Notational Equivalence in Tonal Geometry

Chris Oakden

2020

Main idea: The tonal geometries of Yip (1989) and Bao (1990) are not unique representations and are instead notationally equivalent.

Supporting data: Model theoretic analyses of assimilatory tone sandhi in Pingyao (register) and Zhengjiang (countour).

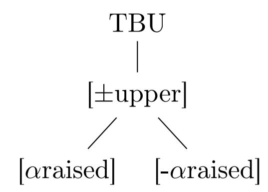
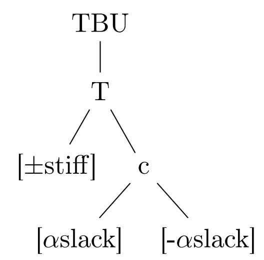
Paper Summary

Introduction (p. 1)

- The tonal representations of Yip (1989) and Bao (1990) have been claimed to vary in their empirical coverage (specifically assimilatory tone sandhi processes in Chinese).
- This view stems from a derivational view of spreading and delinking (despite these two mechanisms not being enough to account for all attested tone sandhi).
- He shows that a copy mechanism is also necessary.
- Oakden uses a computational framework (model theory) to compare the input-output mappings of each representation and finds that they both handle the relevant assimilatory tone sandhi process equally well.
- Furthermore, he shows that the two representations can be freely translated between each other without loss of contrast.
- *****Therefore, he concludes that the representational proposals "do not constitute distinct theories, but are instead notationally-equivalent."*****

A Notion of Notational Equivalence (pp. 1 – 5)

- *Conditions for Notational Equivalence* (Fromkin, 2013)
  - Two models do not differ in their empirical predictions
  - Two models represent the same set of abstract properties, differ only superficially
-

Yip (1989)	Bao (1990)
	

- Refer back to McCarthy (1988) for assumption as to why these predict different types of assimilation.
- Spread (addition of association line between elements) and delink (deletion of an association line) are insufficient to model tone spread because spreading of a contour tone requires the extra assumption of tier conflation.
- Tier conflation is already in the system, so we can get the same mapping with a copy analysis rather than a spreading analysis (see examples (3)-(4) on p.3).
- Derivational vs. serial accounts also make different predictions. To abstract away from this, we can analyze the input-output mappings themselves to avoid theory-specific mechanisms.
- Model theory analysis will show that both theories use QF transductions to explain register assimilation and contour assimilation.
  - Copy mechanism does not overextend the "spirit" of the original theory.
- Both theories are also bi-interpretable (but uses a different definition of bi-interpretability from Strother-Garcia (2019)).

Yip and Bao Models (pp. 5 – 9)

- See (6) on p. 6 for how both models represent 8 tones: L, H, M_1, M_2, HM, MH, ML, LM.
  - Visually, we can already see the similarities. The Bao (1990) model splits the [±u] node from Yip (1989) into three nodes: T, [±u], and 'c'. In the rest of the paper he therefore refers to these as the *spread* (Bao) and *bundled* (Yip) models.
- Focus on spreading and delinking have resulted in claims that the bundled model cannot handle contour or register spread while the separated model can handle both. See examples (10-11) and (13-14) on pp. 8-9 if still uncertain why this is the case.
- Register spread in Pingyao:

tʰuæ	pang	'quit class'	
35	13	base form	/MH.LM/
13	13	sandhi form	[LM.LM]
tɕi	ma	'ride a horse'	
13	53	base form	/LM.HM/
35	423	sandhi form	[MH.HM]

- Contour spread in Zhenjiang:

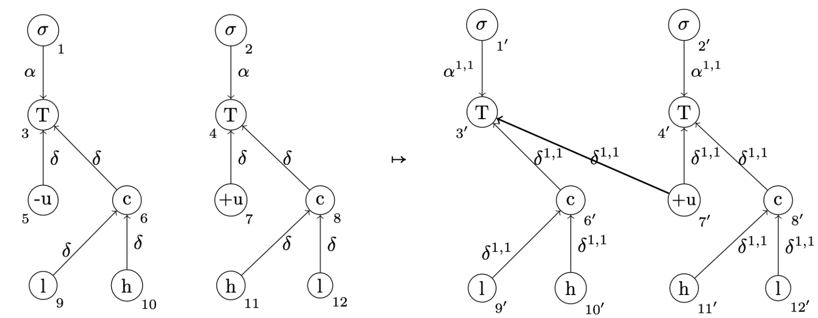
lɛ̃n	to	'lazy'	
31	55	base form	/ML.H/
22	55	sandhi form	[M.H]
ɕĩ	huei	'virtuous'	
35	55	base form	/LM.H/
22	22	sandhi form	[M.H]

