

What's in a name? Phonetic and phonological conditioning of personal names in Taiwan Mandarin

Huang, Po-Hsuan
University of Southern California

Keywords phonetic-phonological conditioning; tonal coarticulation; tonal compatibility; name; Taiwan Mandarin

Introduction & background Naming in Mandarin enjoys high liberty where any character can be used to form a one- or two-character name (cf. Wang & Holmberg, 2023). The structures and conditioning of Mandarin personal names thus offer insight into speakers' linguistic preferences/knowledge. However, previous studies mainly focus on the morphosyntactic (Wang & Holmberg) or semantic/sociolinguistic (Kauyska, 2016) aspects of Mandarin names, and those that do touch upon the lower-level processing of phonetic-phonological conditioning mainly focus on sound symbolism (Hsiao, 2005; Chen & Kenstowicz, 2021). Although Hsiao also investigated one non-sound-symbolic phonetic condition, avoidance of tonal coarticulation (TC) (i.e., the preference for the preceding tonal offset to match the following tonal onset), this condition was dismissed due to the small percentage difference in the collected data. It has been shown, however, that the different levels of linguistic domains are unlikely to be entirely autonomous, and past studies have found both examples of higher-level morphosyntax conditioning phonology (Smith, 2011; Inkelas, 2014) and phonologically-conditioned morphosyntax (Shih & Zuraw, 2017). Therefore, the conditioning of names in Mandarin is unlikely to be entirely blind to the lower-level phonetic/phonological processing. This paper investigates phonetic-phonological conditioning in Taiwan Mandarin (TM) names. Crucially, we revisit the avoidance of tonal coarticulation. In Huang's (2024) corpus study, Taiwan Southern Min speakers were found to avoid TC for more informative tone pairs, while TM speakers did not. Huang stipulates that the fewer tonal contrasts (4 tones) in TM allow for a lower perceptual cost, which means TM speakers do not need to constrain TC for informativeness. It is possible, however, that while it may be the case for word production, the high liberty of name composition in TM may invite conditions that could not be taken care of in words, which are already existent and harder to manipulate than names. Aside from the TC avoidance, I also revisit the effects of the sound-symbolic constraints from Hsiao, where only descriptive percentages were calculated. Also included are the several phonological-phonotactic conditions suggested to potentially condition morphosyntax in Shih & Zuraw. A list of conditions investigated in this study is provided in Table 1.

Methodology In this study, a corpus of TM names was used to investigate the phonetic-phonological conditioning in TM names. Linear-mixed effects models (LMM) were then fitted to observe the significance of the conditions. **Corpus** The Chinese Name List Corpus (Tsai, 2017) was used. This corpus collected the names of Taiwanese high school graduates who attended the Joint College Entrance Examination from 1994 to 2004 and those who attended the General Scholastic Ability Test from 2005 to 2012. **Preprocessing** Since the conditions involved the interaction of syllable pairs, names with only two characters (family name+one-character given name) and names with three characters but with compound (two-character) family names were removed. A total of 780,736 tokens were selected. *pinyin* and *pinyin-to-ipa* (Taubert, 2024) were used to determine the tones and segmental structures of the names. **Conditions** The status of each of the conditions for each unique name was then coded as 0 (non-conforming) or 1 (conforming) and then centered around the means. The criteria for each condition are listed in Table 2. **LMM** The frequency of each unique name was calculated. The log name frequency was taken as the predicted variable; the conditions were taken as the predictors; the first and second characters were taken as the random slopes. Following Shih & Zuraw, stepwise backward elimination was used to extract the final model.

Results Significant conditions are listed in Table 3. Two conditions were found to have main effects on naming frequency: *V.V and HT#. Both promoted the likelihood of a name in TM. HT# was also found to have a stronger effect together with *NC. *TC, while not having a main effect per se, was also found to promote the naming likelihood together with HY#, HT#, and the interactions between HY# vs. HT# and HY# vs. *NC. All other conditions also contributed to the conditioning of naming through multi-way interactions. Three interaction effects were, however, inhibitory. This might be due to the incompatible nature of some of the conditions. *Hongyin*'s, for example, exclude syllables with the /j, q/ glides, which will be heavier-weight than syllables without glides, thus contradicting HW#. Another possible cause may be the innate incompatibility of the conditioned probabilities among characters preferred for naming. In the corpus, for example, a pair has a 74.1% chance of having a high last tone if they combine freely, while among pairs that avoid TC, it is only 42.30%, and they have a 23.9% chance of avoiding TC in free conditions, as compared to 13.9% among those with a high last tone.

Discussion & conclusion The results support the hypothesis that lower-level processing of phonetic-phonological conditions other than sound symbolism may also be present in personal names in TM. Specifically, TM naming has been attested to avoid adjacent vowels, a preference also found in English noun ordering (Wright et al., 2005). Several other crosslinguistic phonological preferences were also found to further facilitate naming likelihood through interactions with other conditions. This suggests that TM naming accesses not only higher-level

morphosyntax but also low-level phonological knowledge. In terms of phonetic conditioning, a trend was found where TM names, in interaction with other facilitatory effects, avoided tonal coarticulation caused by discrepant tonal offset-onset pairs. Such a finding compensates for the lack of TC avoidance for informativeness attested in Huang. While TM speakers would not actively seek to manipulate TC during word production, when higher liberty is given, as in naming, TM speakers would still be inclined to avoid perceptual confusion in certain conditions. Aside from phonetic-phonological conditions, our results also partially conform with Hsiao’s consideration of sound symbolism. Names ending with high tones were found to be preferred, and *hongyin*’s in final syllables were also found to interact in a facilitory direction with other conditions. This confirms that sound symbolism is also likely to play a role in conditioning TM names. Overall, this study examined phonetic-phonological factors that were largely overlooked in past studies and showed that the structures of personal names in Taiwan Mandarin are not only subject to high-level morphosyntax/sociolinguistic drives but are also conditioned by lower-level phonetic and phonological constraints.

