Proceedings of the Society for Computation in Linguistics

Volume 4 Article 16

2021

Learnability of indexed constraint analyses of phonological opacity

Aleksei Nazarov
Utrecht University, a.i.nazarov@uu.nl

Follow this and additional works at: https://scholarworks.umass.edu/scil

Part of the Computational Linguistics Commons, and the Phonetics and Phonology Commons

Recommended Citation
Nazarov, Aleksei (2021) "Learnability of indexed constraint analyses of phonological opacity," *Proceedings of the Society for Computation in Linguistics*: Vol. 4 , Article 16.
DOI: https://doi.org/10.7275/f1zb-5s89
Available at: https://scholarworks.umass.edu/scil/vol4/iss1/16

This Paper is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Proceedings of the Society for Computation in Linguistics by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Learnability of indexed constraint analyses of phonological opacity

Aleksei Nazarov
Utrecht University
a.i.nazarov@uu.nl

Abstract

This paper explores the learnability of indexed constraint (Pater, 2000) analyses of opacity based on the case study of raising in Canadian English (Chomsky, 1964; Chambers, 1973). Such analyses, while avoiding multiple levels of derivation or representation, require the learner to induce indexed constraints, connect these constraints to particular segments in the lexicon, and rank these constraints. An implementation of Round's (2017) learner for indexed constraints, which is an extension of Biased Constraint Demotion (Prince and Tesar, 2004), is used here to test whether a simple learner can rise to this challenge and learn a restrictive analysis of the opaque pattern (i.e., one that restricts raising to its proper phonological context). Three different datasets are used with decreasing evidence for a restrictive analysis, as well as three underlying form hypotheses (two of which entail entertaining multiple underlying forms for the same surface form simultaneously), with decreasing evidence for the phonotactic patterns in the data (cf. Jarosz, 2006). It is found that the learner can find a restrictive analysis of opaque raising in Canadian English, provided that the most informative dataset is used and multiple underlying forms are considered for those data points that contain [t, d, ɾ] after a diphthong.

1 Introduction

To represent phonological opacity (see section 2), Optimality Theory (OT) requires some additional mechanism (Idsardi, 2000; though see Baković, 2011 for some exceptions to this), such as serial extensions of OT (e.g., Bermúdez-Otero, 2003; McCarthy, 2007; Jarosz, 2014). An alternative to such dedicated extensions is to re-use the machinery of indexed constraints (Pater, 2000) already in place to account for lexical exceptions. If indexed constraints refer to individual segments rather than entire morphemes (Round, 2017), a systematic account of opaque mappings is possible, as shown in section 2. Such systematic accounts formalize the link between phonological opacity and exceptionality in phonology. However, they do contain a great amount of additional free parameters (the number and kind of indexed constraints, as well as the number and kind of lexical items attached to each of those). Can such analyses be discovered given a standard OT learner (Biased Constraint Demotion or BCD, Prince and Tesar, 2004) with an indexed constraint-learning extension (Round, 2017), and what are the phonological and morphological requirements on the dataset for these analyses to be discoverable?

The rest of this paper is set up as follows. Section 2 will briefly introduce indexed constraint analyses of opacity, after which section 3 will discuss the inherent learnability challenges. The computational experiment will be described in sections 4 (description of the learner), 5 (simulation set-up), and 6 (results). Section 7 will outline the implications and conclude.

2 Indexed constraint analyses of opacity

Indexed constraints (Kraska-Szlenk, 1995; Pater, 2000) have been proposed as a tool to encode exceptional patterns in the grammar. They are copies of a phonologically defined (universal) constraint that only receive violations for a specified set of morphological affiliations. For instance, a universal constraint like *[+voice] (one violation for every [+voice] segment) might have an indexed variant [+voice]$_i$ (one violation for every [+voice] segment affiliated with a morpheme that has index *i*). This means that morphemes that

Proceedings of the Society for Computation in Linguistics (SCiL) 2021, pages 158-166.
Held on-line February 14-19, 2021

carry the index *i* (e.g. /ada$_i$/) may be prevented from having voiced segments, whereas all other morphemes may have both voiced and voiceless segments, as in Table 1:

/ada$_i$/	*[+vce]$_i$	ID(vce)	*[+vce]
ada$_i$	*!		*
☞ ata$_i$		*	
/ada/			
☞ ada			*
ata		*!	

Table 1: Illustration of indexed constraints.

