Interactions in Chinese: *Designing Interfaces for Asian Languages*

Heiko Sacher

Human–computer interfaces for Chinese-speaking users attract growing attention in the computer business. Impressive growth projections for China's economy and the rapid technology adoption by the so-called Asian Tigers make solutions for oriental languages an important part in today's design strategies.

Heiko Sacher
Interaction Designer
GVO Inc.
2370 Watson Court
Palo Alto, CA 94303
+1-650-354-0620
Fax: +1-650-493-8105
heiko@gvo.com

In 1992, a group of researchers from Apple Computer and the Institute of System Science in Singapore looked at the vision of a "computer for the rest of us" from a global perspective. They decided to study jointly the problems Asians are facing with personal computers, and they founded the Apple-ISS Research Center, Singapore.

In the following years, the team in Singapore took part in an exciting journey into the challenges of advanced interfaces for Asian users. The most tangible result of this work is the world's first commercial human–interface solution that integrates speech and handwriting for Chinese text input—the Apple Advanced Chinese Input Suite. The underlying concept of this product can be seen as a step toward a new interaction paradigm for Asian computing. The innovative user interface design was awarded with the 1997 silver medal in Business Week and Industrial Designers Society of America's Industrial Design Excellence Award. Reflections on the development of this product and some lessons learned from the design practice in Asia are discussed in this article.

Text Input: A Challenge for Chinese Computer Users

Around the globe, over 1.2-billion people speak Chinese. This includes the People's Republic of China, Taiwan, Singapore, and a large community of overseas Chinese in Asia and North America. The adoption of personal computers in the Chinese-speaking population is diverse. A closer look reveals some interesting facts:

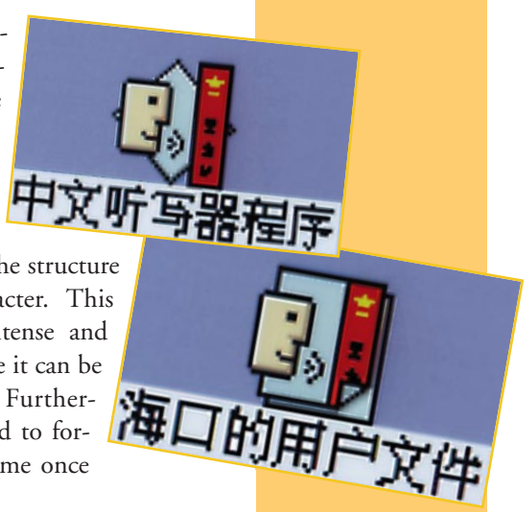
- The United States has one personal computer for every three households.
- Singapore has one personal computer for every three households.
- Taiwan has one personal computer for every 20 households.
- The People's Republic of China has one personal computer for every 100 households.

The significant difference between Singapore and Taiwan seems unexpected. Both countries share a similar economic situation and standard of living. The adoption of technologies such as telecommunication, mass media, and personal electronics is equally high. Why do Taiwanese customers make much less use of personal computers than Singapore customers?

The answer lies in the distinct language patterns of the two societies. In Singapore, English is used as the official language in public, business, and education. Basically, all Chinese Singaporeans are bilingual and speak English as their second or even first language. Taiwanese customers, however, generally speak Chinese only. Inputting Chinese characters with devices designed for Western languages is difficult to learn and use. From the variety of keyboard input methods available, two methods can be seen as emerging: "Hanyu Pinyin" and "Wubi."

The most widely used method is Hanyu Pinyin (Figure 1). It is based on a phonetic system that the People's Republic of China officially adopted in 1979 in order to represent Chinese characters in Western languages. The sound of the spoken character is interpreted in Roman script. Entering a Pinyin expression for input brings up a list of Chinese characters that share the same pronunciation. There can be up to 60 different characters that sound the same. The user has to pick the right character from a list of choices. This method involves multiple steps per character, and customers perceive it as being too slow. The other drawback is that Chinese need to learn the pinyin system first. Furthermore, it seems very peculiar that Chinese users have to describe their characters in a foreign script in order to create a Chinese text on a computer.

The Wubi or "5-Stroke Method" (Figure 2) requires the knowledge of an even more abstract encoding scheme in which a sequence of Roman characters represents the structure of the written character. This technique requires intense and lengthy training before it can be efficiently applied. Furthermore, Wubi users tend to forget the encoding scheme once



