

Project Proposal - Lin632 Compling

Han Li

April 23, 2022

For the project of this class, I want to build a transducer to implement the tone sandhi process in Mandarin Chinese. In this proposal, I will first introduce the tone sandhi process in Mandarin and then try to tell some plan about this automata, and some questions.

1 Mandarin Tone Sandhi

Mandarin Chinese is a four-tone system which differentiate lexical meanings. There are different notations. The first way is to label four tones by T1-T4. The second notation is to use H and L marks [Yip, 2002]; the last one is to use a five-scale tone value system.[Chao, 1930].

Tone label	Yip's Notations	YR Chao's Notations	Description	Gloss
T1	H	55	a high level	mā “Mother”
T2	LH	35	high rising	má “hemp”
T3	L	214	low rising or dipping	mǎ “horse”
T4	HL	51	high falling	mà “scold”

The tone in question is the T3 in Beijing Mandarin. The basic rule is that when two T3 are adjacent to each other, the first tone will change into a T2:

$$T3 \rightarrow T2 / _T3$$

Chao [1930] noted that the citation tone of T3 will also be reduced from 214 to 21 and lose its rising end in the non-sentence-final position. In this case, there will be 3 different surface forms of T3[Chen, 2000]:

context	214	21	35
before 214	-	-	+
utterance final	+	+	-
elsewhere	-	+	-

Another question should be considered is about trisyllabic words. There are three types of tone sandhi process for trisyllabic words:

1. three monosyllabic words: /jiǔ jiǔ jiǔ/ → [jiú jiú jiǔ] “nine, nine, nice”
2. monosyllabic + bisyllabic /mǎi shòu biǎo/ → [mǎi shóu biǎo] “buy (a) watch”
3. bisyllabic + monosyllabic /shòu biǎo chǎng/ → /shóu biǎo chǎng/ → [shóu biáo chǎng] “watch factory”

It seems that the different words will first do the “grouping” and then apply tone sandhi rule from the left.

2 A Tone Sandhi Transducer

The transducer that this project will build is to take a sequence of citation tones as the underlying forms, and output the surface forms. For now, what I can come up with is this automata:

- $Q = \{q_{head}, q_{nonhead}\}$
- $\Sigma = \{T1, T2, T3, T4\}$ or $\{L, H, LH, HL\}$
- $q_0 = q_{head}$
- $F = q_{nonhead}$

3 Questions

1. How should we choose the tone representations (tone values, H/L notations, or just tone labels)
2. How do we implement the free variations or the optionmality?
3. How do we address the “grouping” problem in trisyllabic tone sandhi (different states?)

References

Moira Yip. Tone. Cambridge University Press, 2002.

Yuen-Ren Chao. A system of tone letters. Le maître phonétique, 1930.

M.Y. Chen. Tone Sandhi: Patterns across Chinese Dialects. Cambridge Studies in Linguistics. Cambridge University Press, 2000. ISBN 9781139431491. URL <https://books.google.com/books?id=D328u70WNgMC>.