

# Learning Phonetically and Phonologically Natural Classes through Constraint Indexation

Aleksei Nazarov  
*Utrecht University*

## 1 Introduction

Phonological processes are typically defined over (natural) classes (Chomsky & Halle 1968). These classes are motivated by the desire to write compact grammars (Chomsky & Halle 1968) and generalization of processes to novel segments: for instance, generalization to non-native segments (e.g., Halle 1978); and generalization to segments not seen in the training phase in artificial language learning experiments (e.g., Cristia et al. 2013).

While many models of phonological representation work from the assumption that the set of (potential) natural classes is language-universal (e.g., Chomsky & Halle 1968), there is evidence for some arbitrary and language-specific aspects to class behavior (e.g., Mielke 2004). If segment classes that participate in phonological processes can be language-specific, then there must be a learning procedure by which language-learning infants induce these segment classes using ambient linguistic data.

One way to model this learning process is through contrast detection (Dresher 2014, Sandstedt 2018), whereby an infant detects contrasts in the lexicon, and uses these to find which classes of sounds are in contrast in the language. However, the existing approaches that use contrast (Dresher 2014, Sandstedt 2018, Mayer 2020) are outside of Optimality Theory (OT). In addition, they use at least some domain-specific methods (i.e., learning methods that only apply to linguistic data, which arguably should be specified inside of Universal Grammar (UG)) for detecting contrast rather than domain-general ones (i.e., learning methods that could potentially be applied to all kinds of data and need not be specified in UG): they use “micro-cues” to contrast (Sandstedt 2018) or phonological contrast detection mechanisms (Dresher 2014).

The question explored in this paper is whether segment class induction can be implemented in standard OT using domain-general methods. I will propose (natural) classes can be seen as a consequence of resolving inconsistency with locally indexed constraints (Temkin-Martínez 2010, Round 2017), and I will test this proposal on a set of natural classes included in transparent, opaque, and exceptional processes. This is done because natural classes do not always present themselves on the surface, but may be obscured by other, (morpho)phonological factors: interacting phonological processes and exceptions can make a process non surface-true (there are surface forms that should undergo the process but do not) or non surface-apparent (there are surface forms that undergo the process but should not; McCarthy 1999), meaning that the set of segments in the data that undergo the process not simply defined by phonetics but also by interactions with (morpho)phonology.

Section 2 will present the OT learning implementation (including (locally) indexed constraints) used to test whether this can be done, while section 3 will lay out the toy language case studies used here. Section 4 will show the results of applying the OT learner to the case studies, while section 5 will discuss the implications of these. Finally, section 6 will conclude.

---

\* Many thanks to Adam Albright, Ricardo Bermúdez-Otero, Stuart Davis, Elan Dresher, Gaja Jarosz, Peter Jurgec, Michael Kenstowicz, Athina Kikiopoulou, John McCarthy, Marc van Oostendorp, Joe Pater, Nathan Sanders, Mirella De Sisto, Brian Smith, Nina Topintzi, audiences at the University of Toronto, UMass Amherst, and MIT, and at the 27th Manchester Phonology Meeting. All errors are my own.

## 2 OT approach(es) to contrast

In OT, the concept of contrast has been implemented in various ways: as constraints that refer to contrast between specific morphemes (e.g., Lubowicz 2003), as anti-faithfulness (Alderete 2001), as phonotactic rankings in Stratal OT (Mackenzie 2016), or as mechanisms that divide words into groups that contrast in their phonological behavior: cophonologies (Inkelas & Zoll 2007) or indexed constraints (Kraska-Szlenk 1995, Pater 2000, 2010). Of these, indexed constraints have been the most explicitly linked to contrast induction. Becker (2009), Pater (2010) propose that indexed constraints are induced by the language learner whenever there is an inconsistency in ranking requirements between various inputs (see Tesar 1995 for inconsistency). This idea of two morphemes’ having inconsistent ranking requirements essentially means that there is contrast between these two morphemes: there is no phonological factor that can distinguish between the behaviors of these two morphemes, so therefore these morphemes contrast with each other. In this sense, inconsistency detection can be seen as a cue to contrast: whenever we see inconsistency in Tesar’s (1995) sense (i.e., two morphemes with mutually incompatible ranking requirements), we know there must be some contrast between these morphemes that we have not yet acknowledged. Crucially, this cue is domain-general, since it is computed based on constraint violations only. While OT-constraints are normally used for linguistic applications, it is possible to apply them to non-linguistic data, as well (e.g., Parker & Parker 2004), which means that inconsistency detection could also be used to find contrast in non-linguistic data.

In section 2.1, I will explain how various types of contrast can be expressed in indexed constraint grammars, and I will introduce segmentally local indexation. Then, in section 2.2, I will explain the learning mechanisms that allow to induce segmentally local indexation to encode segment-level contrast. Finally, section 2.3 will show the general setup of the learner.

