

The annotation of information structure in spoken Chinese: a pilot study

Emanuela Cresti, Shuai Luo, Massimo Moneglia, Alessandro Panunzi
LABLITA research unit, University of Florence, Italy

<https://doi.org/10.36505/ExLing-2022/13/0012/000554>

Abstract

The paper sketches a pilot aimed at verifying the consistency of the L-AcT principles and tagging methodology for the annotation of Chinese speech. The pilot confirms both the correspondence between Prosodic Units and Information Units foreseen in this theory and the applicability of its tag-set to Chinese speech.

Keywords: Information Structure, Speech Segmentation, Chinese

Introduction

According to the Language into AcT theory (L-AcT, Cresti & Moneglia 2018), the speech continuum is parsed into prosodic units (PUs) by boundary phenomena that correlate with pauses, pitch reset, lowering of intensity, and lengthening. Prosodic Boundaries are highly perceivable in connection to intentional movements on stressed syllables ('t Hart 1990) and define the edges of PUs, which signal information functions. Terminal prosodic breaks (TB) mark the utterance, i.e. the minimal unit which can be pragmatically interpreted. Non-terminal prosodic breaks (NTB) signal the information units (IUs).

This paper sketches a pilot study aimed at verifying the consistency of the above principles and tagging methodology specifically with respect to the relation between prosody and information structure in Chinese spontaneous speech. The pilot is the preliminary work of a doctoral project which foresees the collection and analysis of a new corpus of spoken Chinese "C-ORAL-CIN" (10 hours, with diaphasic variation) implementing Chinese in the IPIC Information structure data base. In 2. we will show the correspondence with PUs of the IUs of Comment (COM), Topic (TOP), Appendix (APC), Locutive introducer (INT), Parenthesis (PAR) and Bound Comments (COB) and the consistency of their informational value according to L-AcT's tag-set (see. Moneglia & Raso 2014 and the definitions therein).

The annotation of the data set according to L-AcT

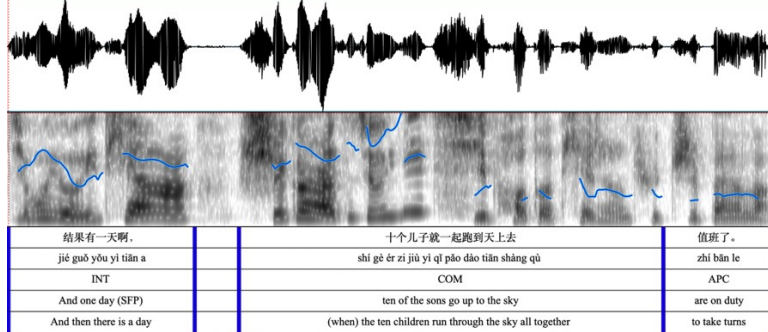
The pilot deals with the annotation of a story told to a child by his mother (the Chinese fairy tale of Hou YI). The data set has been transcribed in characters and transliterated in pinyin. The prosodic parsing of the stretch of speech marking TB and NTB is provided following the perceptual recognition by a

mother tongue. In parallel, the sound wave is analysed through WinPitch and PRAAT, allowing the verification of perceptual judgements. Finally, the word-by-word translation of each chunk in English, and translation in current language is added. In agreement with L-AcT's principles, dedicated layers are provided for the segmentation of the wave in connection with TB and NTB.

The correspondence of terminated sequences with reference units such as *utterances* and *stanzas* has been verified (Izre'el et al.2020) dividing the recording into 90 reference units. To each PU is assigned an IU tag according to the L-AcT tagset. (1) allows to verify the main principle of the L-AcT approach: within an utterance compound by more than one PU, one IU is at the core (the COM), since it bears the illocutionary information. The COM can be pragmatically interpreted for this reason even if the other IUs (in this case INT and APC) are erased.

(1) *LYF: 结果有一天啊, 十个儿子就一起跑到天上去, 值班了。

jié guǒ yǒu yì tiān a/^{INT}shí gè ér zǐ jiù yì qǐ pǎo dào tiān shàng qù/^{COM}zhí bān le//^{APC}
 And one day (SFP)/^{INT}ten of the sons go up to the sky/^{COM}are on duty//^{APC}
 [then there is a day (when) the ten children run through the sky together to take turns]

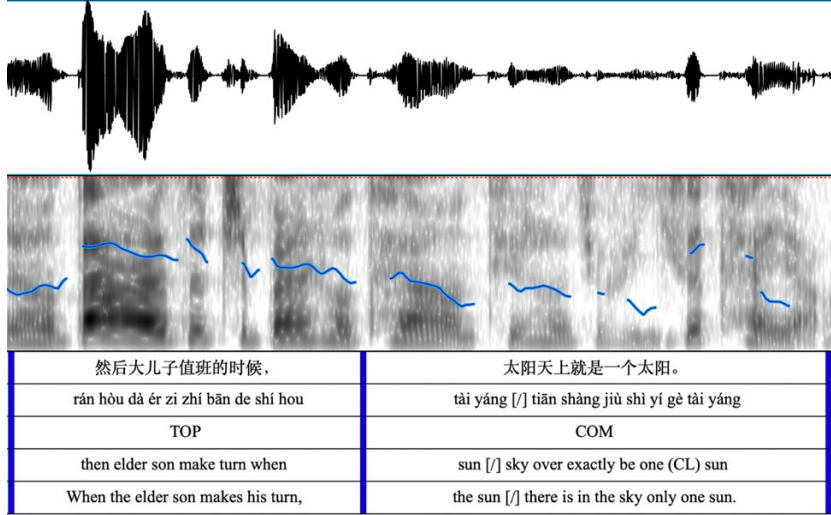


It is worth noticing that, in Chinese, from a semantic point of view, the locutive content in the APC (值班了-zhí bān le) could be in principle a sentence. But this is not the case considering its actual flat prosodic performance, which is not consistent with any possible pragmatic interpretation of the unit. According to L-AcT definition, the APC just *adds unnecessary information to the COM* (4 APC in this text). So, prosody defines the structure of the speech flow, segmenting the same utterance into three IUs instead of parsing it in two utterances.

