Natural class-preserving transductions among phonological representations

Nick Danis nsdanis@wustl.edu

Workshop on Model Theoretic Representations in Phonology Stony Brook University, Septepmber 22-24, 2022

Washington University in St. Louis





consonants and vowels probably should be natural classes



consonants and vowels probably *shouldn't* be natural classes





big questions

- how can we formally compare phonological representations?
- what can we learn from these comparisons?
- what does it mean for us as linguists?
- what does *anything* mean?



medium answers

- two theories can be shown to be formally equivalent using logic and model theory
 - given two representations A and B, a transduction between A and B means that any linguistic rule given with structure A can be translated into structure B, and vice versa
 - Strother-Garcia (2019), Danis & Jardine (2019), Oakden (2020), a.o.





(Strother-Garcia 2019)

medium answers

- not every transduction preserves ideas of *linguistic* equivalence
 - process should respect natural classes, which may be lost in certain transductions
- the property of a natural-class preserving transduction is defined to find those logically equivalent representations that also share linguistic intuitions





introduction definitions example transduction empirical basis existing transductions conclusion

natural-class preserving transductions



- A transduction between two representations A and B is **natural-class preserving** iff the set of all natural class extensions of A exactly match those of B
 - A natural class extension is a maximal set of segments that share some common structural property

nsion: { i u }	shared structure: [+high]
u o	every segment in this set is [+high] no other segment in the theory is [+high]
a	

assimilation: sharing is caring

- assimilation operates over like things
 - Trubetzkoy (1969), Chomsky and Halle (1968), Hyman (1974), Hayes (1986), Clements & Hume (1995) a.o.



Naturalness of Assimilation 😽

the targets and triggers of an assimilation process should constitute a natural class extension

assimilation: sharing is caring



• "Nevertheless, there is empirical evidence in favor of imposing a limitation on the use of variables with different features in different segments. The great majority of examples involve only a single feature, and in other cases there clearly seems to be some intrinsic connection between the features involved in the process of assimilation. At the present juncture, however, we are in no position to formulate these restrictions." (SPE 352)

assimilation: sharing is caring



- Clements & Hume (1995):
 - "Phonological rules perform single operations only." (p. 250)
 - "In the present model, in contrast, assimilation rules are characterized as the association (or "spreading") of a feature or node F of segment A to a neighboring segment B..." (p. 258)
- If assimilation is the result of spreading, then it directly follows from this that the resulting segments will have shared structure and therefore constitute a nontrivial natural class

introduction definitions **example transduction** empirical basis existing transductions conclusion

comparing theories



unified place theory

- consonants and vowels share representational primitives
 - e.g. LABIAL C-place, LABIAL V-place
- Sagey (1986), Clements & Hume (1995), a.o.

vowel feature theory

- vowel place is largely defined by primitives not used to describe consonant place
 - e.g. [+back], [-round]
- Odden (1991), Ni Chiosain & Padgett (1993), Halle et al. (2000), a.o.



full structure

segments



full structure

segments

the transduction: unified \rightarrow v-feature







[u]

the transduction: v-feature \rightarrow unified









unified and v-features are QF-bi-interpretable

and are therefore notational variants?

"A QF transduction is extremely restricted in the degree to which the output can differ from the input because QF is a weak logical language limited to local operations. QF-bi-interpretability can therefore be considered an indication of notational equivalence."

(Strother-Garcia 2019: 39)

enumerating natural class extensions



general procedure:

- given the set of possible segments S, find all subsets of S
- for each subset N, find the shared structure G
 - for each segment not in N, check if it also contains G
 - if true, N is not a natural class extension
 - if false, continue
 - N is a natural class extension

enumerating natural class extensions



	subset	{p, u, p ^w , t}	{p, u, p ^w }	$\{p, p^w, t\}$	{ p , p ^w }	{p, t}	{ u , p ^w }
hared structure	unified	•	• LABIAL	• C-Place	• C-Place LABIAL	Ø	• V-Place LABIAL
largest unique s	v-feature	•	Ø	• Place	• Place LABIAL	• Place -round	• +round

comparing natural class extensions





comparing natural class extensions

 ${p, p^w, t}$

{p, u}

 $\{u, t\}$

 ${p^w, t}$

 ${p, u, t}$



 $NCE_{unified} \neq NCE_{v-feature}$ ${}^{\{p, u, p^w, t\}}_{\{p, p^w, t\}}$ ${}^{\{p, u, p^w\}} transdu{}^{\{p, v, t\}}_{\{u, p^w\}} n between them$ is not natural class preserving

NCE _{unifie}
CATTITIC .



hypothetical assimilation

• $t \rightarrow p / _ u$

Naturalness of Assimilation 🕉

the targets and triggers of an assimilation process should constitute a natural class extension



natural class extension: {p, u, p^w} shared structure: LABIAL

- predicts [p^w] also a trigger
- seems reasonable



- seems wrong
- not even an assimilation rule



comparing theories



(1) Conditions for notational equivalence

- a. Two models do not differ in their empirical predictions.
- b. Two models represent the same set of abstract properties, differing only superficially.

(from Oakden 2021, summarizing Fromkin 2010)

• if we take seriously assumptions like Naturalness of Assimilation 🕉 (in whatever formulation), then a QF-bi-interpretable contrast-preserving transduction is not enough to satisfy (1a) above

introduction definitions example transduction empirical basis existing transductions conclusion



decided to sleep omitted for time, but see the following:

• Danis, Nick. 2020. Phonology needs geometry: Implicit axioms in segmental representation. 2020 Annual Meeting of the LSA. Poster.

2.0.)

• Danis, Nick. Cross-category agreement as reference to general dominance. AMP 2018, UC San Diego. Oct 5-7, 2018. Poster.

Phonology needs geometry: Implicit axioms in segmental representation Nick Danis nadanii@wustl.edu

Main Points -

- Phonological features are organized into "motivated subsets".
 Can a specific feature be in multiple subsets (or classes), depend-
- ing on the segment, or is all membership unique and absolute? • The question here is not of implementation (e.g. sets vs. trees), but rather on the implicit axioms governing the definitions of the sets: is class membership globally assigned or locally (per segment)?
- (One aspect of) of Feature Geometry is the idea that segments have nontrivial structure.
- Evidence from cross-cateogry place interactions supports a segment-specific (geometric) model of segmental representation.

Output of assimilation includes two segments having the same feature (value):



Geometry There exists organizational information about features that must be specified on a per-segment basis

Global Class Assignment (GCA)

 $(\forall f,g) [\texttt{label}(f) = \texttt{label}(g) \rightarrow (\neg \exists C)[C(f) \land \neg C(g)]]$ "If two features f and g are the same (share a label), their class memberships are always identical."

Unpacking the GCA -

- Feature organization is hierarchical (Clements 1985, Sagey 1986, a.o.)
- Classes refers to defined subsets of features, agnostic of dominating nodes vs. sets Place Place = {lab, cor, dors, ...}



- The GCA is an axiom (potentially) governing how the classes are defined, not how they are implemented structurally
- Given an indivual feature, is all class membership determined irrespective of any individual segment?
- Feature theories can be grouped into those that obey the GCA and those that do not

Case Study: [labial] ______. • To what extent are these groups of segments related phonologi-

cally? Rounded vocalics Plain labials / w u k[∞] / p kp /

Feature Class Theory: Obeys GCA

• "Disembodied" feature organization (Padgett 1995, 2002)

- Rounded vocalics = [+round]
 Plain labials = [labial]
- Elsewhere in theory:
 [+round] ∈ V-Place
 [labial] ∈ (C-)Place

Structure can be removed from individual segments as long as class definitions obey GCA

Not all theories of FG can be translated into FCT (contra Cahill and Parkinson 1997)
Rounded vocalics and plain labials not a natural class

 rounded vocaties and plant labrais not a natural class
 Other GCA-obeying theories (non-exhaustive): Chomsky and Halle (1968) (trivially), Ni Chiosáin and Padgett (1993), Halle et al. (2000)