**2.1 Global vs. local indexation** Indexed constraints as defined by Pater (2000) – i.e., constraints that are violated only for certain lexical items or morphemes – encode contrast at the morpheme level: one group of morphemes, indexed  $i$ , undergoes a certain process ( $A_i \gg B$ ), while all other morphemes, not indexed  $i$ , do not undergo it ( $B \gg A$ ), or *vice versa* if  $A$  is a faithfulness constraint. While contrast at the morpheme level is a useful and economical concept, encoding classes of segments requires contrast between specific segments within morphemes, which standard indexed constraints cannot encode.

However, the concept of locally indexed constraints (Temkin-Martínez 2010, Round 2017) provides a solution. Locally indexed constraints are constraints that are violated only for particular segments in particular morphemes: for instance, in the hypothetical morpheme /ma<sub>k</sub>nan/, only the first vowel /a<sub>k</sub>/ might be indexed to the markedness constraint \*V[-nas]<sub>k</sub>N, which prohibits oral vowels indexed  $k$  before nasal consonants given the ranking \*V[-nas]<sub>k</sub>N  $\gg$  IDENT(nas)  $\gg$  \*V[-nas]N, meaning that only the first vowel gets nasalized on the surface: [mãnan]. This constitutes a contrast between two types of vowels: potentially nasalizing /a<sub>k</sub>/ and never nasalizing /a/. In standard (“globally”) indexed constraints, this contrast between the first and second vowel cannot be made: if we have the morpheme /manan<sub>k</sub>/, all of its segments are equally subject to \*V[-nas]N<sub>k</sub>, and the ranking \*V[-nas]N<sub>k</sub>  $\gg$  IDENT(nas)  $\gg$  \*V[-nas]N leads to nasalization of both vowels: [mãñãñ].

In addition to local indexation, I will also treat indexation as being binary (which is inspired by Becker’s 2009 constraint cloning approach): all segments that are not  $k = [+k]$  are explicitly marked as  $[-k]$ . This means that the example from the preceding paragraph will be represented /m<sub>[-k]</sub>a<sub>[+k]</sub>n<sub>[-k]</sub>a<sub>[-k]</sub>n<sub>[-k]</sub>/ – and this makes indexation equivalent to so-called alphabet features (non-phonetic binary features a value for which is assigned to every individual segment) from SPE (Chomsky & Halle 1968). Binary local indexation thus forms “proto-features” of sorts: indices may be linked to a high-ranked markedness constraint, which assigns all segments with that index some phonetic property, which means that this index defines a natural class. For instance, if the constraint \*V<sub>[+m]</sub> is undominated, all underlying segments with the index [+m] must surface as consonants or be deleted, which means that, on the surface, [+m] stands for the class of consonants and is a “proto-feature” that can form the basis of a [syllabic] distinction. Once such “proto-features” that define natural classes are learned, these may be the basis for constructing a language-specific phonological feature system (see Mayer & Daland 2020 for an algorithm to construct a feature system from natural classes).

At the same time, the set of segments that carries a particular index may also be indexed to a lower-

ranked markedness constraint or to a faithfulness constraint, in which case a different type of segment class may be found: a class of random segments, or a class that is only “phonologically natural” (see section 4), or perhaps even a class of unnatural segments. This means that discussion of learning is important: how plausible is it that learners will learn representations of natural classes given data that support these classes?

**2.2 Learning mechanisms** Binary local indices, as described above, lead to a significant increase in the number of grammar hypotheses compatible with any given dataset: instead of just multiple rankings of a fixed set of constraints’ being compatible with the same data, we now have multiple rankings of multiple constraint sets (various combinations of indexed constraints + a fixed set of universal constraints) along with multiple assignments of indices to specific morphemes. Given this explosion of the learning space, can an appropriate grammar still reasonably be found given a plausible dataset? Can a grammar that encodes appropriate “proto-features” be found given data that contain various types of segment classes?

To explore these questions, we need an explicit learner with plausible assumptions that would induce indexed constraints and sets of segments associated with them as well as give these constraints an appropriate ranking. Such learners exist in the form of existing indexed constraint learners (Becker 2009, Pater 2010, Round 2017), in particular, Round’s (2017) segmentally local indexed constraint learner.

This latter learner is an extension of a pre-existing categorical (=non-probabilistic) OT learner: Biased Constraint Demotion (BCD; Prince & Tesar 2004). This learner takes pre-defined OT tableaux with underlying representations (URs), winning (observed) and losing (non-observed) candidates, and constraint violations as input, and it gradually builds a ranking of the pre-defined constraints from high to low based on ranking requirements derived from the constraint violations of winning and losing candidates. The procedure ends once the ranking correctly eliminates all losing candidates and all constraints are incorporated into the ranking.

