association for Computational Linguistics Anthology (up to 2013) that cite an article included in the proceedings of CHI from 1983–2013. (There may be more, but the number is remarkably small.)

In retrospect, the connection between interactive MT and early HCI research is obvious. Kay, Melby, and Church had all conceived of interactive MT as a text editor augmented with bilingual functions. Card et al. identified text editing as “a natural starting point in the study of human-computer interaction,” and much of their book treats text editing as an HCI case study. Text editing is a “paradigmatic example” of HCI for several reasons: the interaction is rapid; the interaction becomes an unconscious extension of the user; text editors are probably the most heavily used computer programs; and text editors are representative of other interactive systems.<sup>7</sup> A user-centered approach to translation would start with text entry and seek careful bilingual interventions, increasing the level of support through user evaluation, just as Bar-Hillel and Kay suggested many decades ago.

### Recent Breakthroughs in Interactive MT

All this is not to say fruitful collaboration is absent at the intersection of AI and HCI. The landmark work of Horvitz and colleagues at Microsoft established mixed-initiative design principles that have been widely applied.<sup>18</sup> Bar-Hillel identified the need to find the “region of optimality” between human and machine; Horvitz’s principles provide design guidance (distilled from research experiences) for finding that region. New insights are appearing at major human/machine conferences such as UbiComp and HCOMP. And the explosion of data generated by companies has inspired tools such as Tableau and Tri-facta, which intelligently assist users in aggregating and visualizing large datasets. However, language applications have largely escaped notice until recently.

When we began working on mixed-initiative translation in 2012, we found



**A user-centered approach to translation would start with text entry and seek careful bilingual interventions, increasing the level of support through user evaluation, just as Bar-Hillel and Kay suggested many decades ago.**



that even post-editing had a mixed experimental record. Some studies found it increased translator productivity, while others showed the classic negative results. At CHI 2013, we presented a user study on post-editing of MT output for three different language pairs (English to Arabic, French, and German). The between-subjects design was common to HCI research yet rare in NLP, and included statistical analysis of time and quality that controlled for post-editor variability. The results showed that post-editing conclusively reduced translation time *and* increased quality for expert translators. The result may owe to controlling sources of confound overlooked in previous work, but it may also come from the rapid improvement of statistical MT, which should cause users to revisit their assumptions. For example, to avoid bias, subjects were not told that the suggestions came from Google Translate. However, one subject commented later that

*Your machine translations are far better than the ones of Google, Babel and so on. So they were helpful, but usually when handed over Google-translated material, I find it way easier and quicker to do it on my own from unaided.*

One of Horvitz’s 12 principles is that a mixed-initiative system should learn by observing the user. Recall the top of Figure 1, in which final translations are returned to the MT system for adaptation. Recent improvements in online machine learning for MT have made this old idea possible. Denkowski et al.<sup>10</sup> were the first to show users can detect a difference in quality between a baseline MT system and a refined model adapted to post-edits. The adapted suggestions required less editing and were rated higher in terms of quality than the baseline suggestions. Updating could occur in seconds rather than in the hours-long batch procedures conventionally applied.

These quantitative successes contrast with the qualitative assessment of post-editing observed in many studies: that it is a “boring and tedious chore.”<sup>8</sup> Human translators tend not to enjoy correcting sometimes fatally flawed MT output. Previously, we showed richer interactive modes have been built and evaluated, but until none improved translation time or quality rela-

tive to post-editing, a mode considered as long ago as the 1962 Georgetown experiment.

Last year we developed Predictive Translation Memory (PTM, Figure 3), which is a mixed-initiative system in which human and machine agents interactively refine translations. The initial experience is similar to post-editing—there is a suggested machine translation—but as the user begins editing, the machine generates new suggestions conditioned on user input. The translation is collaboratively refined, with responsibility, control, and turn-taking orchestrated by the user interface. The NLP innovations that make this possible are fast search and online parameter learning. The interface design is informed by Horvitz's mixed-initiative guidelines, fundamentals of graphical perception, and the CHI 2013 user study results.

In a user study with professional translators, we found that PTM was the first interactive translation system to increase translation quality relative to post-edit.<sup>13</sup> This is the desired result for the dissemination scenario in which human intervention is necessary to guarantee accuracy. Moreover, we found that PTM produced better training data for adapting the MT system to each user's style and diction. PTM records the sequence of user edits that produce the final translation. These edits explain how the user generated the translation in a machine-readable way data that has not been available previously. Our current research is investigating how to better utilize this rich data source in a large-scale setting. This is the motivation for one of Horvitz's best-known recommendations for mixed-initiative system design: minimizing the cost of poor guesses about action and timing.<sup>18</sup>

## Conclusion

We have shown a human-machine system design for language translation benefits both human users—who produce higher-quality translations—and machine agents, which can refine their models given rich feedback. Mixed-initiative MT systems were conceived as early as 1951, but the idea was marginalized due to biases in the AI research community. The new

results were obtained by combining insights from AI and HCI, two communities with similar strategic aims but surprisingly limited interaction for many decades. Other problems in NLP such as question answering and speech transcription could benefit from interactive systems not unlike the one we have proposed for translation. Significant issues to consider in the design of these systems are:

- ▶ Where to insert the human efficiently in the processing loop.
- ▶ How to maximize human utility even when machine suggestions are sometimes fatally flawed.
- ▶ How to isolate and then improve the contributions of specific interface interventions (for example, full-sentence suggestions vs. autocomplete phrases) in the task setting.

