

陳正賢

臺灣大學語言學研究所

論文題目：以計算統計方法及聲學參數研究中文自然語流之韻律短語切分

(Prosodic Phrasing in Mandarin Spontaneous Speech: A Computational-Acoustic Perspective)

指導教授：曾淑娟

論文摘要

Our research investigates the interaction between prosodic phrasing and its grammatical configuration in Mandarin spontaneous speech from a computational-acoustic perspective. We manually segment conversational discourse into clause units (CU), identify their main predicates and perform a semantic role labeling for each CU based on the guidelines proposed in Chinese PropBank. We then propose a comprehensive set of acoustic-prosodic measures to characterize the prosodic units (PU) annotated by Liu and Tseng (2009). Highlighting the gradient and composite nature of prosodic structures, these continuous measures draw insights from both linguistic studies on laboratory phonology and computational modeling of speech prosody in hope to offer a more efficient and theory-independent model of prosodic phrasing for both cross-linguistic comparative studies and its application to natural language processing.

We are primarily concerned with the gradient variation of the PU resulting from its grammatical configuration. The methodological commitment of the present study is to utilize multivariate linear models to evaluate the relation between prosodic phrasing and its interaction with grammar within an integrated model and simultaneously assess the independent effects of multiple hypothesized linguistic factors on different acoustic-prosodic aspects. Our endeavor differs from previous literature in that we look at the influence of the linguistic structures on prosody as a problem of degrees (e.g., How does the linguistic structure contribute to the variation of the acoustic-prosodic measures at the PU boundary?), rather than a problem of a binary classification (e.g., Will the linguistic structure be provided with a PU boundary?).

We first examine how the PU-CU alignment may contribute to the variation of the acoustic-prosodic measures of the PU. This initiative is motivated by the cross-linguistic observation of a strong PU-CU correlation. Previous studies have paid attention to the exact alignment between PU and CU boundaries. We take a computational-modeling approach to examine how the PU-CU alignment may lead to different prosodic structures. Our findings show that the degree to which a PU is coextensive with a CU is systematically reflected in the variation of the acoustic-prosodic measures (such as pitch change, durational pattern, rhythm alteration, change of speech rates). We therefore argue that a clause schema contributes to systematic patterns in prosodic phrasing of Mandarin spontaneous speech.

We further investigate several relevant linguistic factors that might influence the structure of the PU in addition to the CU, including interactionally relevant junctures, phonological factors, relative syntactic positions, and the boundary types of PUs at different linguistic levels. We statistically isolate the prosodic variation of the acoustic-prosodic measures that cannot be attributed to the PU-CU alignment and evaluate the influence of these additional linguistic factors on these acoustic-prosodic residuals. The residualization of the acoustic-prosodic measures enables us to assess the accumulative effects of these linguistic factors on the structure of PUs without the potential confounding of the PU-CU alignment.

Finally, we apply our computational-acoustic representation of PUs in spontaneous speech processing. We explore the possibility of the automatic PU boundary detection by utilizing the comprehensive set of acoustic-prosodic features proposed in this study. Prosodic phrasing creates an intermediate-level segmentation unit that is both linguistically motivated and acoustically prominent. If this intermediate unit can be automatically identified, it would thus provide additional linguistic information for processing spontaneous speech corpora. The objective is to see to what extent our acoustic-prosodic measurements can contribute to the practical task of the automatic PU boundary detection. Encouraging experimental results are obtained and the potential applications of the PU are envisioned. It is suggested that our measures not only have theoretical implications for our conceptual planning and grammatical

structuring in speech production, but also have practical values in the application of the computational modeling.

The contributions of the thesis are threefold. We offer empirical evidence in support of the claim that the clause schema is prosodically indexed in spontaneous speech production. We argue that the look-ahead conceptual planning in our incremental speech production may proceed on a clausal basis as the general outset of the intended proposition is often prosodically anticipated at the onset of the prosodic phrasing. Finally, instead of adopting a syntax-based prosodic hierarchy superimposed upon the analysis of prosody in conversational discourse, we study the prosodic patterning in conversational discourse from a more theoretically-neutral perspective, i.e., relying more on the perceptual acoustic-prosodic cues, whereby the posited prosodic hierarchy can be further verified based on this flat prosodic structure. We suggest that a computational-acoustic representation of one-level PUs may elucidate more comprehensive understanding of the relationship between prosody and grammar in conversational discourse.