However, there is a possibility that the pre-defined tableaux will contain contradictory ranking requirements. For instance, if there is a word where a prenasal vowel undergoes nasalization (/pan/ → [pã̃]) and another word where the same prenasal vowel remains oral in the same phonological context (/apan/ → [apan]), this leads to contradictory ranking requirements  $*V[-nas]N \gg IDENT(nas)$  and  $IDENT(nas) \gg *V[-nas]N$ , respectively. Whenever the learner encounters this situation, this is called inconsistency (Tesar 1995). Becker (2009), Pater (2010) propose to use inconsistency in constraint demotion as a diagnostic that an indexed constraint is necessary. An indexed constraint that applies only to inputs that require one of the rankings is inserted, which resolves the inconsistency. For instance, if all words that require  $*V[-nas]N \gg IDENT(nas)$  are indexed to the constraint  $*V[-nas]N_k$ , the ranking  $*V[-nas]N_k \gg IDENT(nas) \gg *V[-nas]N$  leads to nasalization in nasalizing words and non-nasalization in other words.

Round (2017) extends their proposals to apply to specific loci (specific underlying segments in the lexicon). Whenever inconsistency is detected, all underlying segments that require one of the contradictory rankings involved are indexed to a relevant constraint (see Round 2017 for the selection mechanism), which leads to constraints that are violated only for specific segments in the lexicon. For instance, in the word /manan/ → [mã̃nan], the first vowel requires the ranking  $*V[-nas]N \gg IDENT(nas)$  while the second vowel requires  $IDENT(nas) \gg *V[-nas]N$ . This is solved by giving the first vowel the index [+k]: /ma<sub>[+k]</sub>na<sub>[-k]</sub>n/, adding the locally indexed constraint  $*V[-nas]_{[+k]}N$ , and building the ranking  $*V[-nas]_{[+k]}N \gg IDENT(nas) \gg *V[-nas]_{[-k]}N$ . Thus, Round’s learner, an implementation of which will be used in my learning setup (see section 2.3), starts out with pre-defined universal constraints and a pre-defined set of URs and winning and losing candidates, and ends up with a ranking of universal constraints and locally indexed versions of these universal constraints (i.e., version of these constraints that are violated only by certain underlying segments in the lexicon).

**2.3 Learning setup** The goal of the current learning setup are to account for the (transparent, opaque, exceptional) patterns in the data and to simulate the induction of language-specific (natural) classes in these data using constraint indexation. To allow natural classes to be induced, this learning setup assumes that no phonological features as such as available: these features, as suggested in section 2.1, are taken to be induced at the next step.

Since no features are assumed during the simulations, I assume that there are no faithfulness constraints, since these, for segmental phenomena, refer to phonological features. Rather, the tableaux

given to the learner are purely phonotactic, i.e., feature only markedness constraints; URs only consist of unspecified segment slots that have identifiers on them that are unique to that morpheme (similar to the idea of “colors” in Colored Containment, Van Oostendorp 2008):  $/X_1X_2X_3X_4X_5/ \rightarrow [\text{mānan}]$ . Since, again, no features are assumed, the markedness constraints in the simulations (defined in section 3.2) are based on articulatory and perceptual factors, conceptually similar to Boersma’s (2007) cue, sensorimotor and articulatory constraints.

The goal of this stage is to rank (indexed versions of) these articulatory and perceptual constraints such that this will account for the pattern in the data. The formation of features, plausible URs, and faithfulness is seen as a next step that would build on the output of the step modeled here. Thus, only “proto-features” are induced in the current simulations rather than true phonological features, and the reason why the simulations do not start with universal features in the first place is to make room for language-specific (natural) classes. The schema in (1) summarizes the learning setup used in this paper:

(1) Learning setup summary

**START**

Data:  $/X_1X_2-X_3/ \rightarrow [\text{efe}]$   
 $/X_4X_5-X_6/ \rightarrow [\text{ifi}]$  etc.  
 Natural classes: none  
 Universal constraints:  $*f, *{sf}, *{iu} \dots \{eo\}$  etc. (see section 3.2)  
 Indexed constraints: none  
 Ranking: none

**PROCEDURE**

Rank constraints high to low (BCD, Prince & Tesar 2004);  
 If **inconsistency** detected: add locally indexed constraint (Round 2017)

**RESULT**

Data:  $/X^1_{[-i,-j]}X^2_{[-i,+j]}-X^3_{[-i,-j]}/ \rightarrow [\text{efe}]^1$   
 $/X^4_{[-i,-j]}X^5_{[-i,+j]}-X^6_{[-i,-j]}/ \rightarrow [\text{ifi}]$  etc.  
 Natural classes:  $\{i\}=[+i]$  vs.  $\{u\}=[-i]$   
 $\{f\}=[+j]$  vs.  $\{sf\}=[-j]$  etc.  
 Universal constraints:  $*f, *{sf}, *{iu} \dots \{eo\}$  etc. (as above)  
 Indexed constraints:  $*f_{[-j]}, *{sf}_{[+j]}, *{iu}_{[+i]} \dots \{eo\}_{[-i]}$ ,  
 Ranking:  $*{iu}_{[+i]} \dots \{eo\}_{[-i]}, \dots \gg *{sf}_{[+j]} \gg \dots$   
 $\gg *{iu} \dots \{eo\}, *{sf}$

