

C-COMMAND, PROJECTIONS AND MELODY: MICRO-PARAMETERS OF THE HAN-TEMPLATE

SHANTI ULFSBJORNINN

Department of Linguistics
School of Oriental and African Studies/University of London
Thornhaugh Street, Russell Square
WC1H 0XG London
United Kingdom
su1@soas.ac.uk

Abstract: (South-)East Asian languages often force highly restricted word-structures, leading standard Government Phonology (sGP) to the useful, though problematic, Han-template (regular/augmented) (Beijing Mandarin (Goh 1996), Thai (Denwood 1999), Hongkui To (Xu 2001), and Vietnamese (Ulfsbjorninn 2008b)). We replace it with an obligatory c-command relationship between nuclear heads (xN_1 , xN_2), resulting in xO_2 being obligatorily c-commanded by xN_1 (*V:C). In contrast, the augmented Han-template has obligatory c++2-command, hence, xO_2 is not obligatorily c-commanded by xN_1 (V:C). Concomitantly, *V:C (unlike V:C) languages prohibit l/U in xO_2 . In sGP no non-arbitrary link can be suggested; in GP 2.0, it emerges from micro-parametric c-command-conditions acting on word-structure.

Keywords: Han-template, c-command, phonology, Sino-Tibetan, Tai-Kadai

1. The Han-template

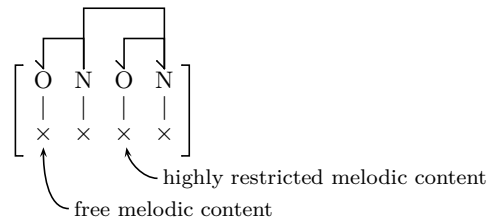
In Beijing Mandarin major-class lexical items may be of the shapes shown in (1). To account for these facts, Goh (1996) proposed that Beijing Mandarin word-domains are highly selective for both their melodic and constituent content. The Mandarin word is said to be comprised, minimally and maximally, of a template of two ON pairs ($O_1N_1O_2N_2$), excluding (1a,k). No constituent is permitted to branch, excluding (1f, 1h–j). Melodic material can be base generated either in O_2 or N_2 but not both, excluding (1g). N_1 is the head of the domain; O_1 , the onset

licensed by the head of the domain, is free for its melodic content (Xu 2001), while only three consonants could be found in O_2 /n, ŋ, ʈ/.

(1) Mandarin word-shapes¹

- (a) *CV
- (b) V: /i:/ ‘year’
- (c) VC /an/ ‘peace’
- (d) CV: /ma:/ ‘horse’ /jy:/ ‘fish’
- (e) CVC /foŋ/ ‘bear’ /lan/ ‘wolf’
- (f) *CV:C
- (g) *CVCV
- (h) *CVCC
- (i) *CCVC
- (j) *CCVCC
- (k) *CVCVCV⁽ⁿ⁺¹⁾

(2) Han-template (Goh 1996)

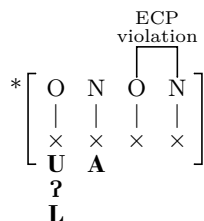


Violation of melodically empty *CV is independently derived from the phonological empty category principle (ECP) (Kaye 1990), which states that a non-p-licensed empty category must be given a phonetic interpretation. The principles surrounding p-licensing forbid the p-licensing of two contiguous empty categories. Therefore, “short, open syllables” are ungrammatical as they necessarily result in two contiguous empty categories. CV: or CVC, on the other hand, are licit (compare (3a) with (3b,c)).

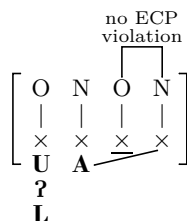
¹ We deliberately omit discussion of Mandarin on/offglides discussed in Živanović–Pöchtrager (2010); as will become apparent later, they do not contribute to the eventual discussion. The reader will see, though, that none of our claims interfere with their described behaviour or representation in *ibid.*

(3) ECP and the Han-template

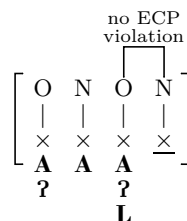
(a) */ma/



(b) /ma:/ 'horse'



(c) /lan/ 'wolf'



2. Problems for the Han-template

2.1. Universal grammar and the Han-template

One major problem with the Han-template is that it groups languages into those which are Han-templatic and those which are not.