Graph Mappings: Empirical Predictions (pp. 9 – 19)

- Relations used in each model:

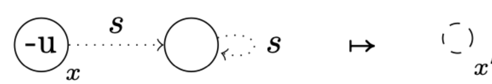
Bundled Model Relation	Separated Model Relation	Label
P_σ	P_σ	syllable
P_{+u}	P_{+u}	+u register
P_{-u}	P_{-u}	-u register
P_h	P_h	h terminal
P_l	P_l	l terminal
	P_T	'T' root node
	P_c	'c' terminal node

- Functions used in each model:
  - α = **association** (between syllable node and root node)
  - δ = immediate **dominance**
  - s = **successor**
- Full models for a [L.MH] sequence shown in (16-17); successor edges omitted in all further visualizations.
- A transduction Γ from input to output is defined logically over connected substructures of the input model. These are formally given in the appendix, and visually explained in the body of the paper.
- **/LM.HM/→[HM.HM] (Separated Model):**

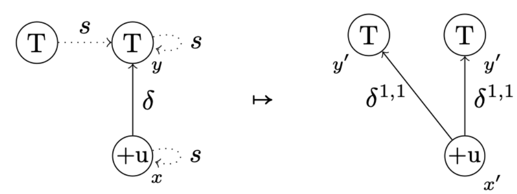


- Two changes in substructures:

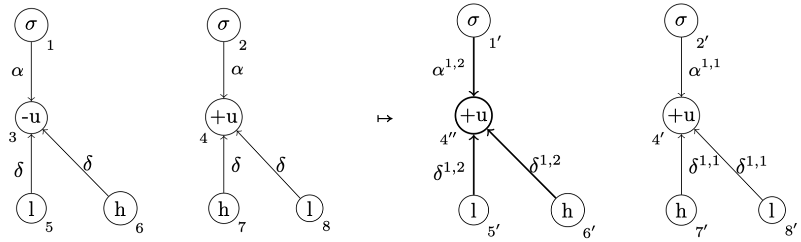
1. (Nodes 5,7)



2. (Nodes 3,4,7)

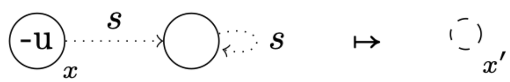


- **/LM.HM/→[HM.HM] (Bundled Model):**

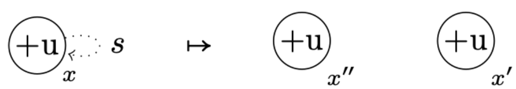


- Two changes in substructures:

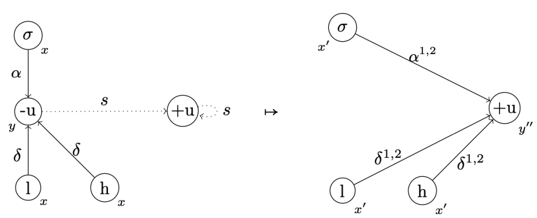
1. (Nodes 3,4)



2. (Node 4)

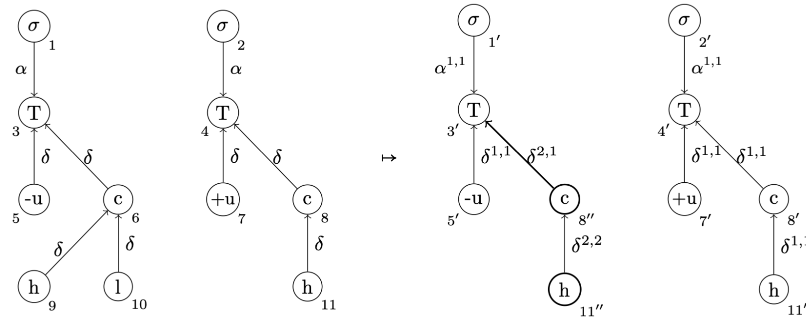


3. (Nodes 1,3,4,5,6)