### 3 Data used for simulations

**3.1 Toy languages** The learner described in sections 2.2-3 is applied to three different toy data case studies based on similar toy data case studies from Prickett & Jarosz (2021). These case studies illustrate transparent, opaque, and exceptional application of a process, which illustrate the potential of the current natural class induction procedure to analyze cases where natural classes are obscured by other (morpho)phonological processes.

Each case study is based on two potentially interacting processes:

<sup>1</sup> For the sake of legibility, segment identifiers are shown as superscripts rather than subscripts here.

## (2) Processes involved in case studies

## a. Vowel Harmony (VH; dominant/recessive):

*raise mid vowels (e) to high vowels (i) in the presence of a high vowel*

Rule

$e \rightarrow i / \{ \_ C_0 i/u, i/u C_0 \_ \}$

Examples

$ef- + -i \rightarrow [ifi]$        $ef- + -e \rightarrow [efe]$

$uf- + -i \rightarrow [ufi]$        $uf- + -e \rightarrow [ufi]$

## b. Palatalization (Pal)

*palatalize s to f before the high vowel i*

Rule

$s \rightarrow f / \_ i$

Examples

$is- + -i \rightarrow [ifi]$

## c. Interaction

VH potentially feeds Pal

$is- + -e \rightarrow isi \rightarrow [ifi]$  (feeding) OR

$is- + -e \rightarrow [isi]$  (counterfeeding)

Three toy languages are constructed based on different interactions between these two processes: transparent palatalization (feeding interaction, 2c), opaque palatalization (counterfeeding interaction, 2c), and exceptional palatalization (transparent palatalization in some stems, no palatalization in others). These interactions are shown over a small set of stems and suffixes listed in (3a). The interactions are illustrated in (3b) (stems ending in /f/ are not shown, since these have no palatalization). In all three languages, the data points /es-e/ (does not trigger either process), /es-i/ (harmony and palatalization target different vowels), and /us-i/ (only triggers palatalization) show up the same. However, the outcomes for /is-e/, /is-i/, and /us-e/ differ. In the transparent language, /is-e/ and /us-e/ undergo palatalization, because the raising of /-e/ feeds palatalization, and /is-i/, like all /si/ sequences, also undergoes palatalization. In the opaque language, these forms do not undergo palatalization, because raised /-e/ does not feed palatalization, while /is-i/, like all /si/ sequences, does undergo palatalization. Finally, in the lexically specific language, the /is-/ forms do not undergo palatalization because /is-/ is marked as not undergoing this process, while /us-e/ does undergo palatalization through feeding because /us-/ is marked as undergoing palatalization.

## (3) Processes involved in case studies

a. Stems: *es-*, *ef-*, *is-*, *if-*, *us-*, *uf-*

Suffixes: *-e*, *-i*, *-u*

## b. Transparent palatalization

$es- + -e \rightarrow [ese]$

$es- + -i \rightarrow [ifi]$

$is- + -e \rightarrow [ifi]$  (feeding)

$is- + -i \rightarrow [ifi]$

$us- + -e \rightarrow [ufi]$  (feeding)

$us- + -i \rightarrow [ufi]$

## Opaque palatalization

$es- + -e \rightarrow [ese]$

$es- + -i \rightarrow [ifi]$

$is- + -e \rightarrow [isi]$  (counterf.)

$is- + -i \rightarrow [ifi]$

$us- + -e \rightarrow [usi]$  (counterf.)

$us- + -i \rightarrow [ufi]$

## Lexically specific palatalization

$es- + -e \rightarrow [ese]$

$es- + -i \rightarrow [ifi]$

$is- + -e \rightarrow [isi]$  (/is-/: no Pal)

$is- + -i \rightarrow [isi]$  (/is-/: no Pal)

$us- + -e \rightarrow [ufi]$  (feeding)

$us- + -i \rightarrow [ufi]$

These data contain various natural classes that a class/index induction algorithm would be expected to internalize. The most obvious classes, shared by all three datasets, are a class of high vowels {i,u} as opposed to the mid vowel {e}, a class of front vowels {e,i} as opposed to the back vowel {u}, and a class of coronal consonants {s,f} as opposed to the labial consonant {f}. For the opaque dataset, there is an additional class of palatalizing vowels ([i] deriving from *-i*: *is-i* [ifi]) as opposed to non-palatalizing vowels ([i] deriving from *-e*: *is-e* [isi]); see also Drescher (2009) for another case of a contrast between palatalizing versus non-palatalizing [i]. For the lexically specific dataset, there is an additional class of palatalizing [s] (the *s* in *es-* and *us-*: *es-i* [ifi], *us-i* [ufi]) as opposed to non-palatalizing [s] (the *s* in *is-*: *is-i* [isi]). These additional classes of segments are defined by a combination of a phonetic description and participation in a process; in this paper, I call these “phonologically natural classes” (see section 4.2 for more detail).