In phonological opacity, one phonological process creates apparent exceptions to the other process (which is then called opaque; see also Kiparsky, 1973; McCarthy, 1999). For instance, opaque raising in Canadian English (Chomsky, 1964; Chambers, 1973) applies before voiceless consonants only, except when /t/ is flapped to voiced [ɾ], which is a systematic “exception” to this pattern:

(2) /raɪd/→[raɪd] ‘ride’
/raɪt/→[rʌɪt] ‘write’
/raɪt-ɚ/→[rʌɪɾɚ] ‘writer’

An indexed constraint analysis of this pattern may have an indexed constraint against unraised diphthongs [aɪ/ʊ] before consonants with the index *i* (*C$_i$/aɪ_; /t/ in /raɪt$_i$/ is such a consonant). This is illustrated in Table 2.

/raɪt$_i$/	*C$_i$/aɪ_	ID(low)	*C/aɪ_
raɪt$_i$	*!		*
☞ rʌɪt$_i$		*	
/raɪd/			
☞ raɪd			*
rʌɪd		*!	

Table 2: Raising with indexed constraints.

2.1 Exceptionless raising with indices

To make sure we have a restrictive analysis, that is, the process applies strictly before voiceless consonants or instances of [ɾ] that alternate with [t], we need to regulate the set of consonants before which flapping happens. Thus, consonants that trigger raising (those that carry index *i*) are voiceless except when flapping applies. At the same time, raising-triggering consonants may co-occur with non-flapped voiced consonants in the same morpheme (e.g., [rʌɪt$_i$] ‘write’, [bʌɪt$_i$] ‘bite’). This means that the property “triggers raising and is voiceless” cannot belong to an entire morpheme: it is localized to a specific segment (see Nazarov, 2019 for further discussion of this point). Therefore, I adopt a segmentally local variant of indexed constraints (Round, 2017), where each individual segment in each morpheme may have its own index. This allows the constraint *[+voice]$_i$ to require surface voicelessness specifically for consonants indexed *i*, but not for any other consonant in the morpheme. This requirement is then outranked by pro-flapping constraints, which force a voiced [ɾ] outcome in certain environments. In Table 3, /aɪ/ raises before the indexed consonant in both ‘write’ and ‘writer’ due to *C$_i$/aɪ_. The *i* consonant surfaces as voiceless in ‘write’ due to *[+voice]$_i$, but as a voiced [ɾ] in ‘writer’ due to undominated *V{t,d}V and *ɾ̥.[1] Thus, raising only occurs only before instances of a consonant indexed *i*, and consonants indexed *i* must be voiceless except when flapped. In other words, raising occurs only before voiceless consonants, or before flaps that alternate with [t].

The indexed consonant in ‘write’ and ‘writer’ is represented as underlying /d$_i$/ in Table 3 for the sake of Richness of the Base, to explicitly show how *i* consonants are required to be voiceless. The same surface candidates would win if the indexed consonant were represented as /t$_i$/ (see Nazarov, 2019).

/raɪd$_i$/	*ɾ̥	*V{t,d}V	*[+vce]$_i$	*C$_i$/aɪ_	ID(low)	ID(vce)
raɪt$_i$				*!		*
☞ rʌɪt$_i$					*	*
raɪd$_i$			*!	*		
rʌɪd$_i$			*!		*	
/raɪd$_i$ɚ/						
raɪt$_i$ɚ		*!				*
rʌɪt$_i$ɚ		*!			*!	
rʌɪd$_i$ɚ		*!	*			
☞rʌɪɾ$_i$ɚ			*			
rʌɪɾ̥$_i$ɚ	*!					

Table 3: Raising before [ɾ] that alternates with [t].

The discussion above explains how raising *only* happens in the right environment. The other half of an exceptionless account of raising is ensuring that raising *always* happens in the right environment. This is done by also including an undominated

[1] Undominated constraints against voiceless nasals and laterals should also be assumed, as well as high-ranked Faithfulness constraints that would preclude other manner changes (i.e., plosive to fricative) that might let /t/ or /d/ surface as voiceless while obeying these constraints.

constraint against voiceless consonants without the diacritic *i*: *[-voice]$_{[-i]}$ (see Nazarov, 2019 for more details about the use of *[-i]*; see the end of section 4 on how it is implemented in the learner). This constraint makes sure that any consonant without the diacritic *i* surfaces as voiced, even if it is underlyingly voiceless. This means that a consonant that does not trigger raising (i.e., one without the index *i*) must be voiced. In addition, if a consonant that does not trigger raising surfaces as [ɾ], this [ɾ] is prohibited from alternating with [t], since it has the marking *[-i]*, and *[-voice]$_{[-i]}$ prevents it from surfacing as voiceless [t].

This illustrated in Table 4, where an underlying /t/ is marked as *[-i]*. This underlying /t$_{[-i]}$/ shows up either as a [d] without raising or as a [ɾ] without raising. This is due to the constraint *[-voice]$_{[-i]}$, which prevents /t$_{[-i]}$/ from surfacing as [t], and due to ID(low), which prevents /aɪ/ from raising outside the environment before a consonant indexed *i*.