Unified Feature Theory: Incompatible with GCA

Unified Feature Theory: Rounded vocalics and plain labials form a natural class (Clements and Hume 1995)

Rounded vocalics	Plain labials
C-Place	C-Place
\frown	\sim
V-Place _C	[labial] _g V-Place _C
[labial]	

The class membership of [labial] can vary segment to segment!
Unified Feature Theory are incompatible with the GCA (and therefore with Feature Class Theory)

 Other GCA-breaking theories (non-exhaustive): Mester 1986, Padgett 1994, Dependency Phonology, Governmet Phonology

Summary

In order for rounded vocalics and plain labials to be a natural class, we must assume Unified Feature Theory
Unified Feature Theory is incompatible with the GCA

 Is there phonological evidence for a natural class of plain labials and rounded vocalics?

Washington University in St.D	DUIS
-------------------------------	------

Natural	latural classhood of labials				
 Vietnar 	nese: $k \rightarrow$		(Kirby 20		
	$V\downarrow C\rightarrow$	Palatal	Velar	Labial-Velar	
	Front	[sec] 'slanting'	*[ek]	*[ekp]	
	Central	*[ac]	[sak] 'corpse'	*[akp]	
	Back	*[oc]	*[ok]	[sokp] 'shock'	

• UFT: Trigger and target of assimilation are both [labial]

[labial] V-place triggers [labial] C-place Assimilation is natural

- FCT: Trigger is [+round], target is [labial]
- [+round] triggers [labial]
 Assimilation **not** natural
- Related processes:
- $\begin{array}{ll} & \mbox{ Mumuye: } [kp] \sim [k^{w}] & (Shimizu 1983) \\ & \mbox{ Aghem: } b \rightarrow \mbox{ gb } / \mbox{ o } _ & (Hyman 1979) \end{array}$
- In order to preserve Naturalness of Assimilation, rounded vocalics and plain labials must be a natural class.
- Natural classhood of labials is only possible assuming UFT.
- If we assume UFT, then the GCA cannot be maintained
- Thus, organizational structure of these place features must be specified on a segment-specific basis.

Thus, phonology needs geometry.

References and Acknowledgements Thanks to the Wah U Linguistic, Brown Bag andience, many others unsmaller her provide the second second second second second second second second the provide the second second second second second second second second the provide second second second second second second second second second the provide second the provide second seco

Cross-category agreement as reference to general dominance Nick Danis Princeton University

Main Points -----

- Building on the Clements & Hume 1995 model, vowels and consonants share place features in a motivated and formally coherent way.
- The (expanded) typology of cross-category agreement motivates the homogeneity of vowel and consonant place features and the use of a syntagmatic AGREE constraint (Bakovic 2000).
- General dominance of root nodes over individual place features is encoded in the model, and cross-category agreement (and faith-

Definitions -----

Naturalness of Assimilation (NoA) is the idea that the rule or constraint causing an asimilation process refers to a single feature: • $X \rightarrow [aF] / _ [aF] \qquad cf. X \rightarrow [aF] / _ [aF]$ $• "[aF] [_aF] \qquad cf. X \rightarrow [aF] / _ [aF]$ Gross-category agreement or assimilation is the interaction of con-

sonant place features with vowel place features.

fulness) constraints crucially make reference to this relation.

......

