Influences of Tone *Sandhi* on Word Recognition in Preschool Children

Dilu Wewalaarachchi and Leher Singh

National University of Singapore
thilanga.d.w@gmail.com, psyls@nus.edu.sg

Abstract

Language learners have to contend with systematic variation in the input in the form of morpho-phonological change. In particular, tone Sandhi refers to a morphophonemic alternation referring to a license tone substitution under particular conditions. Two cohorts of children (3-4 years and 4-5 years) were tested on their ability to recognize words under four different conditions. The first three conditions entailed a change in the form of a word due to i) a native phonological rule (Tone 3 Sandhi), ii) a phonologically illegal tone substitution and iii) absence of a substitution when Tone 3 Sandhi was warranted. The fourth condition consisted of the correct, unaltered form of a word, which served as a control condition. Results demonstrated that 3 to 4 year old children were not able to recognize words that were subjected to Sandhi even when Sandhi was licensed by the phonological context, although they were able to correctly identify instances where the Sandhi rule was omitted, resulting in mispronunciations. By 5 years, children were able to recognize words when Sandhi was applied correctly. Results point to a stabilization of word recognition abilities during the preschool years with respect to suprasegmental morphophonemic change.

Index Terms: language development, tone *Sandhi*, word recognition.

1. Introduction

In order to build up a vocabulary, children must master the full range of sound changes that convey meaning. In addition to mastering the phonological inventories of their native tongues, learners must develop an understanding of phonological processes that influence how phonemes are realized in natural discourse. Examples of phonological processes include phonological assimilation and liaison, both of which have been experimentally investigated in prior developmental research [1] [2] [3]. In both cases of assimilation and liaison, the phonetic realization of a word changes based phonological context. For example, if a speaker produces the words "clean baby", the nasal [n] becomes assimilated to the bilabial sound [b] resulting in a word that resembles [cleam]. This transformation does not occur in other phonological contexts, such as "clean door". Listeners have to spontaneously equate variants of "clean" even if they undergo phonological alternation under particular contexts. As such, it is incumbent upon language learners to 'bypass' the surface transformations caused by assimilation and recover the underlying word. In order to recognize words vulnerable to changes due to assimilation, they must undo the effects of this phonological process and recover the underlying lexical form.

Studies investigating listeners' sensitivities to such transformations reveal a late-emergent ability to compensate for the effects of assimilation. For example, Skoruppa, Mani and Peperkamp demonstrated that while children at 30 months can compensate for place assimilation, they can only do so for familiar native words [1], although when fixation-based paradigms are used, children can compensate for assimilation as early as 24 months [2]. With regards to liaison, a phonological process also resulting in an alternation, Chevrot, Dugua and Fayol found that it was not until 6 years of age that children were able to segment words correctly regardless of whether they were subject to liaison [3].

Tone *Sandhi* is a phonological process that is both similar and different to previously studied processes. Like assimilation and liaison, tone *Sandhi* leads to a phonetic substitution based on the tone context. Specifically, in the case of two adjacent syllables defined by Tone 3, the first syllable changes to Tone 2. Listeners have to automatically recognize that this change is the direct result of tone context and that the first syllable remains linked to Tone 3 even though it is produced in Tone 2. Tone *Sandhi* is different to previous studied phonological processes in that it drives suprasegmental syllable-level changes, rather than segmental consonantal changes.

There have been no experimental investigations of when learners recognize the effects of tone *Sandhi* when learning tone languages. Yet, tone is phonemic in the majority of the world's languages and tone *Sandhi* is present to some degree in all tonal languages [4]. Previous work investigating tone *Sandhi* in early language acquisition has been on individual cases. It appears that at the age of 3, children make frequent errors when negotiating tone *Sandhi* e.g. substituting an incorrect tone in a *Sandhi* context, immediately producing the correct form after an omission of the rule or hesitating before producing utterances corrected for *Sandhi* [5].

The ability to compensate for tone *Sandhi* is an understudied and critical component of language comprehension for tonal language learners. The goal of the current study is to chart the development of sensitivity in native speakers of Mandarin Chinese in which *Sandhi* is commonly observed.

