

# Chinese Character Usage in New York City

By Mathew Kane



**Sooner or later** many project managers will be asked to work on a Chinese translation project where they need to choose between Simplified or Traditional characters. To help them decide, they might have a map or chart detailing which character set is used in which part of the Chinese-speaking world. But it is not enough to consider character set alone. In Chinese, there is a conceptual distinction between the written (中文) and spoken (汉语) language, so you must consider both when placing a translation job.

## Written (中文)

Written Chinese is comprised of characters that convey meaning and sound. Chinese does not possess an alphabet, so Chinese words are not constructed with letters that indicate pronunciation. Instead, each Chinese character represents a single syllable, and Chinese words are comprised of one or more Chinese characters.<sup>1</sup> Writing Chinese requires the memorization of thousands of characters, some of which are com-

---

When starting a translation job involving the Chinese language, it is not enough just to determine if the target audience uses Traditional or Simplified characters.

---

plex and contain a large number of strokes. (By “strokes,” I mean brushstrokes, or the physical lines that make up a character.)

In the 1950s and 1960s, the communist government on Mainland China developed and promoted a Simplified character system, with the goal of increasing literacy. (Former Communist leader of China Mao Zedong [1893-1976] reportedly even considered switching to Esperanto.) The number of strokes in many characters was reduced; for example, the character for “dragon” was changed from 龍 to the simpler 龙.

Simplified characters were adopted in the People’s Republic of China in 1956, but other regions held on to Traditional characters, most notably Taiwan, Hong Kong, and Macau. It should be noted that most Chinese speakers can comprehend the majority of either character set, but they tend to be more comfortable with one over the other.

## Spoken (汉语)

Chinese is comprised of hundreds of different dialects (e.g., Mandarin in Beijing, Cantonese in Hong Kong, Fukienise in Fujian, etc.) that are

largely mutually unintelligible but share the same written characters. These characters, however, do not necessarily mean the same thing across dialects. For example, in Mandarin the character 找 (pronounced “zhao”) means “to look for,” but in Cantonese “to look for” is written 玩 (pronounced “wan”), which means “to play” in Mandarin. So, in spite of the fact that Cantonese speakers in Hong Kong and Mandarin speakers in Taiwan both use Traditional characters, it is possible that someone who grew up in Taiwan speaking Mandarin and writing Traditional characters would not be able to make sense of a phrase in a Hong Kong newspaper written for the Cantonese dialect.

Phrasing and grammar also vary according to region. For example, even though the Taiwanese and Mainland Chinese both speak Mandarin, the Taiwanese say 不客氣 (“bu ke qi”) to express the idea of “You’re welcome,” whereas in Beijing you would say 沒事儿 (“mei shi er”); if you said “mei shi er” in Taiwan, they probably would not know what you are talking about.

For these reasons, when starting a translation job involving the Chinese language, it is not enough just to determine if the target audience uses Traditional or Simplified characters. Because of differences in dialect, it is also necessary to determine the geographical location (e.g., Shanghai, Hong Kong, Singapore, Chinatown New York City, etc.) and, in some cases, the demographic makeup (e.g., age, year of entry, heritage) of the target audience.

### Mass Transit Authority Case Study

So which Chinese character set and dialect would you use for Chinatowns throughout the United States? I faced this very question while working for

---

## Which Chinese character set and dialect would you use when translating public service announcements for New York City’s subway?

---

the Metropolitan Transportation Authority (MTA) in New York City for my former company LinguaLinx.

The MTA subway map contains both Simplified and Traditional characters, but their printed service notices in subway stations only use Traditional characters. For most of the 20th century, New York City’s Chinatown was comprised of Cantonese speakers from areas like Hong Kong, where they write Traditional characters. Since the opening of the People’s Republic of China in 1979, however, the majority of Chinese-speaking immigrants to New York City has come from Mainland China, specifically Fujian Province, where they write Simplified characters, speak a Fuzhou dialect at home, and speak Mandarin as a second language.<sup>2</sup> This is typical in Mainland China. Mandarin Chinese is taught in schools throughout the People’s Republic of China, so the vast majority of Chinese citizens speaks it, although it is not the primary language of the majority of people. As a result, Mandarin is quickly becoming the *lingua franca* of the Chinese-speaking world.<sup>3</sup>

To help determine which character set would be most appropriate for the MTA, I mined data from the *Selected American Community Survey, Five-Year Estimates, Public Use Microdata Sample, 2006-2010*.<sup>4</sup> Managed by the U.S. Census Bureau, the American Community Survey collects data by

mail, phone, and personal interviews. The American Community Survey does not survey character usage, but does ask which language is spoken at home. However, since this is a fill-in-the-blank question, I found that this data alone was insufficient. Some respondents write “Chinese,” while others specify a dialect like Cantonese or Mandarin. “Chinese” could include any number of dialects. Moreover, as discussed previously, language alone does not indicate character usage. For instance, Cantonese is spoken in Hong Kong, where they use Traditional characters, and Guangdong (Canton) in Mainland China, where they use Simplified characters. Similarly, Mandarin is spoken in Mainland China, where they use Simplified characters, and Taiwan, where they use Traditional characters.