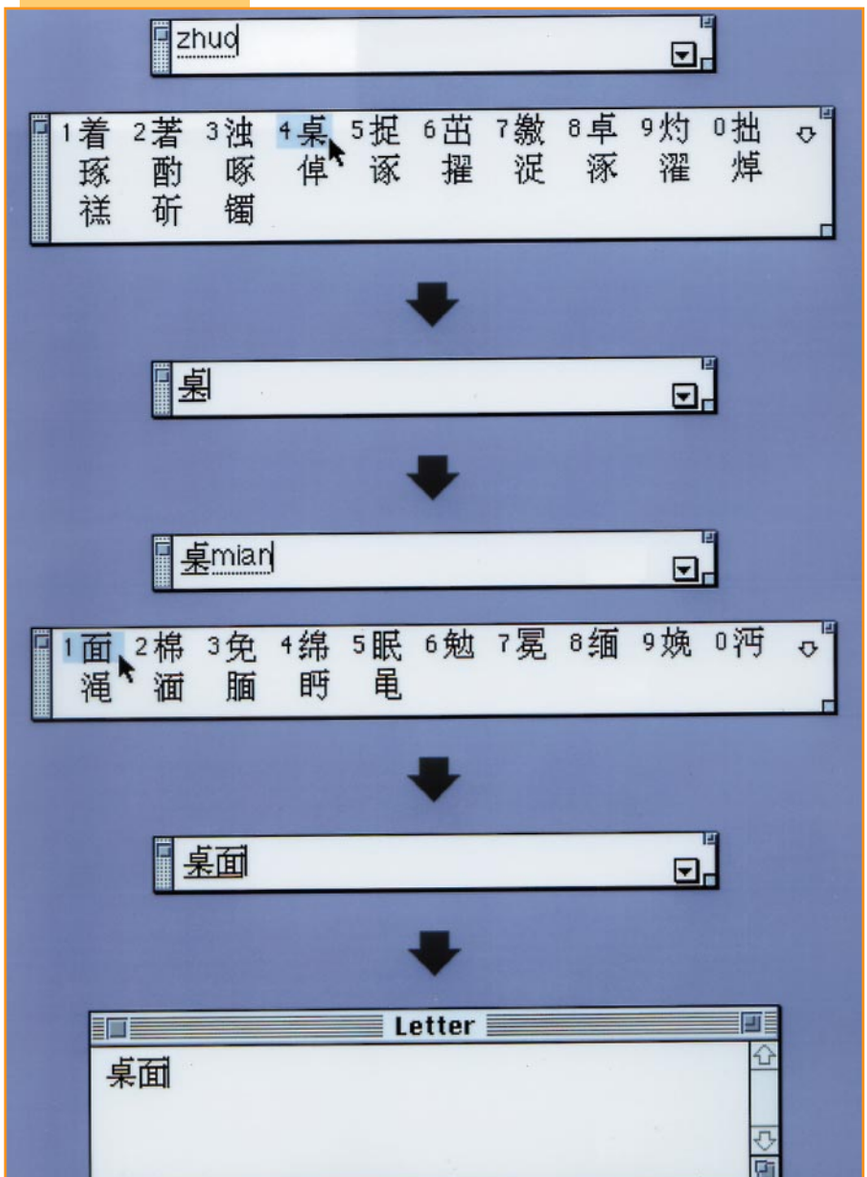


Figure 1 Input of the Chinese word for "Desktop" using the "Hanyu Pinyin" keyboard input method. Entering a phonetic form brings up a choice of characters with the same pronunciations. The user has to pick the desired character.

the method is not applied for a while. However, Wubi experts are able to input Chinese text very quickly and reliably. In fact, in China, this expertise is so valuable that many Wubi users, who are professional computer typists, do text input for a living.

Although today's personal computers can be localized to display Chinese script, creating text remains the main obstacle for all Asians who speak only Chinese. Text input is the key problem that is to be solved before personal computing can become a mainstream activity in Chinese-speaking cultures.

Some Essentials of the Chinese Language

Human interfaces for Chinese customers must address a diversity of issues related to the distinct social, cultural, and economic context in Asia. However, the complexity of the Chinese language does not leave much choice but to focus on the language-related problems first. Many characteristics set it apart from the most basic concepts we know from Western languages [1]. The following explains more.

The Immense Number of Characters

The standard set of characters totals 20,000, and the extended set, which includes names, scientific terms, etc., includes 50,000 characters; 3,000 characters are the minimum for everyday communication.

Chinese text input on a computer becomes mainly a search task. Finding one of the 26 English characters is easy. Western users do not generally perceive searching as part of the input task—unless they are very bad at typing. For the Chinese, however, finding a character out of 3,000 choices is the central part of the input task. Navigating to the desired ideograph often requires complicated interactions that distract user's natural flow of thoughts and essentially conflicts with the actual text creation activity.

Although Chinese are used to coping with approximately 3,000 characters, the handling of characters that goes beyond everyday vocabulary requires active assistance from within the input tool. Learning the written form, pronunciation, and application of unfamiliar characters is a perpetual process. Furthermore, the language profiles of individual users beyond the 3,000 common characters can be diverse, depending on the domain, social context, and education of the users.

Spoken Chinese: Ambiguous Meaning and Unintelligible Dialects

Chinese is spoken with seven major and more than 50 minor dialects. The term *dialect* is somewhat misleading because the various dialects are mutually unintelligible. Although they are based on the same script, the pronunciations differ in a way that is more comparable to English versus German. Although the People's Republic of China established Mandarin in 1956 as the official dialect, it cannot be considered a general solution for human interfaces at this point. A speech interface, which works fine in Beijing, will be useless for most customers in Hong Kong or Singapore where the Cantonese dialect is predominant.

Another issue of spoken Chinese are the homophones: One spoken syllable can mean multiple characters—up to 60 per syllable. "Wu" for example stands for 20 different characters with totally different meanings: *things, 5, fog or dance,* etc. The right character can only be identified from the context or by viewing the written form. Users may sometimes know how to speak a character, but they are not sure about the written form or vice versa. Because of the potentially ambiguous and complex nature of spoken Chinese, speech interfaces for this language are a particular challenge.