2. Methodology

2.1 Participants

Forty-two native learners of Mandarin Chinese were sampled for this study. Two age groups were tested: 3-4 years of age (n=21, M_{age} = 41.3 months, SD_{age} =4.1 months) and 4-5 years of age (n = 21, M_{age} = 53.2 months, SD_{age} = 3.0 months). None of the children had any known developmental delays or disabilities.

2.2 Experimental design

Participants were tested using the intermodal preferential looking paradigm, widely used to obtain on-line measures of word recognition and word learning.

The paradigm utilizes a split screen display with a familiar object (target) on the left/right side and an unfamiliar object (distractor) on the other side. Familiarity and side of screen are counterbalanced across trials. During each trial, children view a familiar and unfamiliar object on screen for 2500 ms (pre-naming phase). Following this, they hear a verbal label for the familiar object (post-naming phase). The proportion of time spent fixating the target during the salience is computed and compared with the proportion of time spent looking at the target during the naming phase. Word recognition is inferred when proportion of total looking to the target increases significantly during the pre- and post-naming phase. Other possible responses include no increase in fixation to target between the salience and naming phase, suggesting that the participant is uncertain about the referent for the verbal label and fixates equally on the target and unfamiliar object (distractor). The trial structure is displayed in Figure 1.

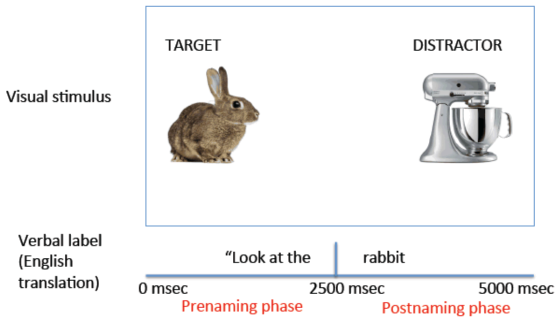

Figure 1: A schematic diagram of the trial structure

12 concrete, imageable, disyllabic nouns that were presumed to be familiar to children aged 3 to 5 were chosen as test stimuli. A post-experiment vocabulary test was conducted to ensure that these early-acquired words were part of individual participant's lexicon.

There were four trial types within each session. In the trial type (six trials), tone *Sandhi* was licensed by the context and correctly applied (*Sandhi* Applied): six words were chosen that comprised two adjacent syllables each spoken with Tone 3 such that the compound underwent the process of Tone 3 *Sandhi* in connected speech resulting in the first syllable produced with Tone 2 and the second with Tone 3. In the second trial type (six trials), *Sandhi* was licensed by context but was not applied (*Sandhi* Omitted). Adjacent syllables in this case were therefore both produced with Tone 3. In the third trial type, *Sandhi* was not licensed by the context and was not applied (Unaltered Correct Label). In this condition, the first syllable was always Tone 2 and the second was always Tone 1. In the fourth trial type, there was a tone substitution from Tone 2 to Tone 3 that was not due to *Sandhi*, resulting in a simple mispronunciation due to an unlicensed tone substitution (Non-*Sandhi* Tone Substitution). For each of these words, pronounced forms were a syllable produced in Tone 3 followed by a syllable produced in Tone 1. This resulted in two trial types that were correctly pronounced: *Sandhi* Applied and Unaltered Correct Label and two trial types consisting of mispronunciations: *Sandhi* Omitted and Non-*Sandhi* Tone Substitution. A summary of trial types is included in Table 1.

