TONAL PHONETIC ANALOGY

Alan C. L. Yu

Phonology Laboratory, Department of Linguistics, University of Chicago
aclyu@uchicago.edu

ABSTRACT

Paradigmatic uniformity effects are commonplace in linguistic change. Recent work has extended this idea to the synchronic domain. At issue here is whether paradigm uniformity holds at the phonetic level. This study offers experimental evidence for phonetic analogy from Cantonese, demonstrating that the phonetic realization of a derived tone may vary in the direction of its paradigmatic neighbor.

Keywords: Phonetic analogy, paradigm uniformity, tone, tonal morphology, Cantonese

1. INTRODUCTION

Paradigmatic effects in phonology and morphology are commonplace in linguistic change. Recent work has extended the idea to the synchronic domain. At issue here is whether paradigm uniformity holds at the phonetic level. Simply put, does *phonetic analogy* exist? Morphologically complex structures often exhibit phonetic properties that differ from those occurring in the simplex forms. Phonetic analogy obtains when non-contrastive phonetic properties associated with the morphological complex structure correlates with those properties associated with its paradigmatically related forms. For example, an optional schwa deletion rule in French creates ostensibly homophonous strings (e.g., *bas ʀetrouvé* [baʁətʁuve] 'stocking found again' → *bas ʀ'trouvé* [baʁtʁuve], contrast *bar trouvé* [baʁtʁuve] 'bar found'). Studies have found that the consonant to the left of the deleted schwa's syllable maintains phonetic qualities that would only be expected if the schwa was still present [1]. Steriade [2] contends that the unexpected phonetic differences are inherited from the related utterance without schwa deletion (e.g. /ʀ/ in *bas ʀ'trouvé* [baʁtʁuve] inherits the onset-like articulation from the /ʀ/ in the related phrase, *bas ʀetrouvé* [baʁətʁuve]).

The claim of paradigmatic uniformity effects at the phonetic level raises question about the separation between representations at the phonetic and phonological levels. If paradigmatic uniformity is enforced through conditions that govern both phonological features and non-contrastive phonetic details, Steriade [2] concludes, the boundary between phonetics and phonology crumbles. As a consequence, detractors of phonetic analogy have attempted to recast the alleged cases of phonetic paradigm uniformity in language more in line with a modular, feed-forward model. For example, the consonantal effect in French schwa deletion has been argued to be a case of masking due to gestural overlap [3].

The matter under investigation is a case of tonal morphology in Cantonese. Cantonese has six tone types (55, 33, 22, 23, 25, 21) and three types of syllable: vowel-final (CVV), plosive-final (CVO), and nasal-final (CVN). In CVO syllables, only level tones are found lexically (55, 33, 22). Besides the lexical low-mid-rising tone (RL), 25 tones may also come about as the result of a set of morphological processes known as *Pinjam*, or *Changed Tone* (RM). That is, 25 tones can be derived from semantically related syllables whose tone is not 55 or 25. *Pinjam* alternants may denote the nominalization of a verbal action ([sou33] 'to sweep' → [sou25] 'a broom'; [pɔŋ22] 'to weigh' → [pɔŋ25] 'a scale') or a familiar and/or diminutive object ([tʰoi21] 'stage, terrace' → [tʰoi25] 'table') and various other complicated semantic nuances.

Given that *Pinjam* syllables have paradigmatic neighbors with different tonal realizations, the possibility of phonetic analogy naturally arises. That is, will the phonetic properties of the paradigmatic neighbor of RM affect the realization of RM itself? The following experiment was set out to answer this question.

2. EXPERIMENT

2.1 Stimuli

The present study dealt with four (near) minimal pairs (i.e. with identical rhyme) of Cantonese

words with a derived 25 tone (see Table 1): each pair contains a syllable with a 22-tone paradigmatic neighbor (RM_{22}) and one with a 33-tone neighbor (RM_{33}). As controls, four lexical 25-tone syllables were also included (RL). These data were originally recorded as part of a larger study looking at the phonetics and phonology of morphological tone in Cantonese.