These questions were anticipated in the translation community long before AI and HCI were organized fields. New dialogue between the fields is yielding fresh approaches that apply not only to translation, but to other systems that attempt to augment and learn from the human intellect. □

## Related articles on queue.acm.org

### AI Gets a Brain

Jeff Barr and Luis Felipe Cabrera  
<http://queue.acm.org/detail.cfm?id=1142067>

### The Future of Human-Computer Interaction

John Canny  
<http://queue.acm.org/detail.cfm?id=1147530>

### A Conversation with Jeff Heer, Martin Wattenberg, and Fernanda Viégas

<http://queue.acm.org/detail.cfm?id=1744741>

## References

1. Bar-Hillel, Y. The present state of research on mechanical translation. *American Documentation* 2, 4 (1951), 229–237.
2. Bar-Hillel, Y. The present status of automatic translation of languages. *Advances in Computers* 1 (1960), 91–163.
3. Beyer, R.T. Hurdling the language barrier. *Physics Today* 18, 1 (1965), 46–52.
4. Birch, A. and Osborne, M. Reordering metrics for MT, (2011).
5. Bisbey, R. and Kay, M. The MIND translation system: a study in man-machine collaboration. Technical Report P-4786, Rand Corp, 1972.
6. Carbonell, J. AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems* 11, 4 (1970), 190–202.
7. Card, S.K., Moran, T.P. and Newell, A. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, 1983.
8. Church, K. W. and Hovy, E. Good applications for crummy machine translation. *Machine Translation* 8 (1993), 239–258.
9. Debons, A. and Cameron, W.J., Eds. *Perspectives in Information Science*, Vol. 10. NATO Advanced Study

Institutes Series. Springer, 1975.

10. Denkowski, M., Lavie, A., Lacruz, I. and Dyer, C. Real time adaptive machine translation for post-editing with cdec and TransCenter. In *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*, (2014).
11. Engelbart, D.C. *Augmenting human intellect: A conceptual framework*. Technical report, SRI Summary Report AFOSR-3223 (1962).
12. Foster, G., Langlais, P. and Lalpalmé, G. TransType: text prediction for translators. In *Proceedings of ACL Demonstrations*, (2002), 93–94.
13. Green, S., Wang, S., Chuang, J., Heer, J., Schuster, S. and Manning, C.D. Human effort and machine learnability in computer aided translation. In *Proceedings of EMNLP*, (2014), 1225–1236.
14. Grudin, J. AI and HCI: Two fields divided by a common focus. *AI Magazine* 30, 4 (2009), 48–57.
15. Grudin, J. A moving target—the evolution of human-computer interaction. *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications* (3rd edition). J.A. Jacko, Ed. CRC Press, 2012.
16. Hauben, M. and Hauben, R. *Netizens: On the History and Impact of Usenet and the Internet*. IEEE Computer Society Press, Los Alamitos, CA, 1997.
17. Hauben, R., Heinz von Foerster, Margaret Mead and JCR Licklider and the conceptual foundations for the Internet: The early concerns of cybernetics of cybernetics. Presentation in Berlin Germany (Nov. 16, 2003).
18. Horvitz, E. Principles of mixed-initiative user interfaces. In *Proceedings of CHI* (May 1999), 15–20.
19. Hutchins, J. From first conception to first demonstration: The nascent years of machine translation, 1947–1954: A chronology. *Machine Translation* 12, 3 (1997), 195–252.
20. Hutchins, J. The origins of the translator's workstation. *Machine Translation* 13 (1998), 287–307.
21. Hutchins, J. Yehoshua Bar-Hillel: A philosopher's contribution to machine translation. *Early Years in Machine Translation: Memoirs and Biographies of Pioneers*. W.J. Hutchins, Ed. John Benjamins, 2000.
22. Kay, M. The proper place of men and machines in language translation. Technical Report CSL-80-11, 1980, Xerox Palo Alto Research Center (PARC).
23. Kay, M. The proper place of men and machines in language translation. *Machine Translation* 12, 1/2 (1998), 3–23.
24. Licklider, J.C.R. Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics* HFE1 1 (1960), 4–11.
25. Melby, A.K. Creating an environment for the translator. In *Proceedings of the Third Lugano Tutorial*. M. King, ed. (Switzerland, Apr. 2–7, 1984). Edinburgh University Press, 124–132.
26. Melby, A.K., Smith, M.R. and Peterson, J. ITS: Interactive translation system. In *Proceedings of the Seventh International Conference on Computational Linguistics*. (1980).
27. Pierce, J.R. (ed.) Languages and machines: computers in translation and linguistics. National Research Council Publication 1416. National Academy of Sciences, Washington, D.C., 1966.
28. Sanchis-Trilles, G., Alabau, V., Buck, C., Carl, M., Casacuberta F. and García-Martínez, M., et al. Interactive translation prediction versus conventional post-editing in practice: A study with the CasMaCat workbench. *Machine Translation* 28, 3/4 (2014), 1–19.
29. Simon, H. A. *The New Science of Management Decision*. Harper, N.Y., 1960.
30. Stocum, J. A survey of machine translation: its history, current status, and future prospects. *Computational Linguistics* 11, 1 (1985), 1–17.
31. Somers, H. and Lovel, H. Computer-based support for patients with limited English. *EAMT Workshop on MT and Other Language Technology Tools*. (2003).
32. Waldrop, M.M. *The Dream Machine: J.C.R. Licklider and the Revolution That Made Computing Personal*. Viking, N.Y. 2001.

**Spence Green** is a co-founder of Lilt, a provider of interactive translation systems.

**Jeffrey Heer** is an associate professor of computer science and engineering at the University of Washington, where he directs the Interactive Data Lab.

**Christopher D. Manning** is a professor of computer science and linguistics at Stanford University.

© 2015 ACM 0001-0782/15/09 \$15.00