(4) Phonological template parameter Xu (2001)

Templatic? ⟨yes, no⟩

(4) is not a true parameter as it must be broken down into several independent parameters to be meaningful. The Han-template comes in at least two forms, regular: [ONON] and augmented: [ONONON] (which we will discuss in detail later). The existence of the augmented template, however, shows that what the parameter in (4) actually regulates is how many ON-pairs may minimally and maximally be present in a word domain. And, if we can have two or three (or four, Xu 2001), why not six?

(5) Phonological template parameter

- (a) Word-domain must contain minimally and maximally 1 ON pair (yes, no)
- (b) Word-domain must contain minimally and maximally 2 ON pairs (yes, no)
- (c) Word-domain must contain minimally and maximally 3 ON pairs (yes, no)
- (d) Word-domain must contain minimally and maximally 4 ON pairs (yes, no)
- ... $n + 1$

The parameter in (5) totally ignores the fact that all Han-templatic languages, augmented or not, revolve around two active nuclei (in that

they license onset material), languages produced by the hypothetical (5a) simply do not exist. Han-templatic languages fit across (5b) and (5c). Languages which demonstrate templatic effects and allow up to four ON-pairs are, we believe, of the Semitic type. Interestingly, if languages allow more than four ON-pairs in a word-domain, most probably, they will not show any templatic effects whatever.

The templatic parameters (5a–d), contravene the spirit of sGP as the prohibition on (for example) [ONONON] templates in Mandarin has nothing to do with interrelations of the constituents within the phonological domain, rather, it is entirely indirectly motivated.

Similarly, (5) does not behave like other parameters of sGP. Take the parameter that explains why words in some languages may end in consonants: the word-final parameter (6). sGP understands this parametric difference as a p-licensing property of the word-final nucleus.

(6) Word-final parameter (Kaye 1990)

Domain-Final Empty Categories are P-licensed ⟨yes, no⟩

If this parameter is set to ⟨yes⟩, words in these languages may show surface forms which end in consonants; if set to ⟨no⟩, then the surface forms in these languages must end in vowels. The word-final parameter is particularly analytically sound as it makes no direct reference to surface effects while it makes direct reference to the constituents within the word-domain. Compare this to (7) which, like (4), makes reference to the surface state in abstraction from internal states.

(7) Ad-hoc alternative to the word-final parameter (to be discarded)

Words can end in consonants ⟨yes, no⟩

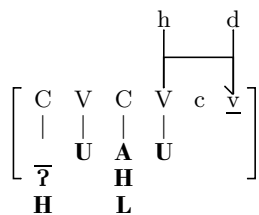
As people, we notice that (7) captures the exact same facts as (6). However, as theoreticians we should always opt for (6) over (7). They are not equal because empirical coverage is not the ultimate aim of a formal theory of language. It is a necessary but not sufficient condition, unlike the necessary and sufficient conditions of a theory of language: economy of theory, and rigour of formalism (Chomsky 1999). With these in mind, we cannot help but notice the arbitrary nature of a parameter which only makes reference to surface states (cf. (5), (7)).

2.2. Principled foundations of the Han-template

We pointed out in **2.1.** that no languages have a template consisting of a single ON-pair and that if a language had more than three ON-pairs it must be of the Semitic type or not show any templatic effects at all. This points to the principled foundation of the Han-template being the interaction of two active nuclei. For this reason we invoke Charette (2008).

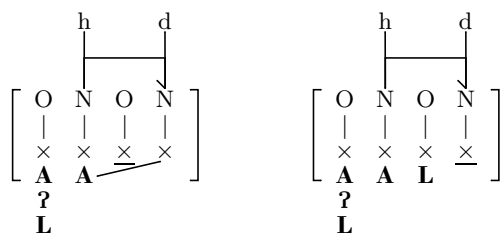
To understand Turkish stress and phonotactics of licit consonant clusters, Charette (2008) posits that all Turkish words must minimally include an exclusively binary, nuclear, dependence relationship where the head (h) is left of its dependent (d) and where the dependent may not be base generated with any melodic material: /kara^h-cv^d/ ‘black’, /e^hv-cv^d/ ‘house’ (8).