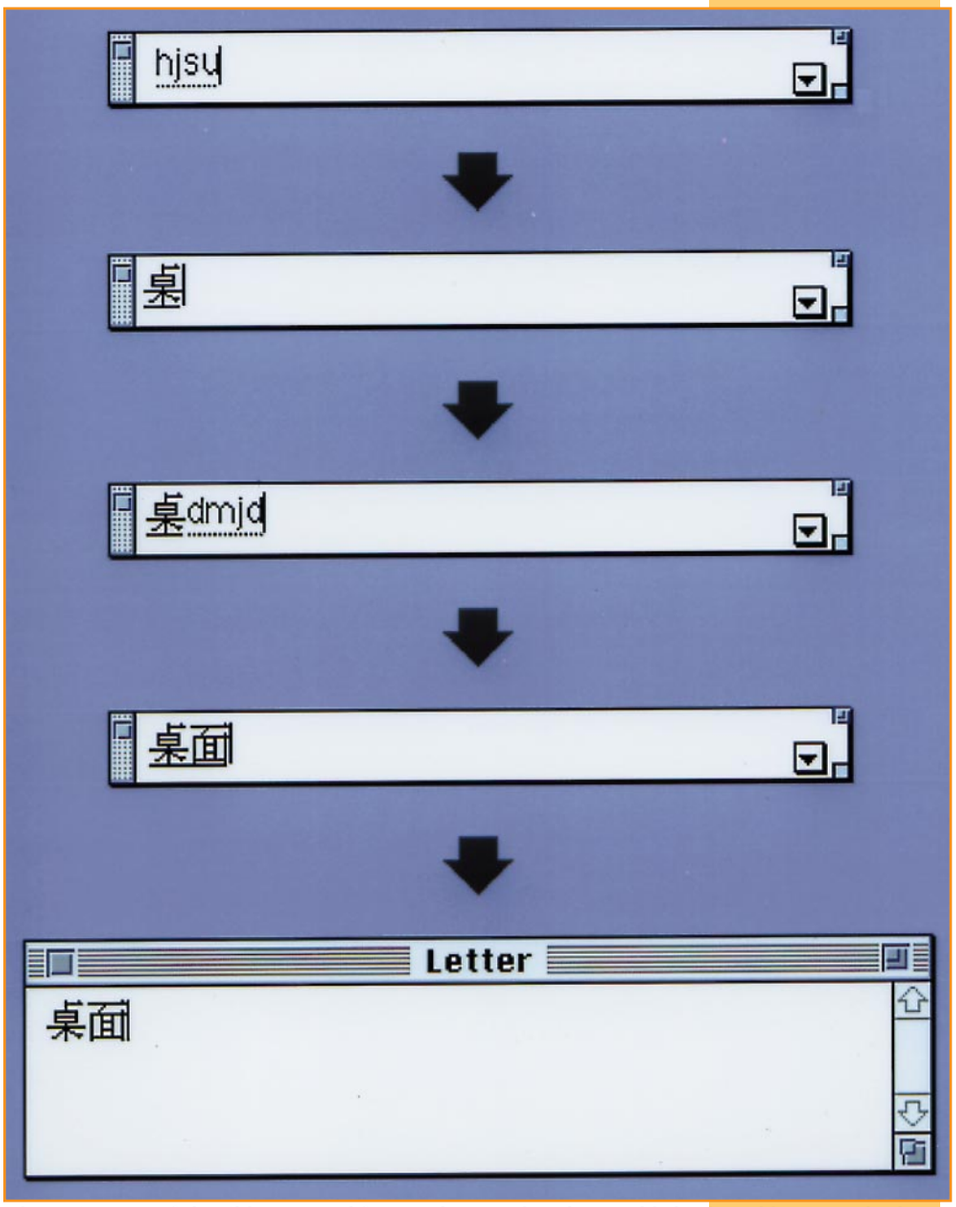


Figure 2 Input of the Chinese word for "Desktop" using the "Wubi" keyboard input method. A specific sequence of roman characters represents the structure of the written Chinese ideograph. Entering this code brings up the character.

Dictionaries Everywhere

The tool to cope with a massive number of characters, as well as the diversity of possible pronunciations, is the dictionary. The functionality of the dictionary represents the core component of Chinese input methods. Considering the frequency of use and close relation with the text-creation process, the dictionary itself has to be

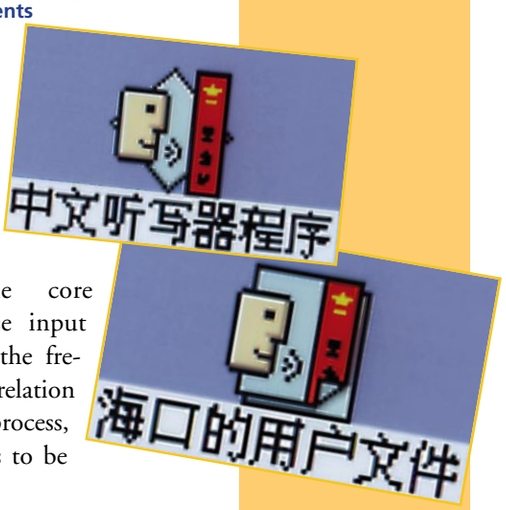


Figure 3 The Apple Advanced Chinese Input Suite integrates a microphone, a tablet and several software components into a comprehensive human interface solution for Chinese text entry.