***Sandhi* Applied**	***Sandhi* Omitted**	**Unaltered Correct Label**	**Non-*Sandhi* Tone Substitution**
手指 shóu zhǐ (finger)	shǒu zhǐ	河边 hé biān (river)	hě biān
小鸟 xiáo niǎo (bird)	xiǎo niǎo	红灯 hóng dēng (traffic light)	hǒng dēng
老虎 láo hǔ (tiger)	lǎo hǔ	楼梯 lóu tī (escalator)	lǒu tī
雨伞 yú sǎn (umbrella)	yǔ sǎn	茶杯 chá bēi (tea cup)	chǎ bēi
水果 shuí guǒ (fruit)	shuǐ guǒ	肥猪 féi zhū (pig)	fěi zhū
脚趾 jiáo zhǐ (toe)	jiǎo zhǐ	男生 nán shēng (boy)	nǎn shēng

Table 1: Trial types within the experimental session

A female undergraduate who was a native speaker of Mandarin Chinese recorded the stimuli in a child-directed manner. Acoustic analyses were conducted to ensure that all stimuli were consistent with desired tone. A graphical representation of acoustic analyses is displayed in Figure 2.

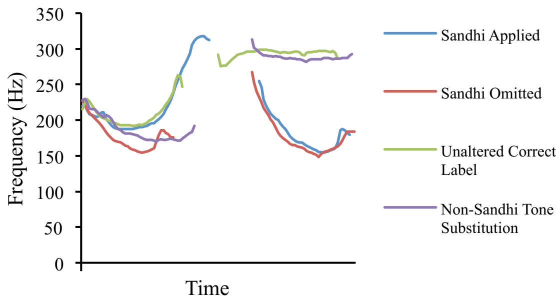

Figure 2: Acoustic Analyses

The experimental design comprised twenty-four randomized test trials, following three "warm-up" trials which were not analyzed. In each test trial, participants saw a side-by-side display of two objects, one familiar and one unfamiliar. They then heard a carrier phrase, *"你看!"* (English translation: "look at the__") followed by the target word.

3. Results

The dependent variable was the proportional total looking (PTL) to target for the pre-naming phase and PTL for the post-naming phase. PTL is defined as the proportion of total looking time at the target (T) compared to the looking time to the target and distracter (D) combined, expressed as T/(T + D). Word recognition is usually indexed by a naming effect which is a significant difference between pre- and post-naming PTL values. Greater post-naming than pre-naming PTL values indicate accurate perception of the lexical tone and correct identification of the labeled object. Preferential looking to the distractor, which would be indicated by smaller PTL values for the post-naming phase, would denote incorrect identification of the labeled object and suggest inaccurate perception of tone. The difference between pre and post-naming values represents a naming effect. Positive naming effects indicate that the participants has mapped the verbal label onto the target, negative naming effects suggest that they have mapped the label onto the distractor (and excluded the target as a potential referent), and naming effects that do not differ significantly from zero represent uncertainty as to the referent for the verbal label. Pre- and post- naming PTL values for the younger sample and the older sample are included in Figure 3 and 4 respectively.

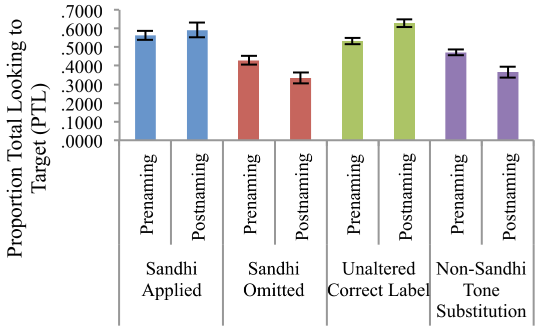

Figure 3: Pre- and post- naming PTL values (younger sample).

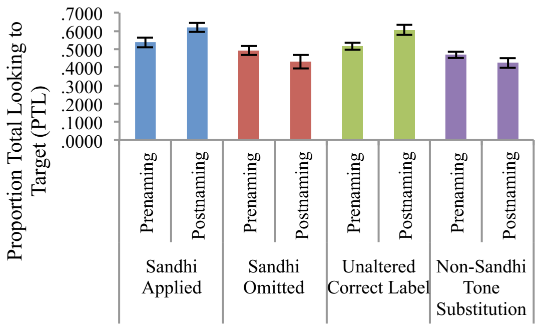

Figure 4: Pre- and post- naming PTL values (older sample).