(8) /kuzú/ ‘lamb’



This h–d structure, which must be present in any Turkish word, allows the Han-template facts to directly fall out of it. The minimal word of Turkish will be minimally (C)VC or CV: . The forms *CVCV, or *CV forms are barred (e.g., see 8, 9). The only difference between Turkish and Mandarin therefore, is that (9a,b) are not just the minimal but also the maximal word-structures.

(9) (a) /da:/ ‘mountain’ (b) /day/ ‘mountain’ (Cypriot Turkish)



This insight allows us to create a non-ad-hoc templatic parameter for Beijing Mandarin. Word-domains must minimally and maximally contain a nucleus which is head to an obligatorily lexically-empty dependent to its right. Part of what this allows us to do is make the Han-templatic languages look less exotic than they used to. The Han-template behaves no differently to most languages' minimal word, and so it shares an identical structure to most languages' minimal word. Beijing Mandarin is consequently only different from the languages listed in (10a) for not allowing word-domains which are larger than this. Although clearly not very explanatorily adequate, we have at least been able to group templatic and non-templatic languages with our templatic parameter.

(10) Templatic parameter

- (a) Minimally all word-domains are h-d
 Yes: English, Mandarin, Dinka
 No: ?Amele, ?Chama, ?Warao
- (b) Maximally all word-domains are h-d
 Yes: Mandarin, Hmong
 No: English, Yoruba, Zulu

We will now show that the nuclear relationship referred to in Charette (2008) fails due to the architecture of the theory it was devised in.

2.3. The limits of sGP in explaining the Han-template

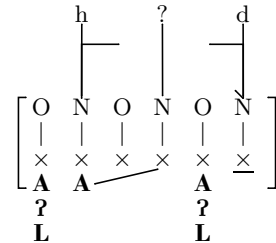
Mandarin is micro-parametrically different from Cantonese (Barrie 2003, 5), Hongkui To, Thai, Lung-Ming. These all allow long vowels to be found before their word-final consonants. They share all word-shape restrictions but for allowing (1e) CV:C. This leads to positing the augmented template.

- | | |
|---|--|
| (11) HongKui To (Xu 2001) | (12) Thai (Denwood 1999; Roengpitya 2000) |
| (a) /ŋu:m/ 'cave' | (a) /cip ⁷ / 'sip' |
| (b) /β ³ o:k ⁷ / 'flower' | (b) /ci:p ⁷ / 'pleat' |
| (c) /ti:n/ 'know' | (c) /mâ:k ⁷ / 'very' |
| (d) /ɬa:ŋ/ 'body/health' | (d) /jà:k ⁷ / 'from' ² |

² Low (/à/) and falling (/â/) tones.

The augmented template can be seen as a Han-template with a recessive ON pair between the standard two ON pairs: $O_1N_1O_aN_aO_2N_2$. In this augmented template, no conditions on O_1N_1 and O_2N_2 differ from the regular template. O_aN_a may only be associated to melody which is spread from the nucleus to its left, and the onset in O_aN_a may never be phonetically interpreted. We will come back to these facts momentarily in building our replacement for the Han-template, but to end this section, we note that the augmented template causes havoc for a Charette (2008)-inspired sGP analysis.

(13) h-d on augmented template



The head of the domain is always, indisputably, N_1 . However, in the augmented template we have three nuclei. To keep our h-d story, we must either abandon binarity of the nuclear relationship, or suppose that the O_aN_a pair is somehow invisible to the nuclear dependence relationships despite being of the same category. The former means we must go backwards and reject the notion that a relationship between binary nuclei is the building block of the template, and the latter introduces a licit yet apparent violation of locality (or a specifically locality violating mechanism such as ‘burial’ of Vs in Scheer 1996). Most damagingly of all, the number of ON pairs in the Han-template no longer falls out from a strong theory of nuclear dependence in word-domains that we can also tag to minimal words of most languages (see 14).