**3.2 URs, candidates, and constraints**

For each of the languages described in section 3.1, a training dataset is made by setting up URs, winning and losing candidates, and constraint violations, as required by

the learner (section 2.2). As mentioned in section 2.3, the URs for this approach do not have any feature content, but do contain segment locus identifiers, so that the learner can distinguish between segments within a single morpheme (e.g. *e* vs. *s* in *es-*) and between the segments in different morphemes (e.g. *s* in *es-* vs. *s* in *is-*).

Each input is mapped to the same set of 48 different candidates, which consist of all possible combinations of {*e*, *i*, *o*, *u*} × {*f*, *s*, *ʃ*} × {*e*, *i*, *o*, *u*}. Most of these candidates do not occur as surface forms in the three languages, which means that these can be excluded by universal constraints. However, whenever there are differences between morphemes (e.g., *us-* is always realized with [u] while *is-* is always realized with [i]), these differences must be accounted for by indexed constraints.

At the outset, each of the 48 candidates receives violations of the 11 universal constraints shown in (4). During learning, indexed versions of these constraints are made according to the procedure sketched in 2.2. As required by this procedure, the violations defined for the 48 candidates and 11 constraints are also linked to specific segment locus identifiers. For example, candidate /X<sub>1</sub>X<sub>2</sub>-X<sub>3</sub>/ [isi] violates the constraint \**si*, which is notated as “X<sub>2</sub>” in the violation column (for constraints that cover multiple segments, a violation is assessed for the leftmost segment); it also violates the constraint \*{iu} twice, which is notated as “X<sub>1</sub>,X<sub>3</sub>”. The full list of universal constraints used in the simulation is as follows:

- (4) Universal constraints used in the simulations
- a. Context-free: \*{iu} (no low F1) \*{eo} (no mid-range F1)  
\**f* \**s* \**ʃ* \*{sʃ} (no coronal fricatives)
- b. Harmony-triggering: \*{iu}...{eo} \*{eo}...{iu}  
\*{ie}...{uo} \*{uo}...{ie} (no F1 differences if low F1 is present)
- c. Palatalization-triggering: \**si* (no narrow constriction at alveolar ridge followed by palatal vocalic gesture)

Though these constraints are notated as applying to segments and classes of segments, their definition is actually given in terms of acoustics or articulation. The context-free constraints on vowels in (4a) may be defined in terms of acoustics (ranges of F1, as indicated in (4a)); the context-free constraints on consonants in (4a) may also receive an acoustic definition (ranges of Center of Gravity), or, alternatively, an articulatory definition (no labiodental fricatives, no alveolar fricatives, etc.). The harmony-triggering constraints in (4b) may be defined acoustically (in terms of F1 differences, as indicated in (4b)). Finally, the palatalization constraint in (4c) may be defined in terms of articulation, as indicated in (4c).

## 4 Simulations and results

**4.1 Simulation setup** The three languages described in section 3.1 are learned using Nazarov’s (2021) implementation of Round’s (2017) local indexation induction procedure (section 2.2), written in R (R core team 2021). 10 runs were performed for each language. This is because Nazarov’s (2021) implementation contains some elements of random choice, so that it is important that the learner explore all options.

For this paper, the interest is whether and how the learner picks up the classes described in section 3.1. Therefore, the evaluation of the simulation focuses on which segments in the lexicon are connected to which indices. The algorithm described in Nazarov (2021) induces a separate index for every new indexed constraint. To make the resulting data better interpretable, indices are consolidated as follows.

For each induced indexed constraint, the set of loci in the dataset may be divided into three categories: the crucially [+*i*] segments, which must be indexed to the current constraint for the training data to be learned correctly, the crucially [−*i*] segments, which may not be indexed to the current constraint, and the neutral segments, which might be indexed either [+*i*] or [−*i*] based on the constraint definition. Two constraints are seen as referring to the same index if they share the same + or − value for at least one locus, and there are no mismatches between their + and − values: all the loci for the + value of constraint A have a + or neutral value for constraint B and *vice versa*, and all the loci for the − value of constraint A have a − or neutral value for constraint B and *vice versa*.<sup>2</sup>

<sup>2</sup> If the + value of constraint A consistently corresponds to the − or “unknown” value of constraint B and *vice versa*, this is also accepted as evidence that A and B refer to the same index, since these +/- values have no intrinsic meaning.