Thus, indexation can yield a restrictive account of raising: raising occurs only and always before either voiceless segments, or before [ɾ] that alternates with [t] on the surface.

/raɪt$_{[-i]}$/	*ɾ̥	*V {t,d} V	*[-vc]$_{[-i]}$	*[+vc]$_i$	*C$_i$/aɪ_	ID (lo)	ID (vc)
raɪt$_{[-i]}$			*!				
rʌɪt$_{[-i]}$			*!			*	
☞raɪd$_{[-i]}$							*
rʌɪd$_{[-i]}$						*!	*
/raɪt$_{[-i]}$ɚ/							
rʌɪt$_{[-i]}$ɚ		*!	*			*	
raɪd$_{[-i]}$ɚ		*!					*
rʌɪɾ$_{[-i]}$ɚ						*!	*
☞raɪɾ$_{[-i]}$ɚ							*

Table 4: No raising before [ɾ] alternating with [d].

2.2 Comparison with other accounts

Various other OT-style accounts have been proposed for Canadian Raising, including Bermúdez-Otero (2003), a serial account, and Pater (2014), a Harmonic Grammar account. Among the non-serial OT accounts, an Output-Output Faithfulness (Benua, 1997) account sketched by Hayes (2004) stands out as a competitor to the current account. Hayes suggests that a high-ranked Output-Output (OO) IDENT-constraint on vowel height ensures that derived forms such as [rʌɪɾɚ] 'writer' and [raɪɾɚ] 'rider' retain the vowel qualities of their respective base forms, [rʌɪt] 'write' and [raɪd] 'ride'. This explains why [ɾ] that alternates with [t] triggers raising, but not [ɾ] that alternates with [d]. At the same time, it does not require the learning of indexed constraints. Why is the complexity of the current account needed?

There are at least two reasons for this. First, Idsardi (2006) mentions a few data points where for him and a few other consultants, raising can lead to alternations in vowel height between base and derived form: [naɪn~nʌɪnθ] 'nine~ninth', [aɪ~ʌɪθ] 'i~ith', [waɪ~wʌɪθ] 'y~yth'. For these forms, the constraint that triggers raising must outrank IDENT-OO(low), while the opposite ranking holds in Hayes' account for 'writer' vs. 'rider' (Hayes, 2004: 190).

Second, OO-Faithfulness does not generalize well to other cases of opacity: it can only account for certain cases opacity where the opaque interaction applies in derived forms. Opaque interactions in which this is not the case, like the ones in Bedouin Arabic (see McCarthy, 2007 for an overview), cannot be accounted for by appealing to OO-Faithfulness. For instance, in the form /gabl/ '...', the epenthetic vowel [i] creates the context in which /a/ would normally raise to [i]; however, this raising does not apply in this form because epenthesis does not feed raising. In this case, there is no morphological base form that /gabl/ is derived from that also has [a] as its first vowel, as would be required for an OO-Faithfulness account. However, an indexation analysis can account for the Bedouin Arabic interactions (see Nazarov, 2020). Broadly speaking, indexation accounts of phonological opacity are an alternative to assuming serialism or opacity-motivated additional representation levels (cf. Turbidity; Goldrick, 2001) in OT.[2] Indexation (Pater, 2000) or other mechanisms of lexicon/phonology interaction like cophonologies (Inkelas and Zoll, 2007) are independently required by exceptionality phenomena.

3 Learnability challenges

Analyses of opacity that make use of indexed constraints do not require the use of multiple

[2] This set of approaches, of course, would also include Sympathy Theory (McCarthy, 1999) and Comparative Markedness (McCarthy, 2003). However, these have their own drawbacks, including the inability to account for multiple interdependent opaque interactions, like in Bedouin Arabic (cf. McCarthy, 2007: 47-56).

derivational levels, which would otherwise pose learnability challenges (see, e.g., Staubs and Pater, 2016). However, such analyses do require the inference of new constraints (which constraints get indexed variants?), the connection of these constraints to particular segments in particular morphemes, and the ranking of these constraints. Is this a problem that can be solved with a (variants of a) standard OT learner? And what kind of data does such a learner need to see to be successful (see, e.g., Kiparsky, 2000; Bermúdez-Otero, 2003 for types of evidence in learning opacity)?

In the case of Canadian English, the indexed constraints *C_i/aɪ_, *[+voice]$_i$, and *[-voice]$_{[-i]}$ must be induced and connected to all underlying segments in the lexicon that surface as voiced; they must be ranked such that *C_i/aɪ_ outranks ID(low), while *[+voice]$_i$, *[-voice]$_{[-i]}$ and *C_i/aɪ_ outrank ID(voice). This will be asked of a variant of Biased Constraint Demotion Prince and Tesar (2004) proposed by Round (2017) that can learn indexed constraints that relate to individual segments. This learner will be briefly described in section 4.