(14) Templatic parameter

(a) Word-domain contains min/max two ON pairs

Yes: Mandarin, Shanghainese

No: go to (14b,c)

(b) Word-domain contains min two ON pairs

Yes: Cantonese, Hongkui To

No: go to various...

- (c) Word domain contains max three ON pairs
 Yes: Cantonese, Hongkui To
 No: go to various...

We will show how GP 2.0, with its arboreal structure, will incorporate the key observation that the Han-template is about a nuclear interaction.

3. The ontology of the Han-template

We adopt the architecture of GP 2.0 (Pöchtrager 2006; Ulfsbjorninn 2008a; Živanović–Pöchtrager 2010). For an easy-to-follow introduction to the framework see Kaye–Pöchtrager (2009) (as severe space limitations make a representative overview impossible). In this architecture, all ‘syllabic positions’ are terminal nodes of their nuclear or onset heads (xN) and (xO). Interaction between terminals is dictated by c-command-types and m(elodic)-command (see Živanović–Pöchtrager 2010). Terminal nodes may also be comprised of an un-annotated x which interacts with other terminal nodes and heads in the structure.

We posit that the Han-template describes a structure of which a formal condition is a binary, xN relationship where xN_1 and xN_2 are in a c-command relationship (15).

- (15) Phonological c-command (Živanović–Pöchtrager 2010)
 Node α c-commands β iff
 α is β 's sister, or
 α 's sister contains β .

The formal definition of Mandarin word-domain is laid out in (16).

- (16) Mandarin word-domain
 $xN\{__\}$ must c-command $xN\{I,U, x, x-x, x\rightarrow x\}$ ³
- (17) Paucity condition
 Lexical domains contain as little phonological material as possible.⁴

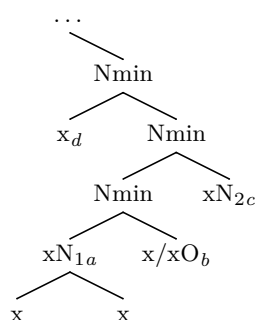
(16) adds to Živanović–Pöchtrager (2010)'s understanding of the rhyme in Mandarin. Added to their structure: xN_1 must be c-commanded by xN_2 .

³ x is /i/, x-x is /ə/, x→x is /a/ (see Kaye–Pöchtrager 2009).

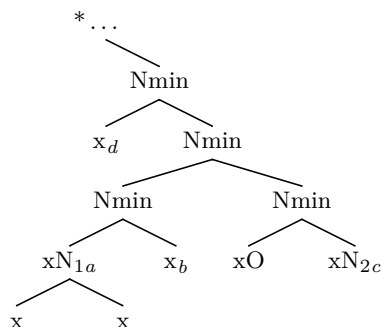
⁴ A general condition on trees in GP 2.0 prohibiting the proliferation of functionally useless structure.

Consequently, as one can see from (18), this means that xN_2 's (labelled "c") projection intervenes between the onglide's terminal node (d) and the rest of the structure. Crucially, this does not affect the c-command and binding conditions explored in their paper; despite the intervening xN_2 projection, the onglide (d) still c-commands all positions contained by its sister.

(18) The rhyme in Mandarin



(19) Illicit rhyme in Mandarin



Crucially, (18) is a well-formed rhyme because c c-commands a , and therefore satisfies (16). Interestingly, because our condition is that xN_2 must c-command xN_1 , it is not possible to posit the structure in (19). This creates a prediction and/or the possibility of a problem. (16) forces the offglide and the final onset (if there is one) to share a projection, meaning they must be in complementary distribution. To represent an offglide preceding xO_2 we would have to arbitrarily posit xO_2 in a projection higher than the offglide, but we do not want to arbitrarily add nodes, so the only place where xO_2 could realistically sit following an offglide is as a sister to xN_2 . This describes structure (19) and is illicit due to the fact that xN_2 no longer c-commands xN_1 (cf. (16)). In short, we predict no reasonable place for xO_2 to sit following an offglide. This prediction is fully borne out. Offglides never occur before final consonants in Mandarin (see a complete table of rhymes in Živanović–Pöchtrager 2010).