In October 1996, the tangible results of the joint efforts that took place in Singapore and California were presented in the form of the Apple Advanced Chinese Input Suite. The product is currently being sold and integrates a microphone, a tablet, and several software components into a comprehensive human-interface solution for Chinese computing (Figures 3, 4, and 5). It features an application for dictation (speaker dependent, discrete word, speech synthesis) and a handwriting input method (user independent, character based) for Chinese. With the input suite, any Chinese-speaking user can instantly enter text both quickly and reliably into all documents and text fields in the Apple Macintosh computer. This is a significant step toward making computers accessible to mainstream customers in Chinese-speaking cultures.

seen as an input tool. It provides active assistance in search, identification, and input of the desired character from the large database. Users can also get complementary information on spoken and written forms in order to disambiguate Chinese expressions. Dictionary services that are integrated smoothly and transparently into compound input tasks drastically increase input speed and ease of use.

Lessons Learned from "Keyboardless" Text Input

The defined mission that the team at the Apple-ISS Research Center in Singapore identified was straightforward: Chinese-speaking users should be able to enter text into the computer based on the language skills that they acquire and apply in daily life. To accomplish the vision of a "keyboardless" alternative, the team developed a range of advanced language technologies specifically for Chinese: handwriting recognition, speech recognition, speech synthesis, and language modeling.

Language-independent issues of speech and pen interfaces, which have been discussed in this magazine and other publications, have been addressed [2–4]. However, many features of our design incorporate specific solutions for Chinese. They are good examples for the requirements and constraints that must be taken into account for human interfaces in an Asian context.

Interdependence of Handwriting and Speech Input

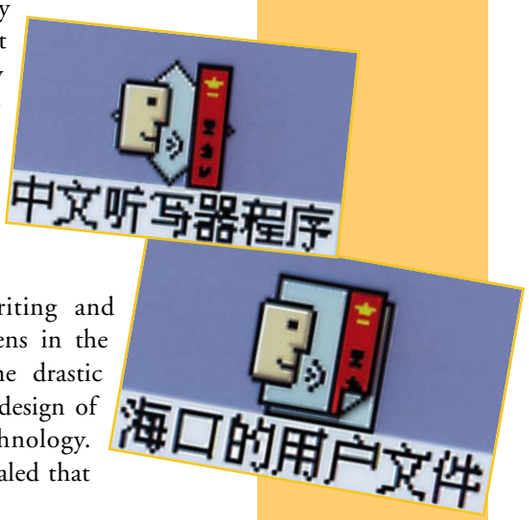
The coupling of speech and handwriting as complimentary tools evolved as an essential requirement for a keyboardless, Chinese user interface. Whereas speech input is generally preferred by Chinese for its speed, it does not suffice as a stand-alone solution. Handwritten

Figure 4 A Chinese name for a folder is created by handwriting. With the Chinese Input Suite users can input Chinese characters directly into all documents or text fields in the Macintosh Finder.

input becomes necessary in cases in which the spoken form of a character is ambiguous or if the user is not familiar with the correct pronunciation of an expression. Most Chinese users prefer to dictate larger passages of the text into the computer and then use the pen to make corrections.

With the current technology, however, truly multimodal interactions that involve speech and pen simultaneously are not supported. Studies done at our center indicate that when dictating text to a computer, Chinese intuitively use both channels to disambiguate spoken expressions. This relates to a technique that exists in the real world. In conversations, Chinese sometimes clarify their spoken statement by moving their finger over the palm of their hand as if they were writing the correlating character. This is an intuitive gesture that Chinese perceive as a simple and effective way to be more precise.

A unified handwriting/speech-recognition framework that is capable of processing both input channels could effectively support and use this behavior. The complimentary information provided by multimodal input could also greatly enhance the accuracy of the character-recognition systems.



The Cost of a More "Natural" Input

Supporting both writing and speaking, as it happens in the real world, had some drastic requirements on the design of the underlying technology. Our user studies revealed that