Table 1: Conditions included in this study.

Condition	Type	Reason of inclusion	Past studies
TC AVOIDANCE (*TC)	Phonetic	Reduction of perceptual pressure	Hsiao; Huang
HEAVY WEIGHT LAST (HW#)	Phonological; phonetic	Crosslinguistic preference; reduction of working memory	Shih & Zuraw
NASAL-VOICELESS SEQUENCE AVOIDANCE (*NC)	Phonological	Crosslinguistic preference	Shih & Zuraw
NASAL OCP (NOCP)	Phonological	Crosslinguistic preference	Shih & Zuraw
VELAR OCP (VOCP)	Phonological	Crosslinguistic preference	Shih & Zuraw
HIATUS AVOIDANCE (*V.V)	Phonological	Crosslinguistic preference	Shih & Zuraw
HIGH TONE LAST (HT#)	Sound-symbolic	The “good is up” metaphor (Lakoff & Johnson, 1980)	Hsiao
HONGYIN LAST (HY#) (<i>Hongyin</i> : syllables without the /j, ɥ/ glides)	Sound-symbolic	<i>Hongyin</i> ’s’ related “wildness and strength”	Hsiao

Table 2: Criteria of the conditions.

Condition	Criteria
*TC	Tones with a tone-level difference smaller than 1 (55+55, 55+51, 35+55, 35+51, 51+21) were classified as not inducing TC.
HW#	The weight of a syllable was taken as the number of segments. A heavier or equal weight of the second syllable versus the first syllable was labeled as conforming.
*NC	A syllable-final nasal followed by a syllable-initial voiceless was taken as non-conforming.
NOCP	Adjacent nasals/velars/vowels at the syllable boundary were taken as non-conforming.
VOCP	
*V.V	
HT#	High-level (55) and the rising (35) tones were classified as high tones.
HY#	Syllables without the /j, ɥ/ glides were classified as <i>hongyin</i> ’s.

Table 3: LMM results.

Condition	Estimate	Std. Error	df	t value	p-value	
Intercept	-0.287	0.011	6463	-25.281	< 0.001	***
*V.V	0.016	0.002	162400	6.506	< 0.001	***
HT#	0.044	0.008	3644	5.760	< 0.001	***
*TC:HY#	0.009	0.002	164600	3.819	< 0.001	***
HW#:HY#	-0.009	0.002	156700	-3.879	< 0.001	***
*TC:HT#	-0.022	0.002	165300	-9.853	< 0.001	***
HT#:*NC	0.017	0.002	161400	6.985	< 0.001	***
*TC:HY#:HT#	0.005	0.002	164800	2.351	0.0187	*
*TC:HY#:*NC	0.006	0.002	165400	2.737	0.0062	**
NOCP:HY#:*NC:VOCP	-0.205	0.052	164900	-3.898	< 0.001	***
*V.V:NOCP:HY#:*NC:VOCP	1.301	0.309	164800	4.212	< 0.001	***

References

- Chen, F., & Kenstowicz, M. (2021). Phonotactics of gender in mandarin given names: Patterns and constraints. In *Proceedings of the Annual Meetings on Phonology*.
- Hsiao, C.-H. (2005). Interaction of Phonology and Semantics in Taiwanese Names – An Optimality Theory Approach. In *SEALS XV* (p. 133).

- Huang, P.-H. (2024). *Informativity and its allocation in tonal coarticulation in Taiwan Southern Min and Taiwan Mandarin*. (Unpublished manuscript)
- Inkelas, S. (2014). *The Interplay of Morphology and Phonology* (Vol. 8). Oxford University Press.
- Kauyska, I. (2016). Linguistic composition and characteristics of Chinese given names. *Onoma*, 51, 161-186.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- Shih, S. S., & Zuraw, K. (2017). Phonological conditions on variable adjective and noun word order in Tagalog. *Language*, 93(4), e317–e352.
- Smith, J. L. (2011). Category-specific effects. In *The Blackwell Companion to Phonology* (pp. 1–25). Blackwell Publishing.
- Taubert, S. (2024). *pinyin-to-ipa (version 0.0.2)*. Computer software. Retrieved from <https://doi.org/10.5281/zenodo.10639971>
- Tsai, C.-H. (2017). *Chinese name list*. <http://technology.chtsai.org/namelist/>. (Accessed 3, Nov., 2024)
- Wang, Q., & Holmberg, A. (2023). The structure of Chinese personal names. *Journal of East Asian Linguistics*, 32(2), 201-244.
- Wright, S. K., Hay, J., & Bent, T. (2005). Ladies first? Phonology, frequency, and the naming conspiracy. *Linguistics*, 43(3), 531–561. doi: 10.1515/ling.2005.43.3.531