The datasets on which the learner will be tested vary in whether they include forms where the opaque process applies transparently (i.e., there is raising without /t/ showing up as a flap), and whether there are alternations. They will also differ in the range of Underlying Representations considered by the learner for each morpheme. See section 5 for a fuller description of the datasets.

4 The learner

The current simulations are done in a model proposed by Round (2017) that learns segmentally local indexation from winner-loser pair data. This model is an extension of Biased Constraint Demotion (BCD; Prince and Tesar, 2004). BCD is a version of Recursive Constraint Demotion (Tesar, 1995) with a Markedness-over-Faithfulness bias to ensure a maximally restrictive analysis (i.e., phonologically determined patterns are privileged over lexically determined patterns).

BCD starts with no ranking and all winner-loser pairs in the corpus. At each step, it selects only those constraints that prefer no losers (=PNL). Out of PNL, it takes just the Markedness constraints and install them at the bottom of the current ranking, while removing all winner-loser pairs from consideration in which the loser has a greater number of violations on the constraints just installed.

If there are no Markedness constraints among PNL, it selects, instead, the smallest set of Faithfulness constraints that will "free up" a Markedness constraint at the next step. If there are multiple such sets, Prince and Tesar (2004) specify a non-deterministic procedure for choosing among different smallest sets that involves exhaustive search and backtracking. In the implementation (section 5), this latter search algorithm is replaced by randomly picking among the smallest sets.

Round (2017) specifies a procedure that induces segmentally local indexed constraints in this context. Whenever two winner-loser pairs in the data have conflicting ranking requirements (=inconsistency), this means that phonological factors alone cannot decide the winner for each input; in this case, the model induces some indexed constraint (Pater, 2010). Which indexed constraint is induced depends on constraint violation loci (CVL; see also McCarthy, 2007): particular underlying segments whose surface realization violates a particular constraint. For each constraint, the learner works out the number of CVL that only favor winners in winner-loser pairs, $\Phi_W - \Phi_L$. The constraint with the greatest $\Phi_W - \Phi_L$ is selected to be cloned into an indexed version (if several constraints are tied for the greatest $\Phi_W - \Phi_L$, one of these is selected at random), and the winner-favoring CVL for that constraint are given the index corresponding to that constraint. The addition of this constraint resolves the inconsistency and allows BCD to continue as usual until all constraints have been ranked.

In the current implementation, indexation is handled in the following way. Whenever an indexed version of a constraint (C) is induced, it is given a new, unique index, for instance, *i* (so the constraint becomes $C_{i)}$.[3] It is then recorded which segment loci have winner-favoring violation marks for this constraint: those segment loci are recorded as [+*i*]; all other segment loci are recorded as [-*i*]. For the violation profile of C_i, only those violations of C that correspond to [+*i*] segments are kept – all violations that correspond to [-*i*] segments are discarded. When evaluating whether two

[3] For best results, constraint definitions should be such that only one locus of violation is possible: for instance, the constraint *C/aɪ_ is only violated at the consonant that follows a diphthong [aɪ], never at the diphthong itself. For Markedness constraints that allow for multiple violation loci, an extension of the implementation would be necessary.

constraints are plausibly referring to the same index (see the bridging assumption in section 5.4), both [+*i*] and [-*i*] segments are considered.

5 Simulations

The algorithm described in section 4 is a batch algorithm (processes all data at once), categorical (non-probabilistic), but non-deterministic. Because the algorithm is non-deterministic, multiple runs of the algorithm have to be done to ensure that all behaviors of the learner can be observed. However, since the algorithm is not truly probabilistic, the number of runs for which each outcome is observed is not directly meaningful.

The algorithm is tested on three data sets that represent different data patterns. Furthermore, three different Underlying Form (UF) hypotheses were used for each of the data sets. This yields 9 different conditions per test. Each of the 9 conditions was tested 20 times (to ensure all non-deterministic paths were explored) with the same constraint set (explained below).

5.1 Conditions: Surface datasets

The surface datasets offered to the learner consist of the Canadian English dataset in Table 5, which I will henceforth refer to as D_1, along with two variants, D_2 and D_3, which are described below. The words in the datasets are chosen to balance the voicing of consonants and provide the basics of the conditioning of raising.

D_1, D_2, and D_3 differ in the evidence they contain for the opaque pattern: whether there is an alternation in terms of flapping (the process that causes opacity), and whether raising (the opaque process) is attested transparently (in this case, before [t]).