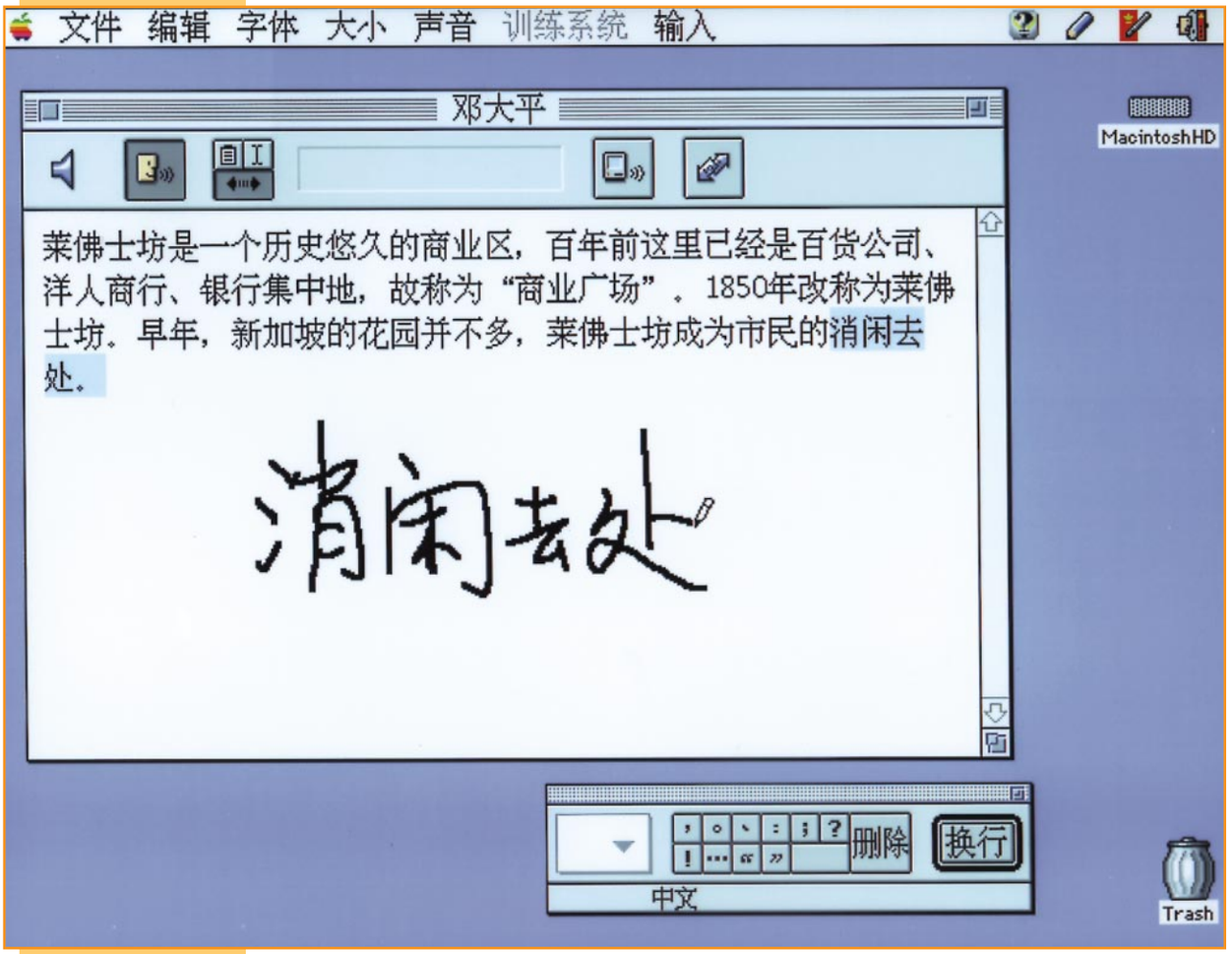


Figure 5 Handwriting and speech are integrated as complimentary tools. Text can be created and edited directly into a document window.

almost all Chinese write cursively. Chinese characters as we know them from books or films are typically the printed form: The strokes are written in separate, straight lines. In cursive writing, however, multiple strokes are connected into a composition of loops. Writing cursive Chinese characters is much faster, but the overall look differs significantly from the printed form. In order to develop a model for this way of writing, along with a strategy for its automatic recognition, a massive data-collection effort became necessary. In collaboration with the Chinese Academy of Science, the Apple-ISS team collected 2.5-million cursive character samples from 5,000 people. A similar approach for speech made sure that common forms of speaking the Mandarin and Cantonese dialects were captured and supported by the product. The following example is the Chinese character for *sun*.

Printed form

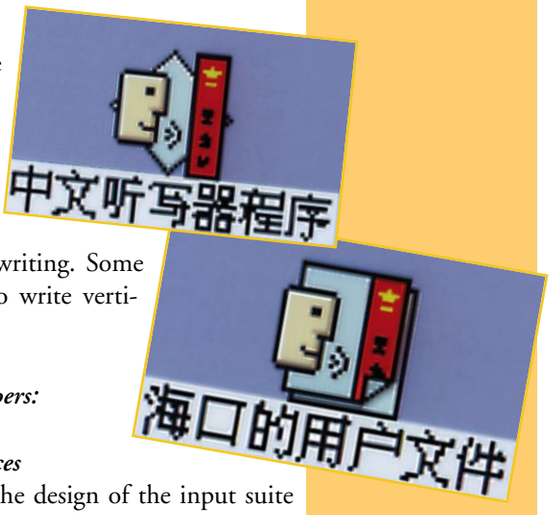
Cursive form

Despite this achievement, the degree of “naturalness” of the text input interaction was constrained significantly by the underlying recognition technologies. The speech recognition is based on a discrete phrase level. As a consequence, users need to put short pauses between individual phrases, which interrupts the natural flow of spoken Chinese. A similar issue arises for characters that are written tightly together. To overcome these issues in the future, the team already completed the development of a continuous speech-recogni-

tion engine for Chinese, along with advanced handwriting segmentation methods. Both technologies effectively lead to interaction techniques in future generations of the system that are much closer to writing and speaking in the real world.

