In this article, Cutler and Chen examine how native and non-Cantonese speakers process tonal information versus segmental structure in spoken Cantonese. Previous research on the prosodic dimension of stress in English suggests stress can be omitted during the process of lexical access in favour of segmental information like vowel quality. While English speakers can afford to eschew prosodic information, Cantonese speakers cannot, as tone is necessary for lexical distinction. Despite this necessity, more research on other tonal languages like Mandarin suggests that processing tonal information is slower in comparison to processing segmental information. To see if this phenomenon holds in Cantonese, Cutler and Chen devised three experiments meant to measure reactions times (RT) and error rates of subjects in processing tonal versus segmental information.

The first experiment was a speeded-response lexical decision task, where native speakers were asked to judge whether an utterance that varied in one or more of three components (onset, vowel, and tone) was a real word or not. It was found that compared with single variation in onset and vowel, and multiply varied utterances, a subject was more likely to label a nonword as real when the only variation was tone. Upon closer analysis, it was discovered that the single tone items varied in their position on the F0 scale. Some items were comprised of distinct tones that were easier to process, others were not. Moreover, tone is processed last, dependent as it is on its tone bearing unit (which is segmental). Combined with the pressure to act quickly, subjects might respond before they process tonal information fully, or before it even arrives.

The second experiment was conducted to circumvent the methodological problems arising from experiment 1. In this experiment, native speakers were asked to judge whether two open syllables were the same or different, forcing subjects to wait for the whole utterance. The syllables were varied like in experiment 1 but the items with tone were divided into two sets: one where the contour tones were very distinct (easy), and one where they were not (hard). When the onsets and vowels were different rather than the same, RTs and accuracy were better. In tones, the opposite is true, with no significant difference between same or different tones. However, RTs with easy tone distinction were faster than those with hard tone distinctions. The high inaccuracy rates with regards to tone differences suggests that speakers were not processing tonal information correctly, which points to the limitation of tonal processing, namely that tone is more dependent on acoustic factors rather than linguistic ones.

To test whether tone is dependent on acoustic factors rather than linguistic ones, Cutler and Chen conducted experiment 2 on native Dutch speakers with no background knowledge of Cantonese. In using non-Cantonese speakers, any linguistic factors were eliminated, leaving only acoustic factors. The results were much the same as they were in experiment 2: differences in onset and vowels elicited faster, more accurate responses, and tone differences did the opposite. However, subjects answered more accurately, and faster than native Cantonese speakers (especially in the easy tone set), which can be attributed to the fact that Dutch listeners have no lexical background in Cantonese, allowing them to focus more on the acoustic features of the language.

In conclusion, these experiments have proven that processing tonal information is slower, and more error-prone than processing segmental structure, whether or not it is used regularly in a language. Tonal information arrives later than segmental information, and must be received in full before it can be used correctly. Moreover, prosodic information in general is beholden to acoustic factors, meaning that speakers must process information outside their lexical codes, which limits the usability of prosodic features like tone.