I decided that the best way to determine which character set and dialect are most common in New York City was to create a model designed to estimate the number of people who would use Traditional characters based on the language spoken, place of birth, as well as age and year of entry to the United States. My sample comprised Chinese speakers (those who filled out “Chinese,” “Mandarin,” “Cantonese,” and “Formosan” in the American Community Survey) living in the New York City Public Use Microdata Sample Areas (PUMAs). PUMAs are subdivisions of U.S. states that contain around 100,000 people. ➡

### Model Variables

Below are the variables that I think best predict if a Chinese speaker would use Traditional characters.

**Place of birth is Hong Kong or Taiwan:** Traditional characters are used in Hong Kong and Taiwan.

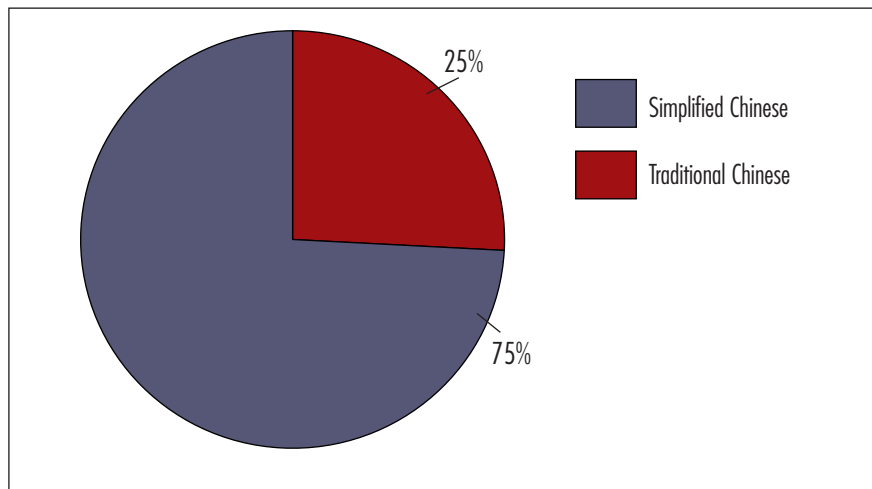
**Place of birth is New York State, age is older than 25:** New York State residents who were born before 1985 most likely were raised by parents who are either from the United States or immigrated to the United States from areas where they use Traditional characters.

**Place of birth is China, age is older than 72:** Simplified characters were adopted in Chinese schools in 1956, so immigrants from China older than 72 would not have been taught Simplified characters in school and would use Traditional characters.

**Place of birth is China, age is 62 to 72, immigrated between 1949 and 1980:** It is possible that some Chinese speakers who list China as their place of birth moved to Taiwan after the communist government came to power in 1949. These people would be at least 62 years old and would probably not have learned Simplified characters because they left Mainland China before Simplified characters were introduced. Significant immigration from the People's Republic of China to the U.S. did not begin until 1980, so the majority of Chinese-speaking immigrants to the U.S. between 1949 and 1980 would use Traditional characters.

**Place of birth is Singapore, age is older than 60:** Simplified characters were adopted in Singapore schools in 1969, so immigrants from Singapore

Figure 1: Traditional/Simplified character breakdown in New York City



older than 60 would not have been taught Simplified characters in school and would use Traditional characters.

### Findings

I used the DataFerrett tool (<http://dataferrett.census.gov/run.html>) to mine the data from the American Community Survey and determine the total populations for each of the aforementioned variables. The populations were totaled to calculate the number of Chinese speakers that would use Traditional characters. I subtracted this number from the total population of Chinese speakers in New York City to determine the amount of Chinese speakers who would likely use Simplified characters. Alyson Slack, project manager for special projects at the Center for Economic Growth in Albany, New York, validated my data.

The results were surprising in that they indicated that long-standing assumptions about the preference for Traditional characters by ethnic Chinese living in the New York City metropolitan area may be outdated. Figure 1 above shows how the percentage breaks out for New York City