Writing and Speaking as a Universal Input Method

As a consequence of the keyboardless approach, the design of speech and pen interaction must support the whole range of potential computer users. Individual variations in how people use the keyboard are rather limited. When it comes to personal preferences and styles of speaking and writing, however, a much larger range of individual usage patterns has to be supported. This has proved to be particularly true for Chinese handwriting. The size of characters and the layout of the drawing area differ significantly depending on experience and individual style of Chinese handwriting. Some users also tend to write vertically (Figure 6).



Unobtrusive Helpers: Integrated Dictionary Services

Many details in the design of the input suite reveal integrated dictionary assistance in compound input tasks. The rationale behind most of this functionality is very specific to Chinese, and there is no comparable equivalent in Western languages. This feature, however, is instrumental in increasing input speed or

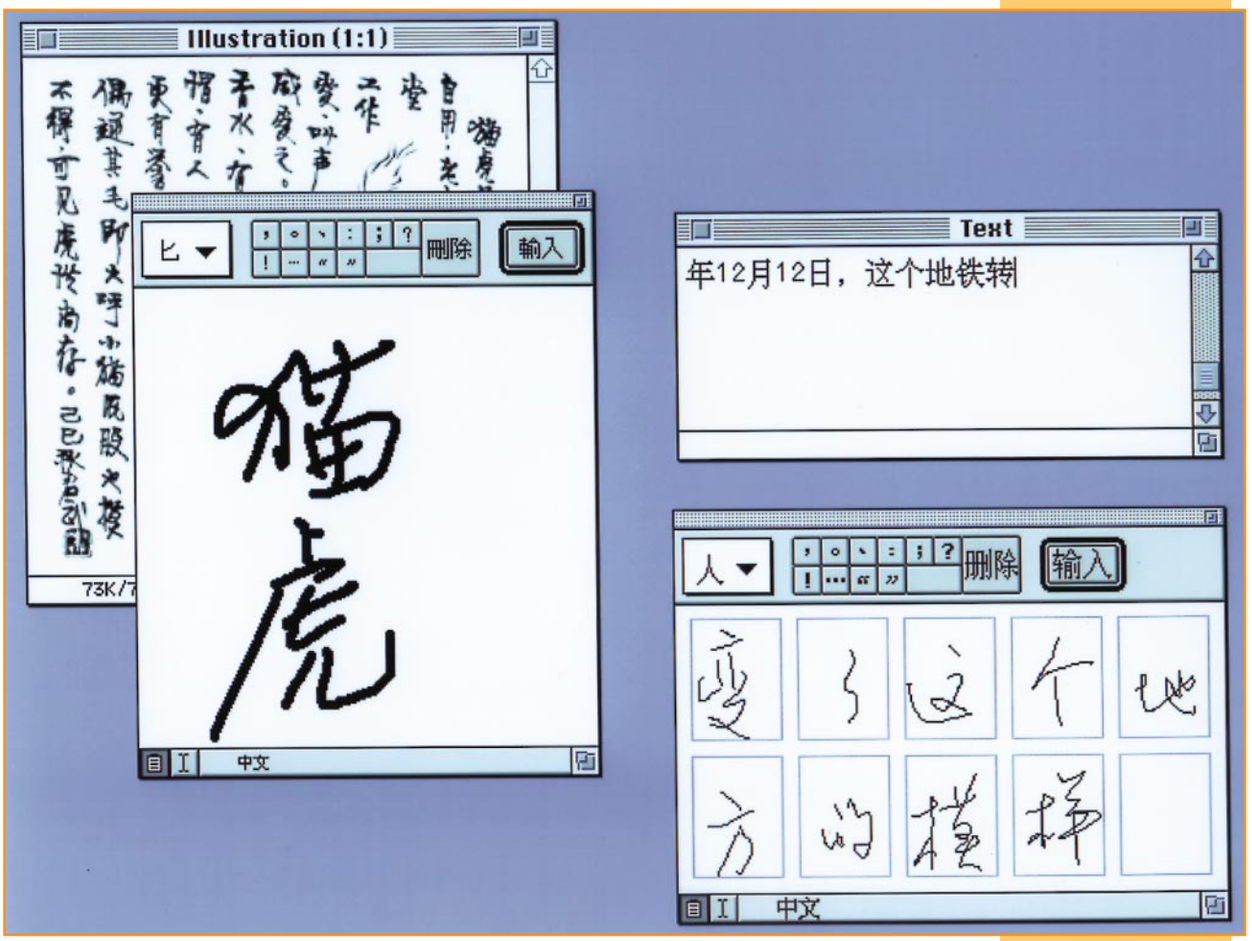


Figure 6 The handwriting input method is customizable according to individual styles, preferences and application contexts. Left: The input area is configured to accomodate large characters, that are written vertically. Right: This user prefers to use the grid for writing multiple small characters in one go.