3.1. Micro-parametric differences in Han-templates

We will show shortly how the structure for the augmented Han-templates result from micro-parametric difference between Mandarin-like

languages and Cantonese-like languages, but before we propose that difference we wish to highlight another micro-parametric difference between these languages: the melodic qualities of xO_2 .

(20) Correlation of parametric differences in Han-templates

	Shanghainese	Beijing Mandarin	Guiliu Mandarin		
Allow V:C?	no	no	no		
Final C's ⁵	ʔ, ŋ	n, ŋ, ʈ	n, ŋ		
	Cantonese	HongKui To	Thai	Lung Ming ⁶	
Allow V:C?	yes	yes	yes	yes	
Final C's	p ⁷ , t ⁷ , k ⁷ m, n, ŋ, j, w ⁷	p ⁷ , t ⁷ , k ⁷ m, n, ŋ ⁸	p ⁷ , t ⁷ , k ⁷ m, n, ŋ	t ⁷ , k ⁷ ?p ⁷ m, n, ŋ	

Initially, the xO_2 consonants of languages which allow V:C do not seem to constitute a 'natural class'; however, there is a striking asymmetry between what can occur at the end of words in *V:C and V:C languages. All the above languages allow nasal structures as xO_2 ; however, *V:C languages do not allow final /m/. In fact, in *V:C languages none of the many xO_1 that contain elements **I** or **U** are tolerated in xO_2 . This is in stark opposition to V:C languages, none of which show this restriction.

In standard GP or Strict-CV this observation cannot even be adequately formalised. The association of melody to nuclei is entirely divorced from the state of a following onset. There can be no principled theoretical link between prohibiting I and U elements in O_2 and simply being a language that uses the augmented (O_aN_a) Han-template (21, 22). Structurally speaking, O_2 is in an identical environment in both (21) and (22), O_2 comes after a nucleus which linearly precedes it.

In stark contrast, GP 2.0 can differentiate this environment/structural condition because arboreal models may invoke c-command. So what is the difference between the Han-template and the augmented Han-template in GP 2.0?

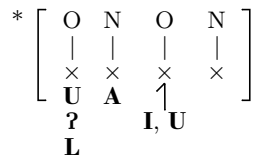
⁵ Counting only xOs , not offglides (which are **I** and **U** annotated on skeletal points) (Živanović-Pöchtrager 2010).

⁶ Central Tai language spoken in Guangxi, China (Ryuichi 1997). Virtually identical data for Siamese and Saek.

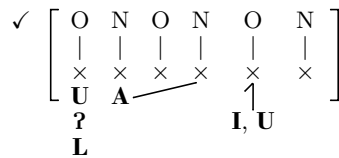
⁷ Data from Barrie (2003). As Onsets when these occur after long vowels (*ibid.*).

⁸ Data from Xu (2001), not an exhaustive, but illustrative list.

(21) Han-template ***I**, **U** in O_2



(22) Augmented **I**, **U** in O_2

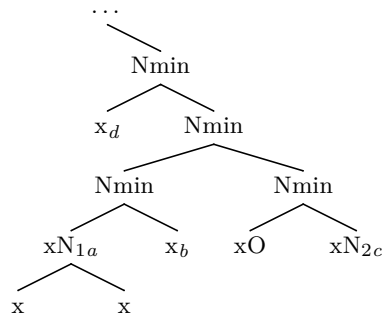


As we have already seen, structure (19) is illicit in Mandarin because xN_2 must c-command xN_1 . We claim, however, that (19) is the augmented template’s structural identity. We claim that the parametric difference between *V:C and V:C languages is that in V:C languages xN_2 must (at least) c++2-command xN_1 .

(23) c++2-command (Živanović–Pöchtrager 2010)

Node α c++2-commands α iff
 α is the highest terminal in the maximal projection immediately containing α ,
 and α ’s mother c-commands β .

(24) “Augmented” template type



In (24), c c++2-commands a . So the word-domain is satisfied. Unlike in Mandarin, however, this leaves b available to produce vowel length before xO_2 . The formal difference between the regular Han-template and the augmented Han-template is the different conditions of c-command between their word-structure’s nuclear heads. It falls out directly from structure (24) that xO_2 need not be c-commanded by xN_1 . While in Mandarin, to have an xO_2 in structure (18) leaves you no option but to place it in b . If it was placed as a sister of c , this would violate Mandarin’s c-command condition. In GP 2.0 these environments of xO_2 across *V:C and V:C languages are distinct—allowing, at the very least, the possibility of differing conditions on xO_2 across these two language types. Although

the underlying principle is not well understood, we propose the condition (25) on xO_2 in *V:C languages.