Although simple utterances occupy an essential role in Chinese speech (43% in the tale), the IUs of TOP, COB and PAR have been retrieved. (2) is an example of TOP/COM utterance, very frequent in Chinese (Li & Thompson 1981), which records 38% of terminated units in mother's speech.

(2) *LYF: 然后大儿子值班的时候, 太阳天上就是一个太阳。

rán hòu dà ér zǐ zhí bān de shí hou /^{TOP} tài yáng [/] tiān shàng jiù shì yí gè tài yáng // ^{COM}
 then elder son makes turn when / ^{TOP} sun [/] sky over exactly be one (CL) sun // ^{COM}
 [When the elder son makes his turn, the sun [/] there is in the sky only one sun.]



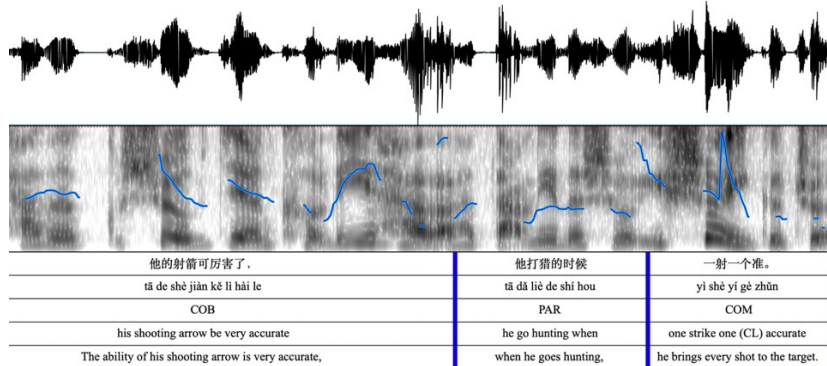
In (2) the TOP is a conditional subordinate sentence, and the COM is a nominal predicate conveying an assertive illocution. Coherently with L-Act definition, the TOP works as *the pragmatic domain relevant for the interpretation of the illocutionary act accomplished by the COM*.

As it is frequent in monologues, terminated units go beyond the utterance limits. In our tale too (19% in mother's speech) sequences ending with a TB are made by Comments bound the one to the other in an adjunctive process (COB). These sequences are called *Stanzas*.

(3) *LYF: 他的射箭可厉害了, 他打猎的时候, 一射一个准。

tā de shè jiàn kě lì hài le / ^{COB} tā dǎ liè de shí hou / ^{PAR} yì shè yí gè zhǔn // ^{COM}
 his shooting arrow be very accurate / ^{COB} he go hunting when / ^{PAR} one strike one (CL)
 accurate // ^{COM}

[The ability of his shooting arrow is very accurate, when he goes hunting, he brings every shot to the target.]



In (3) the first COB is in principle interpretable, but it does not end with a terminal break. On the contrary prosody marks the continuation of the discourse. It corresponds to an assertion which frames the utterance, while the final COM adds an explanation (his shooting-arrow ability is strong).

As it also happens in Romance corpora COBs in a stanza are interpolated by IUs, which add incidental information. These units (PAR), 4 instances in our tale, correspond to prosodic units with less f_0 variation, performed at a lower-mid f_0 values. Accordingly, in (3) PAR, which is a conditional subordinate sentence, shows a f_0 profile at a lower frequency level than the other IUs. The PAR frames the situation, but prosody signals that the circumstance (hunting) is reported “on a different plan” and even if the speaker omits this part, the meaning of the overall terminated sequence still holds.

In conclusion, the research shows that in the Chinese pilot: (a) reference units (*utterance, stanza*) are marked by TB; (b) IUs correspond to PUs; (c) the IU typology of the L-AcT framework are retrieved in spoken Chinese.

References

- Cresti, E., Moneglia, M. 2018. The illocutionary basis of Information Structure. Language into Act Theory (L-AcT), in: Adamou, E., Haude, K., Vanhove, M. (eds.) Information structure in lesser-described languages: Studies in prosody and syntax, 359-401. Amsterdam, Benjamins
- 't Hart, J., Collier, R., Cohen, A. 1990. A Perceptual Study on Intonation. Cambridge, Cambridge University Press.
- Izre'el, S., Mello, H., Panunzi, A., Raso, T. (eds.) 2020. In Search of Basic Units of Spoken Language: A Corpus-Driven Approach, Amsterdam, Benjamins.
- Li, C. N., & Thompson, S. A. 1981. Mandarin Chinese: A Functional Reference Grammar. Berkeley: University of California Press.
- Moneglia, M. & Raso, T. 2014. Notes on the Language into Act Theory, in: Raso, T., Mello, H. (eds.), Spoken corpora and linguistics studies. 468–494 Benjamins,