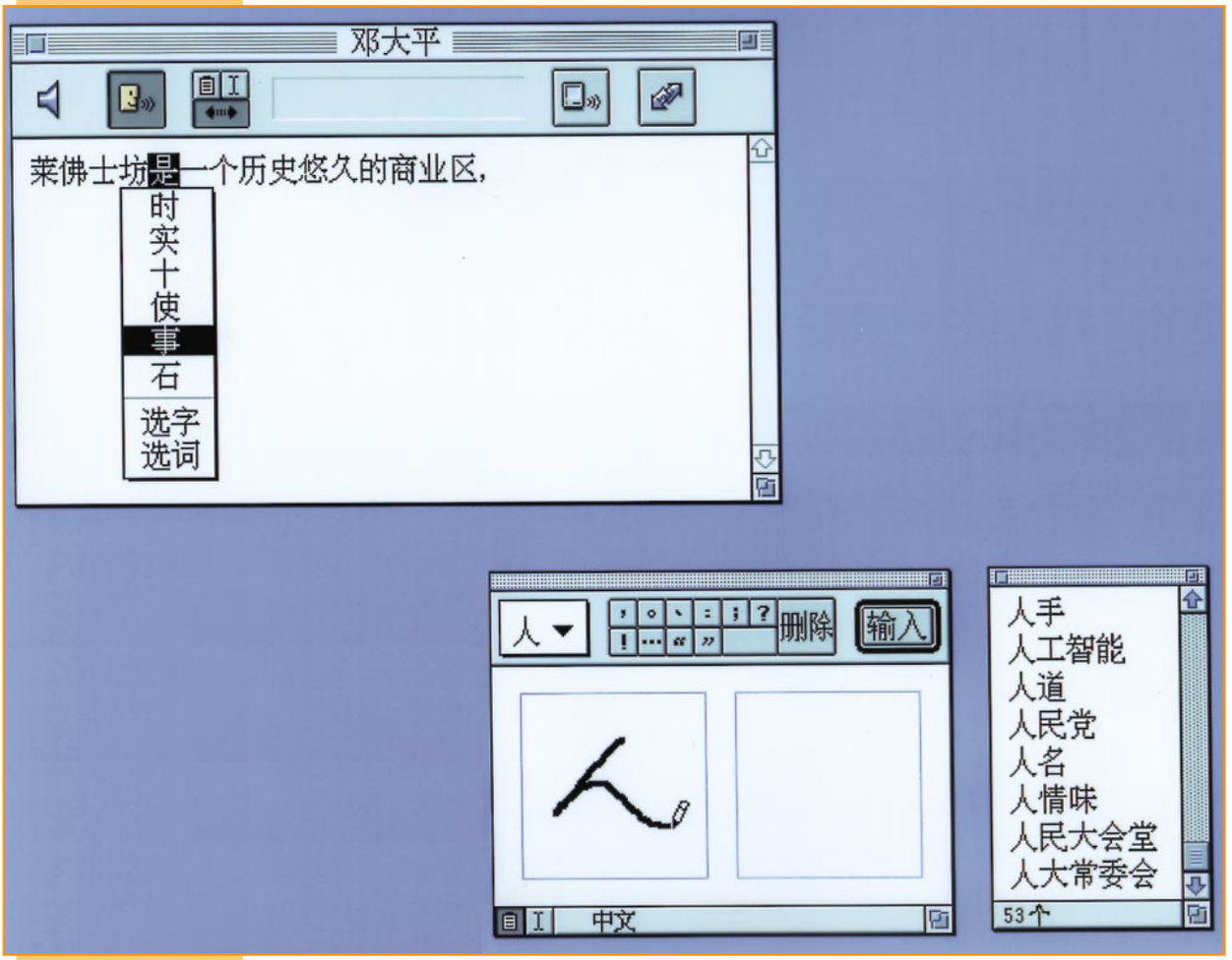


Figure 7 Language specific design solutions are provided to increase input speed or to provide help with unfamiliar characters: The speech homophone list (left) and the handwriting phrase list (right).

helping with unfamiliar characters. An example for speech is the homophone list, and an example for handwriting is the phrase list (Figure 7).

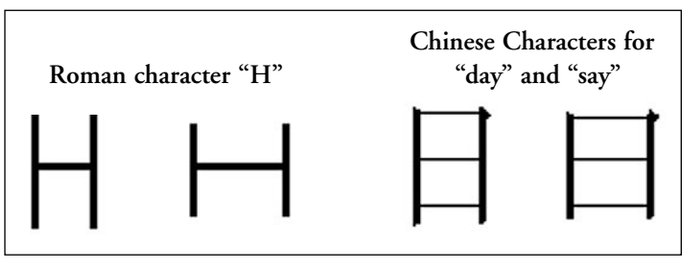
The latter is a good example of how a dictionary actually becomes an input tool. For example, when the user writes the character for *person*, a list will display all phrases based on this character, for example, *man made*, *manpower*, or *people*. The selected phrase then gets inserted into the document. This technique effectively increases the input speed. Although the character for *person* has only 2 strokes, the second character in *man made* is a complex character with 9 strokes. Thus, the user effectively inputs 11 strokes with only 2 strokes and a click. This is tremendously useful considering the fact that phrases can be 4 or 6 characters long, and individual characters can consist of up to 16 strokes.

Chinese also perceive the phrase list indispensable for input of phrases that are made up by characters with which they are not familiar. Users can go through the list, find the correct phrase, and let the list input all of the complicated characters with only one click. Design solutions that integrate automatic dictionary assistance into a compound input task have great potential for Chinese interfaces. Dictionary tools should ideally be totally implicit so that users can focus on the actual text input task.

