Lin-shan Lee, “Voice Dictation of Mandarin Chinese”, Special Section on Signal Processing in Asia, IEEE Signal Processing Magazine, Vol. 14, Nov 4, July 1997, pp. 63-101.

This is a review paper presenting the problems and approaches in speech recognition for Mandarin Chinese. This is one of the two feature articles of the issue. The national flags of Taiwan (ROC) and China (PRC) were both shown on the cover of the issue sent to readers and libraries worldwide, with that of Taiwan (ROC) slightly larger, obviously because people in both Taiwan and China speak Chinese, but the author of the paper is in Taiwan.

This paper first explained the structural features of Chinese language, then reviewed the research path for Chinese speech recognition considering these features. This included how to analyze all the monosyllables of Chinese speech based on the Initials and Finals as in Table 1 of page 69, how to concatenate the Initial Models and Final Models to obtain monosyllable models as in Figure 7 of page 78, how to construct the 23 tone models considering the pitch contour variation as in Table 3 of page 76, how to cluster the Chinese words into classes so as to obtain a compact language model as in Figure 14 of page 86, and the most advanced prototype system at the time, Golden Mandarin III Windows 95 version, as in Figure 21 of page 94.

本文為該期總共兩篇特色論文之一，是一篇完整探討華語語音辨識之回顧性論文(review paper)。本論文讓我國國旗和中國大陸國旗並排在刊物封面上發行至全球圖書館展示及讀者手中，且我國國旗比中國大陸國旗略大，顯然因為海峽兩岸均說華語，而作者在台灣。

本文首先解釋華文語言之特有結構，再根據這些特有結構回顧華語語音辨識的研究歷程，包括如何考慮由所有聲母、韻母組成的所有單音如第69頁的表1所示，如何把所有聲母模型串接所有韻母模型建構出所有單音模型如第78頁的圖7所示，如何考慮聲調變化建構出23個聲調模型如第76頁的表3所示，如何將華文詞彙分群以建立最精簡的語言模型如第86頁的圖14所示，及當時所完成的最進步的「金聲三號windows95版」實驗系統，如第94頁的圖21所示。