Table 5 shows the segmental loci (correspondence indices) used in D_1 – these are the segments whose violations are considered to be the same for the indexed constraint induction system (see section 5). It can be seen that, in D_1, the stems ‘flight’, ‘glide’, ‘sigh’, and ‘vie’ share the same correspondence indices between underived and derived form. This means that there is an alternation between the voiceless [t] in ‘flight’ and the voiced [ɾ] in ‘flighter’. D_1 also features both transparent raising (‘flight’) and opaque raising (‘flighter’).

Underived form	Derived form
[f$_1$ l$_2$ ʌɪ$_3$ t$_4$] ‘flight’	[f$_1$ l$_2$ ʌɪ$_3$ ɾ$_4$ ɚ$_5$] ‘flighter’
[g$_6$ l$_7$ aɪ$_8$ d$_9$] ‘glide’	[g$_6$ l$_7$ aɪ$_8$ ɾ$_9$ ɚ$_5$] ‘glider’
[s$_{10}$ aɪ$_{11}$] ‘sigh’	[s$_{10}$ aɪ$_{11}$ ɚ$_5$] ‘sigher’
[v$_{12}$ aɪ$_{13}$] ‘vie’	[v$_{12}$ aɪ$_{13}$ ɚ$_5$] ‘vier’

Table 5. Dataset D_1.

Whereas D_1 encodes the morphological relationship between derived and underived forms, D_2 (Table 6) does not: ‘flight’ and ‘flighter’ do not share any loci for the purpose of indexed constraint induction and are treated like a pair of unrelated forms. Therefore, D_2 shows no alternation between voiceless [t] in ‘flight’ and voiced [ɾ] in ‘flighter’. D_1 and D_2 are identical otherwise.

Underived form	Unrelated form
[f$_1$ l$_2$ ʌɪ$_3$ t$_4$] ‘flight’	[f$_{14}$ l$_{15}$ ʌɪ$_{16}$ ɾ$_{17}$ ɚ$_5$] ‘flighter’
[g$_6$ l$_7$ aɪ$_8$ d$_9$] ‘glide’	[g$_{18}$ l$_{19}$ aɪ$_{20}$ ɾ$_{21}$ ɚ$_5$] ‘glider’
[s$_{10}$ aɪ$_{11}$] ‘sigh’	[s$_{22}$ aɪ$_{23}$ ɚ$_5$] ‘sigher’
[v$_{12}$ aɪ$_{13}$] ‘vie’	[v$_{24}$ aɪ$_{25}$ ɚ$_5$] ‘vier’

Table 6. Dataset D_2.

Finally, D_3 not only shows no alternations in terms of flapping, but also shows no transparent application of raising. This is because the underived forms of D_1 have been removed. D_1 and D_3 are identical in all other ways.

Underived form	Derived form
-	[f$_1$ l$_2$ ʌɪ$_3$ ɾ$_4$ ɚ$_5$] ‘flighter’
-	[g$_6$ l$_7$ aɪ$_8$ ɾ$_9$ ɚ$_5$] ‘glider’
-	[s$_{10}$ aɪ$_{11}$ ɚ$_5$] ‘sigher’
-	[v$_{12}$ aɪ$_{13}$ ɚ$_5$] ‘vier’

Table 7. Dataset D_3.

D_1 provides the most evidence for a restrictive opaque analysis (raising just before voiceless consonants or [ɾ] that alternates with [t]), since [ɾ] alternates with [t] and we see raising before [t]. D_2 and D_3 provide limited to no evidence for a restrictive opaque analysis: there is no alternation between [t] and [ɾ] and/or raising before [t].

5.2 Conditions: Underlying Forms

In addition to the surface dataset offered to the learner, the Underlying Forms (UFs) the learner was offered in combination with these surface datasets were varied. Since it is plausible that a learner will not know the correct UFs at the outset of learning, the learner was offered multiple UFs for the same surface data point (except under the UF_3 hypothesis). This corresponds to Jarosz’s

(2006) implementation of the phonotactic learning stage (Hayes, 2004; Prince and Tesar, 2004): the learner is offered tableaux with the same winning output but different inputs. The subsequent selection of a particular input representation for each data point is not modeled here.

Three UF hypotheses are considered: each surface candidate considered for a data point is offered as a potential input (*UF₁*), only variation in the underlying voicing of a post-diphthongal [t] or [d] is considered in the inputs (*UF₂*), or only the canonical inputs for a Canadian raising analysis (Chomsky 1964; Chambers 1973) are considered (*UF₃*). These hypotheses are compared for the word 'flighter' in Table 8.

For each surface data point, the competing surface candidates explore all combinations of consonant voicing, sonorancy of post-diphthongal [t, d, ɾ], and diphthong height. For example, the surface data point [flʌɪɾɚ] has 32 surface candidates (including itself): {f, v}×{l, l̥}×{aɪ, ʌɪ}×{t, d, ɾ, ɾ̥}×{ɚ}. Under hypothesis *UF₁*, this surface data point is offered to the learner in 32 different tableaux, each with a different surface candidate chosen as its UF – see Table 8.