A 4 x 2 x 2 repeated-measures ANOVA was conducted with pronunciation type (*Sandhi* Applied, *Sandhi* Omitted, Unaltered Correct Label and Non-*Sandhi* Tone Substitution) and phase (pre- and post-naming) as within-subject factors and age as a between-subjects factor. Results revealed a main effect of pronunciation type, $F(3,120) = 32.83$, $p < .0001$ ($\eta^2 = .44$) a marginally significant two-way interaction of pronunciation type and age, $F(3,120) = 2.12$, $p = .102$ ($\eta^2 = .03$) and a significant two-way interaction type of pronunciation type and phase, $F(3,120) = 13.12$, $p < .0001$ ($\eta^2 = .25$). In addition, there was a main effect of age group, $F(1,40) = 3.5$, $p = .068$ ($\eta^2 = .03$). Analyses were then conducted within age group.

Younger participants (3 to 4 years) Within the younger sample, a 4 x 2 (pronunciation type x phase) repeated-measures ANOVA was conducted with PTL as the dependent measure. Results revealed a significant main effect of pronunciation type (F (3, 60) = 20.96, p<. 0001($\eta^2 = .51$) a non-significant main effect of phase, (F (1,20) = 1.24, p =. 28 (ns) and a significant interaction between pronunciations type and phase (F (3,60) = 10.62, p<. 0001 ($\eta^2 = .35$). Follow-up comparisons focused on pre- and post-naming values for each trial type. In trials where *Sandhi* was correctly applied, participants did not fixate to the target object preferentially upon hearing its t (20) = -.72, p=.48 (ns). Where *Sandhi* was licensed but not applied, participants fixated to the distracter upon hearing the label t (20) = 2.83, p =.010 (Cohen's d: 0.80). In the control trials were words were correctly pronounced, participants selectively fixated to the target upon hearing the verbal label t (20) = -4.35, p <.0001(Cohen's d: -1.06). Finally, when a non-*Sandhi* tone substitution took place, participants selectively fixated to the distracter upon hearing the verbal label t (20) = 3.76, p =.001(Cohen's d: 0.95).

Older participants (4 to 5 years) Within the older sample, a 4 x 2 (pronunciation type x phase) repeated-measures ANOVA was conducted with PTL as the dependent measure. Results revealed a significant main effect of pronunciation type (F (3,60) = 12.48, p <. 0001($\eta^2 = .38$), a non-significant main effect of phase, (F (1,20) = .99, p =. 33 (ns) and a significant interaction between pronunciation type and phase (F (3,60) = 4.41, p =. 007 ($\eta^2 = .18$). Follow up comparisons focused on pre- and post-naming values for each trial type. In trials where *Sandhi* was correctly applied, participants preferentially fixated to the target object upon hearing its label t (20) = -2.40, p =.026 (Cohen's d: -0.68). Where *Sandhi* was licensed but not applied, participants did not preferentially fixate to either target or distracter upon hearing the label t (20) = 1.30, p =.21 (ns). In the control trials were words were correctly pronounced, participants selectively fixated to the target upon hearing the verbal label t (20) = -2.51, p =.021 (Cohen's d: -0.83). Finally, when a non-*Sandhi* tone substitution took place, participants did not preferentially fixate to either target or distracter upon hearing the verbal label t (20) = -2.51, p =.021 (Cohen's d: -0.83).

4. Discussion

Results demonstrate an age-dependent maturation in negotiating tone *Sandhi*. It appears that children's understanding of tone *Sandhi* matures qualitatively between 3 and 5 years of age. In an on-line word recognition task, children between the ages of 3 and 4 were not able to recognize words that underwent phonological alternation due to tone *Sandhi*, treating the altered form as a mispronunciation. However, at this age, children demonstrated an emerging sensitivity to tone *Sandhi* as they appeared to perceive words where *Sandhi* was licensed but not applied as mispronunciations. This suggests that while children between the ages of 3 and 4 may recognize the omission of tone *Sandhi*, they may not yet be able to predict the linguistic outcome of the rule in Chinese. In contrast, by 5 years of age, children seem to have a mature representation of words that account for tone *Sandhi*. They appear to have abstracted a phonological alternation rule and are able to rapidly "undo" the effects of surface changes in fluent speech, discriminating correct and incorrect tonal alternations based on the rules prescribed by *Sandhi*.