**4.2 Phonetic vs. phonological natural classes** The index consolidation procedure described above yields a number of segment classes for each run of the algorithm. These classes will be classified as phonetically natural or phonologically natural, terms which I use as follows. A phonetically natural class is a natural class according to the classic definition (e.g., Chomsky & Halle 1968): a class of segments that is at the intersection of a set of phonetic properties (reified as classificatory phonological features, Chomsky & Halle 1968). Examples of relevant phonetically natural classes in the current datasets include: all high vowels {i,u}, all high back vowels {u}, and all coronal consonants {s,ʃ}. What is meant by a “phonologically natural” class is a class defined by a mix of phonetic properties (as in phonetically natural classes) and phonological properties: undergoing particular phonological processes. Examples of such classes in the current data are palatalizing [i] (all instances of [i] that trigger palatalization of a preceding /s/), palatalizing [s] (all instances of /s/ that undergo palatalization).

Traditionally, such “phonologically natural” classes are not seen as natural classes at all, but rather as an epiphenomenon of natural classes and opaquely or exceptionally applying rules: if [i] fails to palatalize a preceding /s/ in the current opaque dataset (second column in (3b)), this is because palatalization applied before the rule that created the [i]; if [s] fails to palatalize before [i] in the current lexically specific dataset (third column in (3b)), this is because the However, works such as Mielke (2004) and Dresher (2009) do conceptualize sets of segments that are “active” in phonological processes as being relevant to the grammar. Furthermore, representational approaches to opacity like Van Oostendorp (2008) and Nazarov (2021), do conceptualize the set of segments that does or does not undergo a particular segment as a natural class: in Van Oostendorp (2004), undergoers of an opaque or exceptionally applying process can be defined as a combination of a set of features and a set of morphological colors; in Nazarov (2021), indexed constraints are used, where indices correspond to the set of undergoers of an opaque or exceptionally applying process.

**4.3 Summary of results** Table (5) shows the results of the procedure in section 3, implemented as in section 4.1, applied to the data in section 3.1, and interpreted as in section 4.2. The columns of the table shows the expected natural class oppositions that might be found from the data in section 3.1. In terms of phonetic natural classes, we may expect a mid-versus-high opposition, a back-versus-low opposition among the high vowels, and a labial-versus-coronal opposition in the consonants. In terms of phonological natural classes, we may expect a palatalizing versus non-palatalizing [i] opposition for the opaque case, and a palatalizing versus non-palatalizing [s] opposition for the lexically specific case. The rows of the table stand for each of the three toy languages, and the cells of the table indicate how many of the 10 runs for a particular language yield a particular (type of) class.

(5) Table of results: types of natural classes found for each language. Cells indicate numbers of runs.

Language	i,u vs. e	u vs. e,i	s/ʃ vs. f	-V <sub>pal</sub> vs. -V <sub>nonpal</sub>	/es-,us-/ vs. /is-/	Other phonological
Transparent	10	10	10	0	0	0
Opaque	10	10	10	10	0	10
Lexical	10	10	10	0	10	10

As can be seen in (5), all 10 runs for each language behave entirely consistently. There is some variation between the actual grammars learned for each language, but these grammars yield largely equivalent sets of classes. As is guaranteed by Constraint Demotion-based approaches (Tesar 1995), all grammars are 100% successful in accounting for the training data.

As may be expected (see section 4), the transparent language only yields phonetic natural classes. The algorithm finds exactly those classes that are expected: high versus mid vowels: {i,u} vs. {e}, back versus front vowels: {u} vs. {e,i}, and coronal versus labial consonants: {s,ʃ} vs. {f}.

For the opaque and lexically specific languages, the same phonetic natural classes are found, but in addition, appropriate phonological natural classes are also found. For the opaque language, as predicted, palatalizing and non-palatalizing (suffix) vowels are distinguished (-V<sub>pal</sub> vs. -V<sub>nonpal</sub>), where -i (and sometimes -u) is/are contrasted to -e. Strictly speaking, this can be unified with the high vs. mid distinction, a phonetically natural class, since the class of palatalizing suffixes does not conflict with the class of high

vowels (there are no vowels among the high vowels in the data that must be excluded from the class of palatalizing vowels). For the lexically specific case, as predicted, palatalizing and non-palatalizing consonants are distinguished (palatalizing: *es-*, *us-*; non-palatalizing: *is-*; the stems *ef-*, *if-*, *uf-* are classed with the palatalizing consonants since they never surface with [s] before [i]). The exact same phonetic natural classes are found as for the transparent language: high versus mid vowels, back versus front vowels, and coronal versus labial consonants.