Background and Data -

Padgett 2011 a.o.: the extent to which vowels can cause change in consonantal place of articulation is limited to palatalization.
 Table 1: Cross-category typology of Padgett 2011
 Within-category Cross-category
 Within-attegory Cross-category

V-to-C	$\sqrt[4]{e} \rightarrow [u] / w$ (Kabardian)	✓/i/ → [u] / p, m_ (Mapila Malayalam)
C-to-V	$\sqrt[4]{T/} \rightarrow [T^j] / _i, e$ (Russian)	$\checkmark/k/ \rightarrow [tj] / _i (Slovak)$ $\divideontimes/k/ \rightarrow [p] / _u (Unattested?)$

 In Vietnamese, back, round vowels cause velar coda consonants (stops and nasals) to become labial-velar double articulations, in addition to front vowels causing palatalization. (Kirby 2011, Pham 2006, Thompson 1965)

Table 2: Rhyme restrictions in Vietnamese (summarized from Kang et al. 2016)

$V\downarrow C\rightarrow$	Palatal	Velar	Labial-Vel:
Front	[sec] 'slanting'	*[ck]	*[ekp]
Central	*[ac]	[sak] 'corpse'	*[akp]
Back	*[oc]	*[ok]	[sokp] 'shock'

 Vietnamese fills the empirical gap of a non-palatalization consonant-to-vowel cross-category assimilation process.

Representation ------

 Geometry for representations explicitly includes transitive association relation A_T (assumed to be dominance).



Figure 1: Model of representation for [0kp]

 Consonants and vowels share the core set of place features [labia], [coronal], and [dorsal] (following Clements & Hume 1995).
 Grammar here does not make crucial reference to tier structure.

so it is not encoded.

Grammar — 1. Cross-category AGREE_X [LAB]:

(a) $(\forall x, y) [\delta(x, y) \land isLab(x) \leftrightarrow isLab(y)]$ (b) $isLab(x) \equiv (\exists y) [Root(x) \land lab(y) \land A_T(x, y)]$

"For all root nodes x, y in a some domain δ, x generally dominates [labial] iff y generally dominates [labial]."
2. No Rounding ("C") :

 $\begin{array}{l} (\neg \exists w, x, y, z) [\texttt{Root}(w) \land \texttt{C-pl}(x) \land \texttt{V-pl}(y) \land \texttt{lab}(z) \land \\ A_T(w, x) \land A_T(w, y) \land A_T(y, z) \end{array}$

Table	3: Cross-catego	ry agre	ement in Vietn	amese
/ok/	AGREE, LAB	*C*	IDENT [DORS]	*COMPLEX

a.	okp				•
b.	ok	* W			L
с.	ok*		* W		L
d.	op			* W	L

- The back, round vowel contains labial and dorsal V-place features, which totally agree with the C-place features of the double-articulation $\overline{kp}.$

 Palatalization works in a parallel fashion: an input such as /ck/ is realized as [ec], where [c] is phonologically dorsal and coronal C-place: there is cross-category coronal assimilation, while the dorsal C-place of the input is preserved (see Danis 2017 for a full analysis/discussion).

 Class behavior of place: Feature Class Theory (Padgett 2002) or constraint summation (Danis 2017)

PRINCETON UNIVERSITY

 Models where phonetic rounding is expressed via [+round] (e.g. Halle et al. 2000, Ni Chiosáin & Padgett 1993, a.o.) cannor straightforwardly capture interaction with non-round labials while obeying NoA, as in Vietnamese.

 Halle et al. 2000's critique of C&cH: actually modeling crosscategory interactions is cumbersome (though see e.g. Selkirk 1988 and Urek 2016).

а.	$rt_1 \longrightarrow rt_2$	b. $rt_1 \longrightarrow rt_2$	$rt_1 \longrightarrow rt_2$
	1 1	1 1	1
	Cpl ₃	Cpl ₃ Cpl ₇	$Cpl_3 \longrightarrow Cpl_7$
	Vpla		Vpla /
	tor dor	dor da	a day a day
	John John	lab.	lab lab

Figure 2: Vietnamese under the C&cH model requires either transplanar spreading (a.) or a derivational process of spreading (bi.) and promotion (bii.) for cross-category assimilation

Cross-category faithfulness interactions also occur, and can be modeled with a parallel $IDENT_X$ constraint:

Mumuye [mzm]: kp~k^w alternation (Shimizu 1983)

	/kp/	*COMPLEX	IDENT _X	*C ^x	IDENT _c	*C*
a.	k ^w				•	•
b.	kp	* W			L	L
с.	k		* W		•	L
d.	p		* W			L
c.	p ^v			* W	•	L

 The general domination model plus cross-category AGREE_x/ IDENT, straightforwardly captures both cross-category agreement and faithfulness processes while obeying Naturalness of Assimilation, supporting a model of unified place features.