UF₂ holds that all the UF of all consonants except post-diphthongal [t, d, ɾ] equals the surface form. Diphthongs are always unraised (/aɪ/), while, for each post-diphthongal [t, d, ɾ], /t/ and /d/ are both considered as potential underlying variants (see Table 8). *UF₂* reflects the learner's knowing the phonemic contrasts, but not yet knowing the phonemization of instances of [t, d, ɾ] due to voicing and sonorancy neutralization.

Finally, *UF₃* is the same as *UF₂*, except that post-diphthongal [t, d, ɾ] are always given their canonical voicing (/t/ for 'flight', 'flighter', /d/ for 'glide', 'glider'), as also illustrated in Table 8.

UF₁	*UF₂*	*UF₃*
/fl̥aɪtɚ/	/flaɪtɚ/	/flaɪtɚ/
/vl̥aɪtɚ/	/flaɪdɚ/	
/flaɪtɚ/		
...		
/vlʌɪtɚ/		
/fl̥aɪdɚ/		
...		
/fl̥aɪɾ̥ɚ/		
...		
/fl̥aɪɾɚ/		
...		
/vlʌɪɾɚ/		

Table 8. UFs considered for [flʌɪɾɚ] 'flighter'.

UF₁ corresponds to a stage where the learner has not learned anything about the UFs yet, but gives the learner maximal evidence for the phonotactic restriction against /aɪ/ + voiceless consonants as well as the one against [t, d] in the flapping context (since, no matter the underlying form, these phonotactic restrictions are observed).

UF₂ and *UF₃* correspond to a stage where the learner has learned the phonemic contrasts of the language and has (almost) finished learning underlying forms. *UF₂* gives the learner evidence of the phonotactic restriction against [t, d] in the flapping context, while *UF₁* does not necessarily do so.

5.3 Tableau setup: constraint set

Section 5.2 already outlined the surface candidates that are put into each tableau for each data point: these explore all combinations of consonant voicing, sonorancy of post-diphthongal [t, d, ɾ], and diphthong height. For each data point, one or multiple tableaux are considered, depending on the UF hypothesis (section 5.2). The initial constraint set for all simulations is the same, and consists of a range of Markedness constraints regarding voicing, diphthong height, and flapping, as well as Faithfulness constraints for the phonetic features manipulated ([voice], [sonorant], [low]):

*[+voice], *[-voice] (1)
*aɪ, *ʌɪ, *C̥/aɪ_, *C/aɪ_, *C̬/ʌɪ_, *V/ʌɪ_
*ɾ, *ɾ̥, *l̥, *V{t,d}V
IDENT(voice), IDENT(son), IDENT(low)

Each simulation starts with the constraint set in (1), but the learner adds indexed constraints formed from one or several of these constraints, as defined towards the end of section 4.

5.4 Defining restrictiveness

As stated in sections 1 and 2, the desideratum is an analysis of opacity that is restrictive: the opaque process will apply in the correct environment when confronted with new data (i.e., other underlying forms with other patterns of indexation). As mentioned at the end of section 4, the current learner does not reuse indices between indexed constraints, so an analysis exactly like in section 2, where *[+voice]$_i$ and *C$_i$/aɪ_ refer to the same index, *i*, is impossible to obtain directly with this learner. Instead, the following bridging assumption is used:

For any two indexed constraints C_i and C_j, if the segment loci with index *i* are a subset of the segment loci with index *j*, while the segment loci without index *i* ([-*i*] segments) are NOT among the segment loci with index *j*, consider *i* and *j* to be the same index. Index *j* in this context can also be the complement of an index in the analysis: [+*j*] segments = [-*k*] segments. (2)

In other words, if an analysis has a raising constraint (e.g., $*C_i$/aɪ_) that refers to a subset of the consonants that are required to be voiceless (e.g., by being subject to $*[+\text{voice}]_j$ or by being in the complement of the consonants subject to $*[-\text{voice}]_k$), the analysis is considered to be restrictive, because these two constraints are now assumed to refer to the same index. In this case, only consonants that are voiceless or alternate with a voiceless consonant may trigger diphthong raising before them. All other types of analysis are considered non-restrictive, because they will allow raising before voiced consonants other than [ɾ] that alternates with [t] (see also section 2.1).

6 Results

The results of the simulations described above are summarized in Table 9, which displays, for each of the 9 conditions described in section 5, the number of runs out of 20 that converged to a restrictive analysis (as defined at the end of section 5).