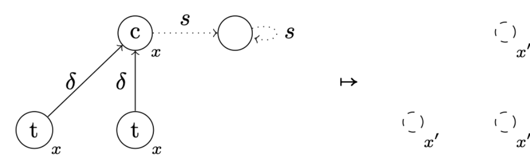
- "The bundled model requires non-size-preserving QF logic to model register assimilation because it has the emulate spreading as deletion *plus* copying. For cases of this type, the separated model captures the process using more restrictive, size-preserving logic" (p. 16).

- **/ML.H/→[M.H] (Separated):**

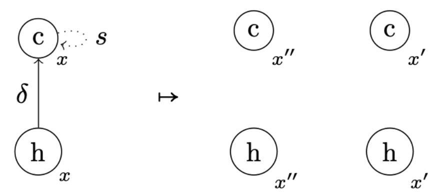


- Four changes in substructures:

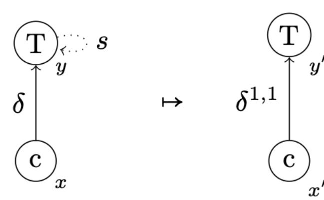
1. (Nodes 6,9,10)



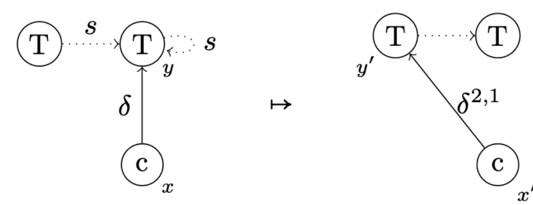
2. (Nodes 8,11)



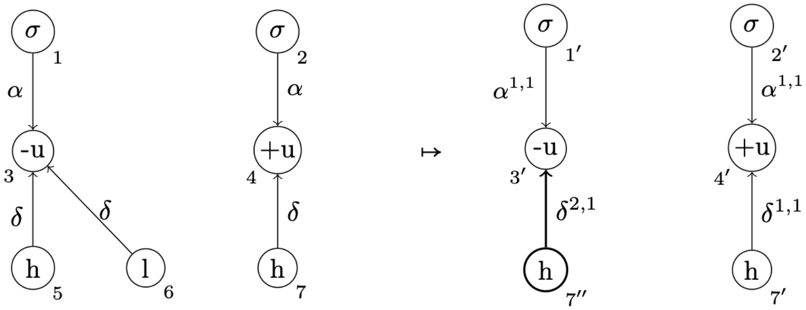
3. (Node 4)



4. (Nodes 3,4,8)

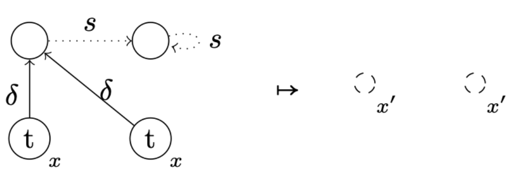


- **/ML.H/→[M.H] (Bundled):**

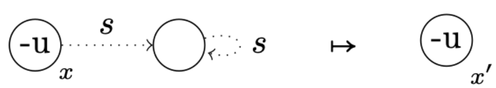


- Four changes in substructures:

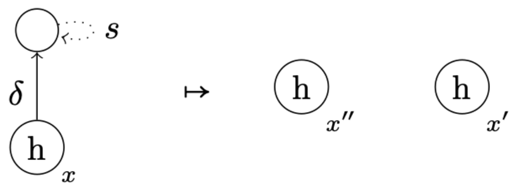
1. (Nodes 3,4,5,6)



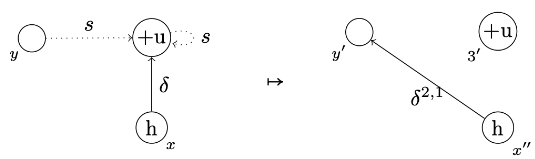
2. (Nodes 3,4)



3. (Nodes 4,7)



4. (Nodes 3,4,7)



- "This section has shown that both models *can* represent these processes when formalized as mappings over graph structures, and that they do so with the *same* logical complexity threshold...From the computational perspective, bundled and separated models of representation satisfy [the first condition] of notational equivalence as they do not differ in their empirical consequences" (p. 19).