Table 1: Stimuli analyzed in the present study

RM_{22}	pin25	'convenience'
	tɐi25	'brother'
	pɑk25	'white'
	tsɑp25	'mixture'
RM_{33}	p^hin25	'a slice'
	kɐi25	'solution' < 'to calculate'
	p^hɑk25	'a racket' < 'to hit'
	kɑp25	'a pigeon'
RL	kɐi25	'dialogue'
	tɐi25	'bottom'
	pin25	'to deprecate'
	tin25	'a collection'

2.2 Subjects

Eight native speakers of Hong Kong Cantonese (three males and five females) were paid a nominal fee for participating in the production experiment. The subjects were all born and raised in Hong Kong. With the exception of two speakers who have been in the United States for eight years, the other subjects have been in the United States for less than two years at the time of the recording; all are second-language speakers of English. None reported any speech or hearing problems.

2.3 Procedure

The original set of recordings contain (near) minimal pairs of CVO, CVV, and CVN syllables that differ in tonal specifications (i.e. RL, RM, and the level tone paradigmatic neighbor of the RM syllables). To ensure that the appropriate pronunciations were rendered, target words were first presented as part of a disyllabic word phrase (e.g., 影相 [jiŋ25 sœːŋ25] 'glasses') due to the fact that the Chinese characters representing words with a deriving rising tone are often written the same way as its non-derived counterpart (e.g., 相 /sœːŋ/ stands for both 'face' /sœːŋ33/ and 'photograph' /sœːŋ25/). The semantic and pronunciation differences are only transparent when the character is used in the appropriate context. To minimize the subjects' awareness of the focus of the experiment, the target syllables/phrases were embedded and randomized within a larger set of tokens of different syllable structures and different tone types. The subjects recited the list of Cantonese disyllabic words/phrases written in the carrier phrase /ŋɔ23 wui23 tuk˥22 ___ pɛi35 nɛi23 t^hæŋ55/ 'I am going to say __ for you to hear' three times. Each subject recited, in total, three hundred sentences (i.e. 10 target tokens x 3 syllable types x 3 tone versions x 3 repetitions + 10 fillers x 3 repetitions = 300 tokens). The randomized tokens are divided into three lists. Subjects were given the opportunity to take breaks between lists. Subjects' productions were recorded in a sound-proofed room using a Marantz PMD670 solid state recorder and a SONY C-48 condenser microphone which was placed approximately six inches from the subject's mouth.

2.4. Acoustic measurement

The recordings were digitized at a sampling rate of 48kHz, and all measurements were made with PRAAT. The following parameters were measured: the f_0 minima and maxima of the rising contour and the duration of these f_0 landmarks relative to the onset of voicing. Only tokens with clear f_0 contours (i.e. tokens not affected by excess creakiness) were included in the analyses.

2.5 Results

Table 2 summarizes the results of the f_0 minima and maxima measurements. Mean f_0 values pooled across three repetitions were z-score transformed for analysis in order to reduce variation across subjects. A three-way ANOVA revealed significant main effects of syllable type ($F(2, 110)=46.26$, $p < 0.01$) and of f_0 landmarks ($F(1, 110)=626.16$, $p < 0.01$). There is a significant two-way interaction between tone type and syllable type ($F(3, 110)=11.25$, $p < 0.01$) suggesting that the realization of f_0 varies in different syllable structures depending on the existence and the nature of the tonal neighbor. A significant three-way interaction is found as well ($F(3, 110)=4.05$, $p = 0.01$), suggesting that the realizations of f_0 at the

f_0 minima and f_0 maxima differ in different syllable types depending on the tone type.