(25) **I** and **U** annotated in an onset head (xO) may not be c-commanded by xN .

4. Conclusion

We have shown how sGP's Han-template is irreconcilably stipulative as it struggles to be understood by a simple interplay of non-ad-hoc parametric settings. We replaced this with an arboreal structure where two nuclear heads xN_1 and xN_2 are in an obligatory c-command relationship. We then linked two seemingly unrelated micro-parametric differences between Mandarin-like languages (*V:C) and Cantonese-like languages (V:C). The former forbid vowel length before xO_2 , and forbid **I** or **U** to be annotated to xO_2 . We showed that sGP, due to its architecture, cannot even formalise a link between these two facts, as in both cases xO_2 occurs in an identical environment (they both linearly follow an interpreted (unlicensed) nucleus (whether this be N_1 or N_a) and both linearly precede a licensed domain-final nucleus). We argue, however, that to explain the regular from the augmented template in GP 2.0 necessarily allows for xO_2 in *V:C languages to be in a different environment from xO_2 in the V:C languages. The former is obligatorily c-commanded by xN_1 while the latter is not. We claim therefore that **I** and **U** may not be annotated to xOs c-commanded by xN_1 . Although the mechanism for this condition is not understood, the argument in this paper brings us closer to understanding the Han-templatic languages than any previous GP analysis as it can, at the very least, formalise a link between our two micro-parametric differences.

References

- Barrie, Mike 2003. Contrast in Cantonese vowels. In: Toronto Working Papers in Linguistics 20: 1–19.
- Charette, Monik 2008. The vital role of the trochaic foot in explaining Turkish word endings. In: *Lingua* 119: 46–65.
- Chomsky, Noam 1999. An interview on Minimalism. Unpublished manuscript, edited by Adriana Belletti and Luigi Rizzi, University of Siena.
- Denwood, Ann 1999. A Government Phonology approach to Thai tones: Observation and hypothesis. In: *SOAS Working Papers in Linguistics and Phonetics* 9: 173–90.

- Goh, Yeng-Seng 1996. The segmental phonology of Beijing Mandarin. Doctoral dissertation. SOAS–University of London, London.
- Kaye, Jonathan D. 1990. ‘Coda’ licensing. In: *Phonology* 7: 301–30.
- Kaye, Jonathan D. – Markus Alexander Pöchtrager 2009. GP 2.0. Paper presented at the GPRT6, Piliscsaba, Hungary, April 25, 2009.
- Pöchtrager, Markus Alexander 2006. The structure of length. Doctoral dissertation, Universität Wien.
- Roengpitya, Rungpat 2000. A study of vowel length in Thai. In: *Journal of the Acoustical Society of America* 107: 2805.
- Ryuichi, Kosaka 1997. A phonology of the Bu Dai (Thu Lao) language. In: *Journal of Asian and African Studies* 53: 137–53.
- Scheer, Tobias 1996. Une théorie de l’interaction directe entre consonnes. Contribution au modèle syllabique CVCV, alternances e–∅ dans les préfixes tchèques, structure interne des consonnes et la théorie X-barre en phonologie. Doctoral dissertation, Université Paris 7, Paris.
- Ulfsgjorninn, Shanti 2008a. The formal dynamism of categories: Stops vs. fricatives, primitivity vs. simplicity. In: *Poznań Contemporary Studies in Linguistics* 41: 121–146.
- Ulfsgjorninn, Shanti 2008b. N''' = domain. Paper presented at CUNY Phonology Forum on the Syllable, CUNY.
- Xu, Zhuo 2001. Phonological analysis of Hongkui To. Doctoral dissertation, Guangdong University.
- Živanović, Sašo – Markus Alexander Pöchtrager 2010. GP 2.0 and Putonghua too. In: *Acta Linguistica Hungarica* 57: 357–80.