	UF₁ (all SFs)	*UF₂* (voicing variation)	*UF₃* (fixed)
D₁ (all data)	*15/20*	*7/20*	*0/20*
D₂ (no alternations)	*0/20*	*0/20*	*0/20*
D₃ (no transparent raising)	*0/20*	*0/20*	*0/20*

Table 9. Number of restrictive outcomes per condition.

As mentioned in section 5, the specific numbers in the cells are not particularly informative, as the learner is not meaningfully probabilistic: it is only the difference between the values 0 (restrictive analysis never found), $0 < x < 20$ (restrictive analysis may be found), and 20 (restrictive analysis always found) that is relevant here. As can be seen, restrictive analyses are found for those conditions in which all data are presented to the learner (D_1) and the UFs exhibit some variation, at the very least in the underlying voicing of [t, d] (UF_1, UF_2). All conditions with D_2 and D_3, as well as all conditions with UF_3, lead to non-restrictive analyses only.

Note that a non-restrictive analysis is always found for all conditions. This has to do with the fact that the current implementation of indexed constraint selection chooses randomly when there is a tie between constraints with the greatest $\Phi_W - \Phi_L$ value, so that the selection of a restrictive analysis is decided by chance if multiple analyses are available. A different, more sophisticated and principled model of selecting indexed constraints might be able to remedy this shortcoming, but the current model was chosen for its simplicity.

However, it must also be noted that indexed constraint analyses could be found for all datasets and all UF hypotheses, regardless: the indexed constraint induction mechanism was able to resolve inconsistency in a way that led to some consistent analysis of the data.

7 Discussion/conclusion

From the results shown in section 7, we can learn (at least) three things. First, the indexed constraint learner described in section 4 can indeed learn restrictive analyses of opaque raising in Canadian English (Chomsky 1964; Chambers 1973) in terms of indexed constraints without derivational ordering. Second, the learner needs to have evidence for a flapping alternation that makes raising opaque (because only the D_1 dataset leads to restrictive analyses) to be able to produce a restrictive analysis. Third, indexed constraint induction must apply before the UFs of all morphemes are completely determined in order to produce a restrictive analysis. Specifically, the learner must have evidence that [t, d] are disallowed in the flapping environment and map to [ɾ] (which is achieved by considering the mappings /t/ → ɾ and /d/ → ɾ).

This result means that indexed constraint analyses of opacity, even though complex (see section 2), are viable even when the learner is maximally simple (section 4). This has implications for evaluating non-derivational accounts of opacity versus derivational ones: the learnability of derivational accounts has been shown before (e.g., Jarosz, 2016), and the learnability of non-derivational accounts with a fixed number of levels of representations has also been shown (Boersma and van Leussen, 2017), but the learnability of an indexed constraint analysis of opacity had not been shown before.

Some important issues for future work remain. One of these is the categorical nature of the current learner, which leads to learnability statistics that are difficult to interpret (see section 6). A probabilistic learner will be able to use more information from the data to choose between various hypotheses, restrictive or not, and will be able to take gradient data into account as well. Finally, a probabilistic learner would make the comparison to other learnability results easier (e.g., Jarosz, 2016; Boersma and van Leussen, 2017; Nazarov and Pater, 2017).

Another issue is the fact that the learner induces each indexed constraint with a new index, not allowing co-indexation between constraints. This is an issue because a restrictive account of Canadian English raising requires two constraints to be co-indexed (see section 2). It has been solved through a bridging assumption in this current implementation (see section 5.4), but a more principled solution within the learner would make the current results even stronger.

Finally, the learnability of a broader range of cases of opacity needs to be considered in this framework: do particular opaque interactions provide greater problems for learning indexed constraints? Are there particular interactions that are problematic? The main result in this paper is an encouraging starting point to start exploring these questions.

References

Baković, Eric. 2011. Opacity and ordering. In John Goldsmith, Jason Riggle, and Alan Yu (eds.), *The Handbook of Phonological Theory*, 2nd ed., 40–67. Wiley-Blackwell. https://doi.org/10.1002/9781444343069.ch2

Benua, Laura. 1997. *Transderivational identity*. PhD dissertation, University of Massachusetts Amherst.

Bermúdez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In Jennifer Spenader, Anders Eriksson, Östen Dahl (eds.), *Variation within Optimality Theory*, 25–36. Stockholm University.

Boersma, Paul, and Jan-Willem van Leussen. 2017. Efficient evaluation and learning in multilevel parallel constraint grammars. *Linguistic Inquiry* 48(3): 349–388. https://doi.org/10.1162/ling_a_00247

Chambers, Jack. 1973. Canadian raising. *Canadian Journal of Linguistics* 18: 113–35. https://doi.org/10.1017/S0008413100007350

Chomsky, Noam. 1964. Current issues in linguistic theory. In J. Fodor, J. Katz (eds.), *The structure of language: Readings in the philosophy of language*, 50–118. Prentice-Hall.