Table 2: Summary of mean f_0 values and standard deviation (in Hz) at the Turning Point and the f_0 peak across syllable (CVN, CVV, CVO) and tone type (RM_{22}, RM_{33}, RL).

		f_0 minima		f_0 maxima	
Syllable	Tone	Mean	SD	Mean	SD
CVN	RM_{22}	149.04	36.84	190.68	47.15
	RM_{33}	157.18	36.39	196.85	45.71
	RL	145.77	34.51	186.07	46.68
CVV	RM_{22}	135.99	34.45	178.45	42.38
	RM_{33}	135.87	33.18	169.03	47.86
	RL	141.49	32.57	180.58	43.83
CVO	RM_{22}	156.71	33.10	183.55	38.77
	RM_{33}	149.20	35.57	185.71	44.17

In order to examine the nature of the interaction effects, a series of post-hoc pairwise analyses using Bonferroni adjusted alpha levels of 0.004 per test (0.05/14) was conducted. With respect to the CVN syllables, significant differences between 22 and 33 (i.e. a mean difference of 8.14 Hz; $p = 0.004$) and between 33 and 35 (i.e. a mean difference of 11.47 Hz; $p = 0.001$) were found at the f_0 minima and significant differences between 22 and 33 (i.e. a mean difference of 6.17 Hz; $p = 0.004$) and between 33 and 35 (i.e. a mean difference of 10.78 Hz; $p < 0.001$) were found at the f_0 maxima. With respect to the CVV syllables, there is a significant difference at the f_0 peak between 33 and 35 ($p = 0.004$), but the difference between 22 and 33 did not reach the adjusted level of significance ($p = 0.011$). No significant difference in f_0 is found concerning the CVO syllables.

A two-way repeated measures ANOVA was used to examine the effect of tone and syllable type on pitch excursion time (i.e. time interval between adjacent f_0 minima and maxima in the target syllable). Only a significant main effect of syllable type was found ($F(2, 55)=6.37$, $p < 0.01$). Post-hoc analyses using Bonferroni adjusted alpha levels of 0.017 per test (0.05/3) show significant differences in excursion time between CVO and CVV ($p < 0.017$) and between CVO and CVN ($p < 0.01$). The mean excursion time in CVN syllables (157 msec.) were generally longer than that in CVV syllables (134 msec.), but the difference did not reach significance ($p = 0.11$).

Finally, in order to examine the effect of tone and syllable type on the slope of the pitch rise, the excursion speed of the rising pitch contour was calculated by dividing the excursion size by the excursion time. Excursion size was calculated as the pitch difference between adjacent f_0 minima and maxima in the target syllable in semitones (in order to make the data from individual speakers, especially across genders, more comparable). The excursion speed is fastest in the CVO syllable and the slowest in the CVN syllable. However, such differences did not reach statistical significance; a two-way ANOVA showed no significant main effects of tone and syllable or interaction effect between these factors.

3. DISCUSSION

The experimental results suggest that the surface realization of a derived rising tone is affected by the tonal specification of its paradigmatic neighbor. However, the extent of this paradigmatic effect is modulated by the shape of the host syllable; a significant effect of tonal neighbor is only apparent in nasal-final syllables and marginally so in vowel-final ones. In particular, the higher the tone of its paradigmatic neighbor, the higher the f_0 profile of the derived rising tone. The just-noticeable difference in fundamental frequency is about 1 Hz [4]. Thus the differences of 8 Hz at the f_0 minima and 6 Hz at the f_0 maxima are above the perceptual threshold of pitch perception. Given that there is no significant difference in pitch excursion speed, it can be concluded that the pitch range of RM_{33} is approximately 6~8 Hz above that of RM_{22}. Crucially, this difference matches closely the difference between mean f_0 of lexical 22 and 33 tones in Cantonese.

Note that the f_0 difference observed here cannot be explained purely as a result of a derived environment effect. While the differences between the f_0 minima and maxima values of the RM_{33} and the RL CVN syllables are approximately 10 Hz, no such significant difference is observed between the f_0 measures of RM_{22} and RL. The lack of an effect between RM_{22} and RL, which echoes the fact that the mean f_0 of lexical 22 and RL are generally comparable to each other, suggests that the significant differences between RM_{33} and RL are due to the tonal level of the paradigmatic neighbor of the RM_{33} syllables, rather than a simple difference in phonetic implementation between derived vs. lexical tones.