Speaking and Drawing with a Typewriter

An important marketing requirement of Apple's Advanced Input Suite was to make it run on standard desktop computer configura-

tions and the common peripherals. Although fulfilling this requirement drastically increased the affordability and compatibility of the final product, implementing Chinese-specific interaction techniques based on a device configuration that evolved totally from a Western context revealed some serious drawbacks. Generally controlling the variety of functionally distinct input devices (keyboard, mouse, microphone, tablet) turned out to be a complex task for users. Particularly problematic was the hand–eye coordination when writing Chinese on an opaque tablet. Unless users are experienced, the handwriting of Chinese as well as of Western users gets distorted. Chinese handwriting is significantly more sensitive toward this phenomena. For example, the overall proportions have relevance for the meaning of a Chinese ideograph. A distorted Roman character normally just looks ugly, but for Chinese, a slightly distorted but otherwise correctly written character may switch to a completely different meaning.



The shift toward speech and pen as the main interaction techniques should also have its counterpart in the design of the physical devices. At that moment, users are in a situation in which they "dictate to a typewriter." Future concepts should also look at greater integration of speech and pen interaction on the level of industrial design. The characteristics of an Asian interaction paradigm should be incorporated equally into software and hardware and should result in a completely new conceptual model for an "Asian personal computer."

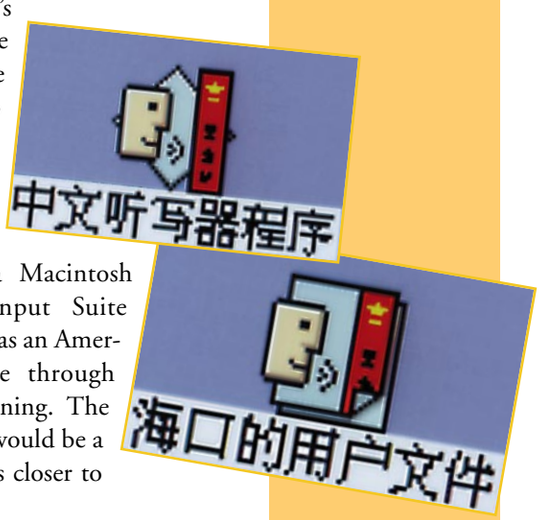
Some Other Impressions

In the design practice in Singapore, it became obvious that little knowledge on how to design human interfaces specifically for Asian users exists in the design community. Furthermore, the predominant, Western model of personal computing turned out to represent a design constraint in a double sense. On one hand, the actual design of standard hardware and software elements did not always allow the envisioned solutions for Chinese to be implemented. Second, it was sometimes difficult for designers, as well as for participating users, to switch off the pre-existing Western conventions about personal computers and get to the bottom of what was really important in an Asian context.

Working with Asian users in the design process also turned out to have some distinct aspects. Politeness, formal attitudes, and different conventions toward expressing personal views can have significant impact on usability studies. The common perception of computer technology as being expensive, foreign, and somewhat intimidating adds to that phenomena. Users, close to desperation while testing a prototype, sometimes concluded with "I am sure I can get used to it."

Conclusion: The Future of Asian Computing

In many Western cultures, tools that did not exist a few years ago, such as word processing or electronic mail, have become so ubiquitous that we can barely live without them. This is yet to happen in most parts of Asia. Solutions that enable mainstream customers in Asia to interact with computers in their own language will play a key role in this development. Apple's Advanced Chinese Input Suite can be seen as a first step toward this new generation of human-interface solutions for Asians.



At this point, a Macintosh running Apple's Input Suite could be interpreted as an American who has gone through intense Chinese training. The final goal, however, would be a computer that comes closer to

Figure 8 The Apple Advanced Chinese Input Suite is the first integrated input system that enables users to enter Chinese in a way that is natural, intuitive and easy-to-use.

being a native Chinese. It would feature an interaction design that does not require any interface layers that are in conflict with the natural means of Chinese communication. This will inevitably require solutions that go beyond a peripheral modification of the existing computing platforms. Core language capabilities, such as language modeling, dictionary services, and input methods, need to be integrated on a system level. Consequently, individual applications can effectively use this platform and incorporate profound Asian characteristics with a consistent user experience. This must be reflected in a unique industrial design that functionally and symbolically resembles the essentials of Asian language computing.

The team from the Apple-ISS Research Center recently founded a new research group: The Asian Interactive Systems Group at the Kent Ridge Digital Labs in Singapore. The group is following up this challenging research direction, which may eventually lead to an entirely new concept for a personal computer that embodies a truly Asian user experience.

Acknowledgments

The author thanks Gareth Loudon, Jean-Luc Lebrun, Gurminder Singh, and Monika Hartwig for their input on this article.

References

1. Li, C. N., and Thompson, S. A. *Mandarin Chinese, A Functional Reference Grammar*. University of California Press, Berkeley, 1981.

2. Schmandt, C. V*oice Communication with Computers*. Van Nostrand Reinold, New York, 1994.

3. Tappert, C. C., Suen, C. Y., and Wakahara, T. "The state of the art in on-line handwriting recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence 12.8*(August, 1990), pp. 787–808.

4. Yankelovich, N. "How do users know what to say?" *interactions 3.6*(1996), pp. 32–43.