Graph Mappings: Structural Differences and Bi-interpretability (pp. 19 – 24)

- Definition of model bi-interpretability (Friedman and Visser, 2014):

  We note that an interpretation $K : U \to V$ gives us a construction of an internal model $\widetilde{K}(\mathcal{M})$ of U from a model M of V. We find that U and V are bi-interpretable iff, there are interpretations $K : U \to V$ and $M : V \to U$ and formulas F and G such that, for all models $\mathcal{M}$ of V, the forumla F defines an isomorphism between $\mathcal{M}$ and $\widetilde{M}\widetilde{K}(\mathcal{M})$, and, for all models $\mathcal{N}$ of U, the formula G defines an isomorphism between $\mathcal{N}$ and $\widetilde{K}\widetilde{M}(\mathcal{N})$.

- For this paper: **IF** the model signature for the separated model is equal to the outcome of the separated model being translated into the bundled model and then back into the separated model **AND** the model signature for the bundled model is equal to the outcome of the bundled model being translated into the separated model and then back into the bundled model **THEN** the two models are bi-interpretable.

- Γ^{sb} is the transduction from separated to bunched and Γ^{bs} is the transduction from bunched to separated. (See (42) on p. 20 for visualization).

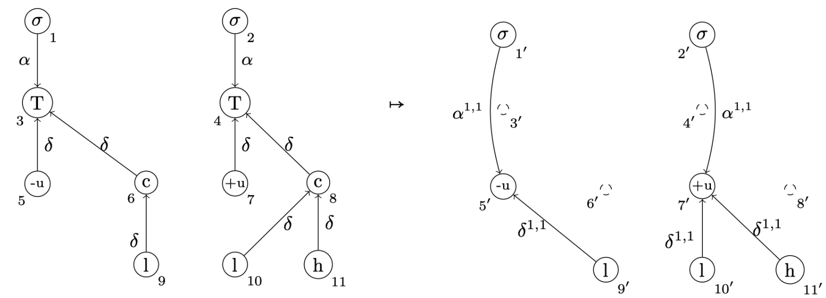
- Is $\mathcal{M}_s = \Gamma^{bs}(\Gamma^{sb}(\mathcal{M}_s))$? ...yes

- Is $\mathcal{M}_b = \Gamma^{sb}(\Gamma^{bs}(\mathcal{M}_b))$? ...also yes

- The two models are therefore bi-interpretable.

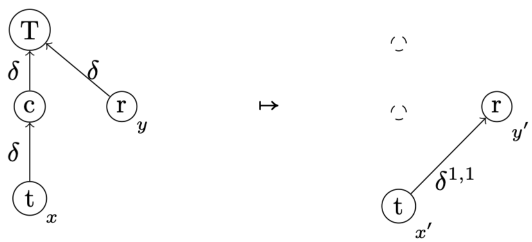
- I show the individual transductions below, but see Section 5.2 (pp. 23–24) and the appendix if interested in the details on bi-interpretability.

- **Separated to Bundled:**

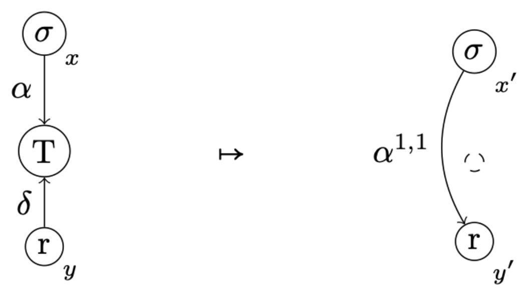


- Two changes in substructures:

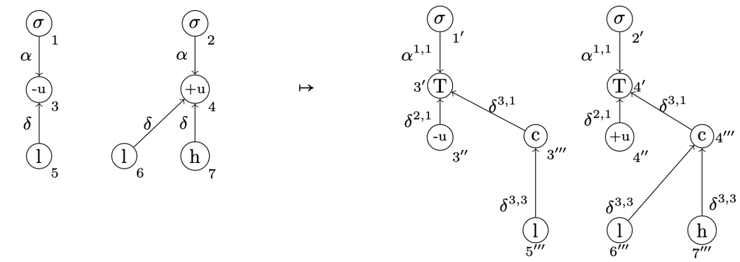
1. (Nodes 3,5,6,9/4,7,8,{10,11})



2. (Nodes 1,3,5/2,4,7)

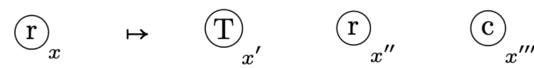


- **Bundled to Separated:**

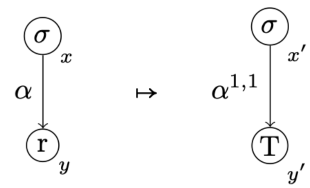


- Five changes in substructures:

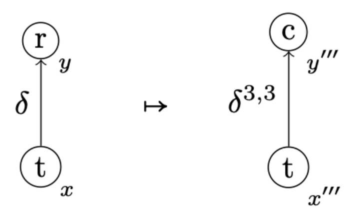
1. (Nodes 3/4)



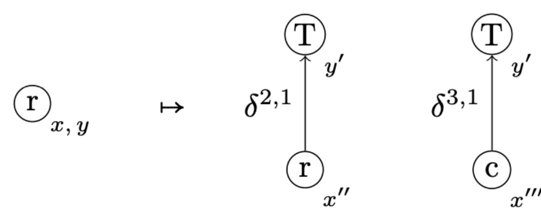
2. (Nodes 1,3/2,4)



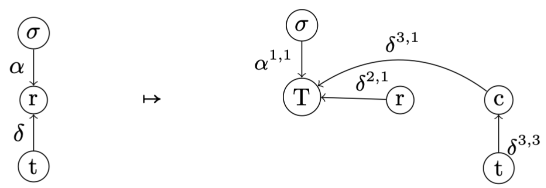
3. (Nodes 3,5/4,{6,7})



4. (Nodes 3/4)



5.



(Nodes 1,3,5/2,4,{6,7})

Discussion (pp. 24 – 29)

- Size preserving QF allows spreading while non-size-preserving allows spreading + copying.
- Copying is necessary and is in fact present in feature geometric theories due to "tier conflation".
- Why is it necessary?
  - Imagine a falling tone spreads from one syllable to the following two syllables (see (53) on p. 25). Without tier conflation, it is predicted that the single falling pattern would align over all three syllables. That is not what seems to happen. Each syllable gets an identical falling pattern. This means that the original tone contour had to be copied onto the next two. So copying is already there (though not by direct name!).
- Furthermore, the process of Spread → Delink → Tier Conflation results in the same input-output mapping (with the same structure) as Copy → Delink → Re-associate.
- Oakden takes this as evidence that these copying analyses preserve the spirit of the original analyses.
- The computational analysis shows that spreading with tier conflation and copying are *formally* indistinguishable since the realize the same map.
- "Determining whether this generalizes to other processes (including those for which other rules intervene between spreading and tier conflation) defined over these representations or others is beyond the scope of the current paper" (p. 27).
- This paper also used a new definition of bi-interpretability. Previous definitions are closer to what he calls *mutual interpretability* with a definition from Enayat and Wijksgatan (2013).
  - *Suppose U and V are first order theories. U is interpretable in V, written $U \trianglelefteq V$, if there is an interpretation $\mathcal{I} : U \to V$. U and V are mutually interpretable when $U \trianglelefteq V$ and $V \trianglelefteq U$.*

Conclusion (pp. 29 – 30)

- "The first result is that the models do not differ in their empirical predictions as previously claimed."
- "...the second result is a proof that any structural difference between the representations is superficial."
- "The purpose of this paper is not to propose a new tonal model or advocate one model over another. Instead, its aim is to establish a formally-rigorous procedure for determining whether two competing models comprise two distinct theories of representation."