This state of affairs finds natural expression in connectionist/exemplar-based models of linguistic

representations. A category is defined by a "cloud" of exemplars (i.e. memorized tokens). Exemplars within the category are organized by similarity across any number of salient dimensions, producing internal structure in category space. Within an exemplar-based model of the production-perception feedback loop [5], the production of the *Pinjam* category label begins with the activation of the relevant exemplar cloud. Since most exemplar-based models [6, 7] assume some form of averaging across multiple exemplars in the construction of a production target, the activation of one exemplar cloud triggers the partial activations of closely related exemplars (e.g., the base exemplars). The partial activation of the base exemplars in turn exerts an influence back onto the *Pinjam* exemplars. The differences in mean f_0 level between RM and RL are thus a manifestation of this paradigmatic influence. Crucially, when a rising tone has no paradigmatic neighbors (i.e. RL), no deviation from the canonical f_0 profile is possible.

To be sure, this type of spreading activation approach to analogical influence between exemplars predicts that non-paradigmatic, but semantically or phonologically related exemplars might exert an influence in the construction of a *Pinjam* phonetic target since the activation of a derived word activate the base exemplars as well as exemplars of semantically/ phonologically related forms. While further research is needed to ascertain the magnitude of this type of non-paradigmatic influence, it is not surprising that the influence of the base exemplars is so robustly observed here given the multiple associations between the derived and base forms; derived and base forms are related in all three dimensions: phonological, morphological, and semantic. Note, however, that paradigmatic influence is not always apparent in speech production. When the duration of the host syllable is not sufficiently long, no analogical effect is found (recall that the paradigmatic effect is only apparent in CVN syllables, and not in CVV or CVO syllables, both significantly shorter in duration than CVN syllables [8]). Since the realization of f_0 is tightly tied to syllable rhyme duration [9], the lack of a tonal paradigm uniformity effect in CVO syllables, and only marginally so in CVV syllables is not at all surprising.

4. CONCLUSION

This study presented novel evidence in support of phonetic analogy. Previous reported cases of phonetic analogy have largely been restricted to languages within the Indo-European family (particularly from within the Germanic and Romance sub-branches). Thus, the discovery of phonetic analogy in a typologically different language (Cantonese is a Sino-Tibetan language) strengthens the case for its status as a genuine linguistic phenomenon. Second, while it has been claimed that "any sound property or any cluster of properties may give rise to paradigmatic leveling" [2: 332], the validity of such a claim is not at all obvious. Kim & Jongman [10], for example, found the manner neutralization in Korean coda obstruents to be complete (i.e. word-final coronals are phonetically realized as [t]). Thus, to be able to demonstrate empirically that tone is a possible "analogizable" feature is pivotal toward mapping the typology of analogically extendable non-contrastive phonetic features.

5. REFERENCES

[1] Fougeron, C. and Steriade, D. 1997. Does deletion of French schwa lead to neutralization of lexical distinctions? *Proc. Eurospeech 97* **2**: 943-946.

[2] Steriade, D. 2000. Paradigm uniformity and the phonetics-phonology boundary. In: M.B. Broe and J. Pierrehumbert (eds.), *Laboratory Phonology V*. Cambridge University Press: Cambridge. 313-334.

[3] Barnes, J. A. and Kavitskaya, D. 2002. Phonetic analogy and schwa deletion in French. *Proc. 27th Berkeley Linguistic Society*, 39-50.

[4] Moore, B. C. J. 1997. *An introduction to the psychology of hearing*. 4th ed. San Diego: Academic Press.

[5] Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In: J. Bybee and P. Hopper (eds.), *Frequency effects and emergent grammar*. John Benjamins: Amsterdam. 137-157.

[6] Wedel, A. B. 2006. Exemplar models, evolution and language change. *Linguistic Review* **23**: 247-274.

[7] Pierrehumbert, J. 2002. Word-specific phonetics. In: C. Gussenhoven & N. Warner (eds.), *Laboratory Phonology VII*. Mouton de Gruyter: Berlin. 101-140.

[8] Gordon, M. 200.1 A typology of contour tone restriction. *Studies in Language* **25**(3): 423-462.

[9] Zhang, J. 2001. *The effects of duration and sonority on contour tone distribution-typological survey and formal analysis.* PhD, Linguistics, UCLA.

[10] Kim, H. and Jongman, A. 1996. Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics* **24**: 295-312.