References and Acknowledgements -

Thanks to Akin Akinlabi, Jaye Padgett, & Adam Jardine for comments on earlier versions; my dissertation committee for comments on the chapter from which this poster is expanded; and to Eric Bakwie for digging though his office for old manuscripts.

tuttor context: characteristic principal and the second second

d.	C-to-V /
must be	 In Vietnam (stops and addition to 2006, Thor

introduction definitions example transduction empirical basis existing transductions conclusion

Oakden (2021)









- Oakden (2021) provides a nonsize-preserving QF transduction (above) between two theories of tone sandhi (left), arguing for notational equivalence
- is this transduction also natural class preserving?



Oakden (2021)

• yes

• because Python

[8]: 1 yip = { 'L' : {('s','-u'), ('-u','l')}, 'H' : {('s', '+u'), ('+u', 'h')}, 3 'M1' : {('s','-u'), ('-u','h')}, 'M2' : {('s','+u'), ('+u','l')}, 5 6 'HM' : {('s', '+u'), ('+u', 'h'), ('+u', 'l')}, 7 'MH' : {('s', '+u'), ('+u', 'l'), ('+u', 'h')}, 'ML' : {('s','-u'), ('-u','l'), ('-u','h')}, 8 'LM' : {('s','-u'), ('-u','l'), ('-u','h')} 9 10 } 11 12 bao = { 13 'L' : {('s','T'), ('T','-u'), ('T','c'), ('c','l')}, 'H' : {('s','T'), ('T','+u'), ('T','c'), ('c','h')}, 14 15 'M1' : {('s','T'), ('T','-u'), ('T','c'), ('c','h')}, 'M2' : {('s','T'), ('T','+u'), ('T','c'), ('c','l')}, 16 'HM' : {('s','T'), ('T','+u'), ('T','c'), ('c','h'), ('c','l')}, 17 'MH' : {('s','T'), ('T','+u'), ('T','c'), ('c','h'), ('c','l')}, 18 'ML' : {('s','T'), ('T','-u'), ('T','c'), ('c','h'), ('c','l')}, 19 'LM' : {('s','T'), ('T','-u'), ('T','c'), ('c','h'), ('c','l')}, 20 21 } 22 23 compare theories(yip, bao)

```
Natural classes unique to theory 1:
Natural classes unique to theory 2:
Natural classes in common:
('LM', 'M1', 'ML')
('LM', 'ML')
('HM', 'M2', 'MH')
('H', 'HM', 'M2', 'MH')
('L', 'LM', 'M1', 'ML')
('H', 'HM', 'LM', 'M1', 'MH', 'ML')
('H', 'HM', 'L', 'LM', 'M1', 'M2', 'MH', 'ML')
('HM', 'L', 'LM', 'M2', 'MH', 'ML')
('H', 'HM', 'MH')
('L', 'LM', 'ML')
('HM', 'MH')
('HM', 'LM', 'MH', 'ML')
```

introduction definitions example transduction empirical basis existing transductions **conclusion**

summary



- the use of features/subsegmental structure sorts segments into natural classes based on shared structure
- while transductions can translate from model A to model B, the predicted natural classes present in the system as a whole may still differ in a linguistically significant way
- the definition of natural class preserving transductions is a start to identify comparisons of theories which still involve differing linguistic predictions

going forward



- the definition of *natural class preserving* is based on the representation themselves—can this property be identified by investigating the transduction rules alone?
- how else can transductions themselves be compared and evaluated from a linguistic standpoint?
- how strongly should our metatheoretical assumptions and expectations about linguistic processes be formalized?





& thank you Adam Jardine (for helping with the transductions like 3 years ago) and the audience at the Wash U Linguistics Brown Bag on Sept. 16