As indicated in the last column of the table, additional phonological natural classes are found. For the opaque language, the other phonological natural class is that of the consonants in the stems *ef-*, *if-*, *uf-*, *es-* to the exclusion of *is-*, *us-*, indexed to the first position in the constraint \*si, to whose second the suffixes *-i* (and for some runs, *-u*) is/are indexed (which form the set of high/potentially palatalizing suffix vowels, see above). This is because the algorithm always finds two indices for constraints that encompass two positions, like \*si: one set of for the first position, and another set for the second position (see also Nazarov 2021). The set of stems *ef-*, *if-*, *uf-*, *es-* as followed by *-i* is chosen for a highly ranked indexed \* $s_{[+i]i_{[+j]}}$  because these stems never surface with the configuration [si] (cf. *is-e* [isi], *us-e* [usi]).

For the lexically specific language, the other phonological natural class is the class of consonants that never surface as [j] – thus, those consonants in which we do not see the effects of palatalization. This encompasses the consonants in *ef-*, *if-*, *uf-*, *is-* to the exclusion of those in *es-*, *us-*; these former consonants are indexed to the constraint \*ʃ. This phonologically natural class is a pseudo-complement of the natural class of palatalizing consonants; only *es-*, *us-* undergo palatalization, so *ef-*, *if-*, *uf-*, *is-* is the complement of this, but, as indicated above, the algorithm actually classes all instances of *f* with the palatalizing consonants, since these have in common with *es-*, *us-* that they conform to the constraint \*si (i.e., never show up as [s] before [i]).

## 5 Discussion/conclusion

The question asked in the introduction is whether natural class induction through contrast detection is possible in standard OT. As can be gleaned from the results in section 4.3, natural classes can indeed be accurately diagnosed using inconsistency resolution (see section 2.2), which is a domain-general technique used in the learning of standard OT grammars. These natural classes can be found even when the contrasts are tied up with processes that are opaque or exceptional. The phonetically natural classes in the data (high vs. mid, front vs. back, coronal vs. labial) are found, and both phonologically natural classes that are expected (palatalizing/non-palatalizing vowels for the opaque dataset, palatalizing/non-palatalizing consonants for the lexically specific dataset) are also found by the algorithm. The algorithm does find a one additional phonologically natural class each for the opaque and the lexically specific dataset, respectively. For the opaque case, this is a set of consonants created sympathetically during the creation of the palatalizing vowels class (see section 4.3), while for the lexically specific case, this is a pseudo-complement of the set of consonants necessary to account for the lexically specific pattern (see section 4.3).

As indicated in section 2.3, the sets of segments thus found can be used as the basis for building a set of formal features (Mayer & Daland 2020), on the basis of which a more traditional OT grammar can be built. Alternatively, a model such as this, in which all contrasts are explicitly expressed through indexed constraints, could also be taken as an inspiration for a different, more parallel and multilevel account of phonology (cf., e.g., Boersma 2007, Boersma & Van Leussen 2017). Either direction would need further working out and testing.

The current results, while not yet a full-scale case study on real language data, do have interesting implications for the role of contrast in phonological systems. While contrast has previously been conceptualized as a something external to the grammar (e.g., Dresher 2009), something defined over specific words in the lexicon (e.g., Lubowicz 2003, Flemming 2017), or something defined over existing features (e.g., Mackenzie 2016), this paper shows the feasibility of seeing contrast as an emergent property, which comes about as a side effect of learning constraint-based grammars. The added benefit of such an approach is that contrasts are not merely stated, but also automatically incorporated into a grammatical analysis.

Two important directions for future work are working this methodology out to a real language data case study and incorporating statistics into the learner, which will give the learner more flexibility in extracting trends from the lexicon. Statistical models for indexed constraint induction do exist (e.g.,



Nazarov 2018), and further developing these and adapting them to the local contrast detection/natural class learning task would be a natural direction to pursue. However, apart from these immediate direction for development, a step beyond the narrow confines of indexed constraint theory could be a productive direction, as well. For instance, working out similar contrast detection models in the context of Cophonology Theory (Inkelas and Zoll 2007), Harmonic Grammar with Scaling Factors (Hsu & Jesney 2016) and/or mixed effect models (Zymet 2018) would be a very important and worthwhile effort, the results of which would shine more light on the nature of contrast detection and natural classes as arising from general-purpose learning strategies and concurrent learning of representations and grammars. Finally, working out the steps leading from initial phonotactic grammars as learned within the current simulations to a fully fledged OT grammar (as alluded to above) would be a very important step to further contextualize the current work and connect it to other approaches.