Goldrick, Matt. 2001. Turbid output representations and the unity of opacity. In Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-Yung Kim (eds.), *Proceedings of the Northeast Linguistic Society* 30, 231–245. Amherst, MA: GLSA.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the early stages. In: René Kager, Joe Pater, and Wim Zonneveld (eds.), *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Idsardi, William. 2000. Clarifying opacity. *The Linguistic Review* 17(2–4): 337–50. https://doi.org/10.1515/tlir.2000.17.2-4.337

Idsardi, William. 2006. Canadian raising, opacity, and rephonemization. *Canadian Journal of Linguistics* 51:21–28. https://doi.org/10.1017/S0008413100004011

Inkelas, Sharon, and Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45: 133–171. https://doi.org/10.1515/LING.2007.004

Jarosz, Gaja. 2006. *Rich lexicons and restrictive grammars: Maximum Likelihood learning in Optimality Theory*. PhD dissertation, Johns Hopkins University.

Jarosz, Gaja. 2014. Serial Markedness Reduction. In John Kingston, Claire Moore-Cantwell, Joe Pater, Robert Staubs (eds.), *Proceedings of the 2013 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America. https://doi.org/10.3765/amp.v1i1.40

Jarosz, Gaja. 2016. Learning opaque and transparent interactions in Harmonic Serialism. In Gunnar Ólafur Hansson, Ashley Farris-Trimble, Kevin McMullin, Douglas Pulleyblank (eds.), *Proceedings of the 2015 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America. https://doi.org/10.3765/amp.v3i0.3671

Kiparsky, Paul. 1973. Abstractness, opacity and global rules. In: Osamu Fujimura (ed.), *Three Dimensions of Linguistic Theory*, 57–86. Tokyo: TEC.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17 (2–4): 351–66. https://doi.org/10.1515/tlir.2000.17.2-4.351

Kraska-Szlenk, Iwona. 1995. *The phonology of stress in Polish*. PhD dissertation, University of Illinois, Urbana-Champaign.

McCarthy, John. 1999. Sympathy and phonological Opacity. *Phonology* 16(3): 331–399. https://doi.org/10.1017/S0952675799003784

McCarthy, John. 2003. Comparative markedness. *Theoretical Linguistics* 29: 1–51. https://doi.org/10.1515/thli.29.1-2.1

McCarthy, John. 2007. *Hidden generalizations: Phonological opacity in optimality theory*. London: Equinox Publishing.

Mielke, Jeff, Mike Armstrong, and Elizabeth Hume. 2004. Looking through opacity. *Theoretical Linguistics* 29(1–2): 123–139. https://doi.org/10.1515/thli.29.1-2.123

Nazarov, Aleksei. 2019. Formalizing the connection between opaque and exceptionful generalizations. *Toronto Working Papers in Linguistics*, 41:1. https://doi.org/10.33137/twpl.v41i1.32767

Nazarov, Aleksei. 2020. Bedouin Arabic multiple opacity with indexed constraints in Parallel OT. In Hyunah Baek, Chikako Takahashi, and Alex Hong-Lun Yeung (eds.), *Supplemental Proceedings of the 2019 Annual Meeting on Phonology,* Washington, DC: Linguistic Society of America. https://doi.org/10.3765/amp.v8i0.4687

Nazarov, Aleksei, and Joe Pater. 2017. Learning opacity in Stratal Maximum Entropy Grammar. *Phonology* 34(2): 299–324. https://doi.org/10.1017/S095267571700015X

Pater, Joe. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17(2): 237–74. https://doi.org/10.1017/S0952675700003900

Pater, Joe. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, 123–54. Equinox.

Prince, Alan, and Bruce Tesar. 2004. Learning phonotactic distributions. In René Kager, Joe Pater, and Wim Zonneveld (eds.), *Fixing Priorities: Constraints in Phonological Acquisition*, 245–91. Cambridge: Cambridge University Press.

Round, Erich. 2017. Phonological exceptionality is localized to phonological elements: The argument from learnability and Yidiny word-final deletion. In Claire Bowern, Laurence Horn, Raffaella Zanuttini (eds.), *On looking into words (and beyond): Structures, relations, analyses*, 59–97. Jena and Berlin: Language Science Press. https://doi.org/10.5281/zenodo.495467

Staubs, Robert and Joe Pater. 2016. Learning serial constraint-based grammars. In: John McCarthy and Joe Pater (eds.), *Harmonic Grammar and Harmonic Serialism*. London: Equinox Press.

Tesar, Bruce. 1995. *Computational Optimality Theory*. PhD dissertation, University of Colorado.