5. Conclusions

Tone *Sandhi* is a common phonological processes that applies to the most widely spoken human language: Mandarin Chinese. As a result, tone *Sandhi* is a source of complexity in speech with which a large proportion of language learners must negotiate. An appreciation of the phonetic consequences of tone *Sandhi* is necessary to be able to accurately recognize words in languages where tone is phonemic such as Mandarin Chinese. However, there have been no experimental investigations of the effects of tone *Sandhi* on word recognition. The current study aims to investigate this question and results suggest that knowledge of tone *Sandhi* matures by 5 years of age in word recognition.

Phonological processes, such as assimilation, *Sandhi* or liaison, lead to phonemic changes that are licensed within the language. In other words, one phonetic segment has been substituted for another, crossing a phoneme boundary. One might expect that these derived forms are akin to mispronunciations for the young language learner. However, children are very sensitive at an early age to true mispronunciations that are not licensed, detecting tone substitutions in Mandarin Chinese between 18 and 24 months [6]. By contrast, phonological alternations that are licensed by phonological processes, such as assimilation, appear to mature later. The current study demonstrates that sensitivity to phonemic alternations at the level of tone matures later by 5 years of age, after sensitivity to the effects of assimilation which is evident at 2.5 years of age [1] and even earlier using more sensitive fixation-based paradigms [2]. Even later than that, prior investigations of liaison suggest that effects of certain phonological processes are not fully appreciated until age 6 [3].

In combination, these studies point to i) a late emergent course of development with regard to sensitivity to phonological processes and ii) to asynchronous development of sensitivity to morphophonemic change caused by phonological processes. Children's ability to 'undo' the effects of morphophonemic alternations in their native languages is appears to be staggered and to follow a protracted course of development. It is possible that assimilation processes are treated early in development due to the reliability of such changes in human languages. Voicing assimilation and place assimilation are cross-linguistic phonological processes thought to emerge from contextual pressure placed on the articulatory apparatus [7]. In contrast, tone *Sandhi* rules in Mandarin are language-specific and involve substitution of two highly confusable tones [5]. Arguably more complex than this are liaison rules that generate multiple exemplars for each instance of liaison [3] in contrast to tone *Sandhi* where the alternation is between two forms. As such, the competition between possible candidates may be greater in liaison versus in cases of tone *Sandhi*, which may lead to later acquisition of liaison in word recognition tasks versus tone *Sandhi*.

In sum, this serves as the first experimental study of the effects of tone *Sandhi* on spoken word recognition in children. An important component of tone language acquisition, tone *Sandhi* creates a morphophonemic change that needs to be accommodated in spoken language processing. The current study points to a sophisticated knowledge of the role of tone *Sandhi* in native speakers by age 5.

6. References

[1] Skoruppa, K., Mani, N. & Peperkamp, S. (2013). Toddlers' processing of phonological alternations: Early compensation for assimilation in English and French. *Child Development, 84*(1), 313-330.

[2] Skoruppa, K., Mani, N., Plunkett, K., Cabrol, D., & Peperkamp, S. (2013). Early Word Recognition in Sentence Context: French and English 24-Month-Olds' Sensitivity to Sentence-Medial Mispronunciations and Assimilations.Infancy, 18(6), 1007-1029.

[3] Chevrot, J.-P., Dugua, C. & Fayol, M. (2009). Liaison, word segmentation and construction in French: a usage-based account, *Journal of Child Language*. 36 (3), 557-596.

[4] Gandour, J. T. (1978). The perception of tone. In Fromkin, V. A. (Eds) (1978). *Tone: A Linguistic Survey.*

[5] Li, C. N., & Thompson, S. A. (2008). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language, 4(02).*

[6] Singh, L., Tam, J-H, Chan, C. & Golinkoff, R.M. (2014) Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Developmental Science, 17,*1, 94-109.

[7] Cutler, A. (2012) *Native Listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.