## References

- Alderete, John. 2001. *Morphologically governed accent in Optimality Theory*. (Outstanding dissertations in linguistics.) New York: Routledge.
- Becker, Michael. 2009. *Phonological Trends in the Lexicon: The Role of Constraints*. Amherst, MA: University of Massachusetts Amherst dissertation.
- Boersma, Paul. 2007. Some listener-oriented accounts of h-aspiré in French. *Lingua* 117:1989–2054.
- Boersma, Paul & Jan-Willem van Leussen. 2017. Efficient evaluation and learning in multi-level parallel constraint grammars. *Linguistic Inquiry* 48(3):349–88.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Evanston.
- Cristià, Alejandrina, Jeff Mielke, Robert Daland & Sharon Peperkamp. 2013. Constrained Generalization of Implicitly Learned Sound Patterns. *Journal of Laboratory Phonology* 4(2): 259–85.
- Dresher, B. Elan. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.
- Dresher, B. Elan. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41(2):165–81.
- Flemming, Edward. 2017. Dispersion Theory and phonology. In Mark Aronoff (ed.), *The Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- Halle, Morris. 1978. Knowledge unlearned and untaught: What speakers know about the sounds of their language. In Morris Halle, Joan Bresnan & George Miller (eds.), *Linguistic Theory and Psychological Reality*, 294–303. Cambridge, MA: MIT Press.
- Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the early stages. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Fixing Priorities: Constraints in Phonological Acquisition*, 158–203. Cambridge: Cambridge University Press.
- Hsu, Brian & Karen Jesney. 2016. Scalar positional markedness and faithfulness in Harmonic Grammar. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society 51*, 241–55. Chicago, IL: Chicago Linguistic Society.
- Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45(1): 133–71.
- Jarosz, Gaja. 2006. *Rich Lexicons and Restrictive Grammars: Maximum Likelihood Learning in Optimality Theory*. Baltimore, MD: Johns Hopkins University dissertation.
- Kraska-Szlenk, Iwona. 1995. *The Phonology of Stress in Polish*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign dissertation.
- Lubowicz, Anna. 2003. *Contrast Preservation in Phonological Mappings*. Amherst, MA: University of Massachusetts Amherst dissertation.
- Mackenzie, Sarah. 2016. Consonant harmony in Nilotic: Contrastive specifications and Stratal OT. *Glossa* 1(1):12.
- Mayer, Connor. 2020. An algorithm for learning phonological classes from distributional similarity. *Phonology* 37(1):91-131.
- Mayer, Connor & Robert Daland. 2020. A method for projecting features from observed sets of phonological classes. *Linguistic Inquiry* 51(4), 725–63.
- McCarthy, John. 1999. Sympathy and phonological opacity. *Phonology* 16(3):331–99.
- Mielke, Jeff. 2004. *The Emergence of Distinctive Features*. The Ohio State U dissertation.
- Nazarov, Aleksei. 2018. Learning within- and between-word variation in probabilistic OT grammars. In Gillian

- Gallagher, Maria Gouskova & Sora Yin (eds.), *Supplemental Proceedings of the 2017 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.
- Nazarov, Aleksei. 2021. Learnability of indexed constraint analyses of phonological opacity. *Proceedings of the Society for Computation in Linguistics* 4, 16.
- Oostendorp, Marc van. 2008. Incomplete devoicing in formal phonology. *Lingua* 118: 1362–74.
- Parker, Steve & Mónica Parker. 2004. Optimality theory and ethical decision making. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 48:2.
- Pater, Joe. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17(2):237–74.
- Pater, Joe. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, 123–54. London: Equinox.
- Prickett, Brandon & Gaja Jarosz. 2021. Modeling the Acquisition of Phonological Interactions: Biases and Generalization. In Ryan Bennett, Richard Bibbs, Mykel Loren Brinkerhoff, Max J. Kaplan, Stephanie Rich, Amanda Rysling, Nicholas Van Handel & Maya Wax Cavallaro (eds.), *Proceedings of the 2020 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.
- Prince, Alan & Bruce Tesar. 2004. Learning phonotactic distributions. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Fixing Priorities: Constraints in Phonological Acquisition*, 245–91. Cambridge: Cambridge University Press.
- R core team. 2021. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Round, Erich. 2017. Phonological exceptionality is localized to phonological elements: The argument from learnability and Yidiny word-final deletion. In Claire Bowern, Larry Horn & Raffaella Zanuttini (eds.), *On looking into words (and beyond): Structures, relations, analyses*, 59–97. Berlin: Language Science Press.
- Sandstedt, Jade. 2018. Feature specifications and contrast in vowel harmony: The orthography and phonology of Old Norwegian height harmony. Edinburgh: University of Edinburgh dissertation.
- Temkin–Martínez, Michal. 2010. *Sources of Non-Conformity in Phonology: Variation and Exceptionality in Modern Hebrew Spirantization*. Los Angeles, CA: University of Southern California dissertation.
- Tesar, Bruce. 1995. *Computational Optimality Theory*. Boulder, CO: University of Colorado Boulder dissertation.
- Zymet, Jesse. 2018. *Lexical propensities in phonology: Corpus and experimental evidence, grammar, and learning*. Los Angeles, CA: University of California Los Angeles dissertation.