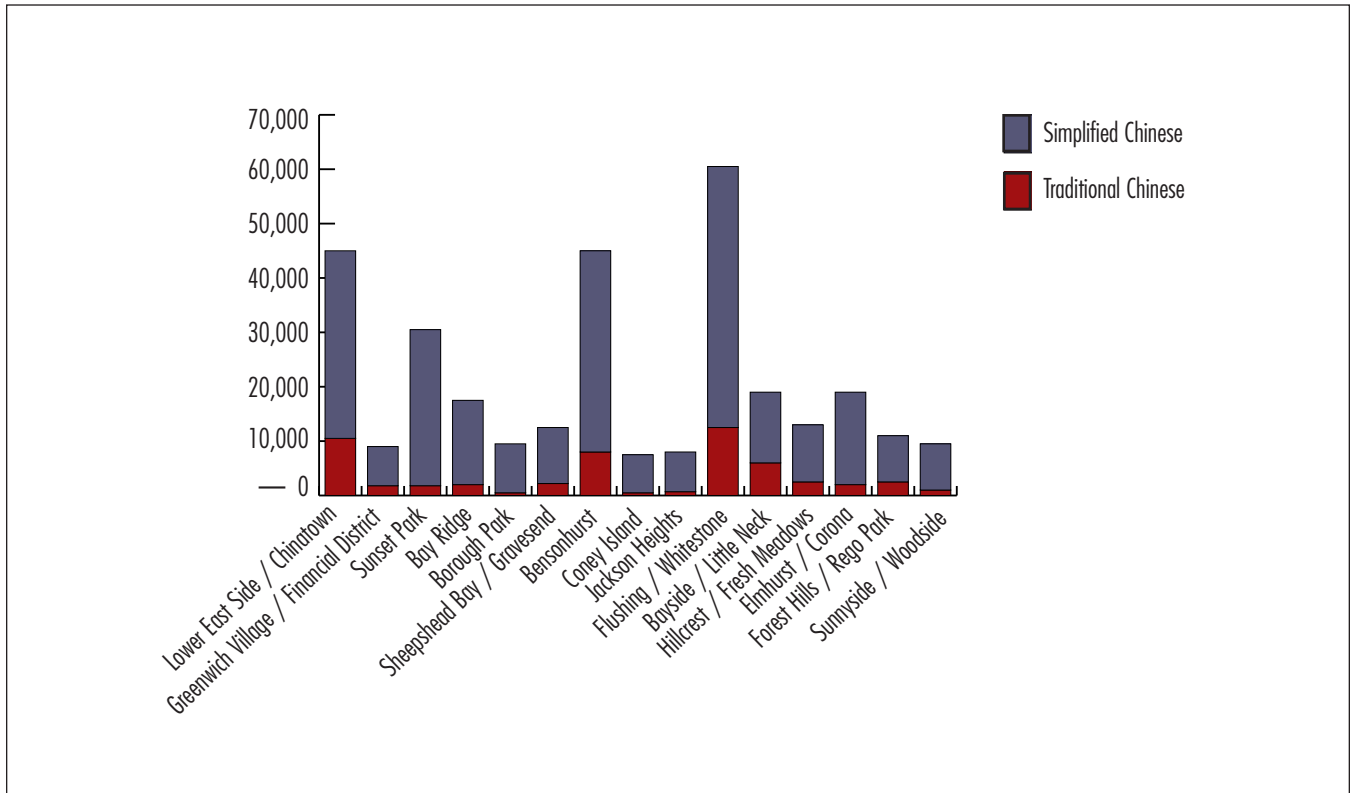
as a whole. Three quarters of the population in question can be expected to prefer Simplified characters. Figure 2 on page 23 shows the breakout for areas in New York City with significant Chinese-speaking populations.

### Presenting the Findings

I finished collecting my findings earlier this year and presented them in case study form to upper management at LinguaLinx. Management approved the case study and gave the green light to present the study to MTA, which at the time had been a client for about one year. The sales department handed out printed copies of my case study to key MTA decision makers during an unrelated meeting at MTA's headquarters in New York City. During the meeting, the sales team summed up the findings and reiterated the recommendation that MTA switch from Traditional to Simplified characters for their printed service notices in New York City subway stations. The individuals from MTA said they would review the case study.

The decision to switch the character set was not taken lightly.

**Figure 2: Traditional/Simplified character breakdown by New York City neighborhood**



Community acceptance of such a significant language change is critical, so it was necessary for MTA to receive buy-in from all pertinent departments and decision makers.

After receiving the necessary approvals in July, MTA accepted my recommendation and instructed LinguaLinx to translate their service notices into Simplified instead of Traditional characters for all subway stations with significant populations of Chinese speakers. MTA creates service notices on an ongoing basis, so

LinguaLinx will translate this content into Simplified Chinese continuously.

The MTA never requested that I write a case study. I did it of my own volition, and it was a surprise when LinguaLinx presented it. As a result, the case study helped strengthen the business relationship between the company and MTA, and provided LinguaLinx with a model for analyzing demographic data for other clients.

**Notes**

1. “Written Chinese,” [www.omniglot.com/chinese/index.htm](http://www.omniglot.com/chinese/index.htm).

2. Foner, Nancy. *New Immigrants in New York: Completely Revised and Updated Edition* (New York: Columbia University Press), 149.

3. Semple, Kirk. “In Chinatown, Sound of the Future is Mandarin,” *The New York Times* (October 21, 2009), <http://bit.ly/Chinatown-Mandarin>.

4. *U.S. Census Selected American Community Survey Five-Year Estimates, Public Use Microdata Sample, 2006-2010*, <http://bit.ly/Community-Survey-US Census>.



For more on ATA's Certification Program, visit [www.atanet.org/certification/index.php](http://www.atanet.org/certification/index.php).