

# *Post-nasal voicing in Japanese classifiers as exceptional triggering: implications for Indexed Constraint Theory*

PHILLIP BURNES

*University of Ottawa*

[pburn036@uottawa.ca](mailto:pburn036@uottawa.ca)

and

KEVIN McMULLIN

*University of Ottawa*

[kevin.mcmullin@uottawa.ca](mailto:kevin.mcmullin@uottawa.ca)

---

## *Abstract*

Indexed constraints are often used in constraint-based phonological frameworks to account for exceptions to generalizations. A point of contention in the literature on constraint indexation revolves around indexed markedness constraints. While some researchers argue that only faithfulness constraints should be indexed, others argue that markedness constraints should be eligible for indexation as well. This article presents data from Japanese for which a complete synchronic analysis requires indexed markedness constraints but argues that such constraints are only necessary in cases where a phonological repair applies across a morpheme boundary. We then demonstrate that algorithms for learning grammars with indexed constraints can be augmented with a bias towards faithfulness indexation and discuss the advantages of incorporating such a bias, as well as its implications for the debate over the permissibility of indexed markedness constraints.

**Keywords:** exceptions, learning, indexed constraints, Japanese phonology

## *Résumé*

Les exceptions aux généralisations sont souvent captées dans les cadres théoriques basés sur les contraintes par l'usage des contraintes indexées. Dans la littérature sur les contraintes indexées, il y a une controverse entourant les contraintes de marquage indexées. Tandis que certains chercheurs affirment que seules les contraintes de fidélité devraient être éligibles pour

---

For helpful feedback on the material presented here, we thank the audience at the 2018 Montréal-Ottawa-Toronto Phonetics/Phonology Workshop (McMaster University, Ontario), three anonymous CJL/RCL reviewers, and the CJL/RCL editors. This research was supported by the Social Sciences and Humanities Research Council of Canada.

l'indexation, d'autres insistent que les contraintes de marquage devraient elles aussi être admissibles à l'indexation. Cet article présente des données japonaises dont l'analyse synchronique nécessite des contraintes de marquage indexées, mais avance que de telles contraintes ne sont nécessaires que pour des cas dans lesquels une réparation phonologique franchit une frontière morphologique. Nous démontrons ensuite qu'il est possible d'incorporer un biais vers l'indexation des contraintes de fidélité dans les algorithmes qui apprennent les grammaires comportant des contraintes indexées et discutons des avantages d'un tel biais, ainsi que de ses implications pour le débat sur l'admissibilité des contraintes de marquage indexées.

**Mots-clés:** exceptions, apprentissage, contraintes indexées, phonologie japonaise

## 1. INTRODUCTION

Japanese phonology contains many inconsistencies, seemingly driven by multiple sub-lexicons and morpheme-specific processes. On the one hand, there are sounds that are contrastive in some portions of the lexicon but exist only as allophones elsewhere. For example, [f] exists only as an allophone of /h/ before [u] in native Japanese words (e.g., /pune/ → hune<sup>1</sup> → [fune] 'boat') but can appear before any vowel in loanwords from English (e.g., *fairu* 'file', *feisubukku* 'Facebook', *fōku* 'fork', *firutā* 'filter', and *futto-bōru* 'football'). On the other hand, there are processes that apply in some areas of the lexicon but are inert elsewhere. For example, the famous process of *rendaku* or sequential voicing applies to compounds whose second member is a native Japanese word, but generally does not apply to compounds whose second member is a loanword (Itō and Mester 2003). This aspect of Japanese phonology inspired the introduction of indexed constraints (Fukazawa 1998, Fukazawa et al. 1998, Itō and Mester 1999, Itō and Mester 2001), which have remained influential within constraint-based frameworks as a method for modelling various kinds of irregularity. The descriptive utility of indexed constraints is unquestionable, and a method for learning them has been developed by Pater (2004, 2006, 2010) as a supplemental module to the Error-Driven Constraint Demotion algorithm for learning constraint rankings (Tesar and Smolensky 1998). There are, however, some unresolved theoretical questions surrounding indexed constraints.<sup>2</sup>

The most prominent of these has been the debate in the literature over whether we must allow only faithfulness constraints to be indexed (Fukazawa 1998, Fukazawa et al. 1998, Itō and Mester 1999, 2001) or whether we must allow both faithfulness and markedness constraints to be indexed (Pater 2000, 2004, 2006, 2010; Flack 2007, Gouskova 2007, 2012; Jurgec 2010). Furthermore, with respect to learnability, how does an algorithm choose which constraint to copy when it has more than one option? In this article, we will add to the evidence that markedness constraints should be eligible for indexation by analyzing some data from Japanese. We

<sup>1</sup>Many instances of [h] and [f] in native Japanese words likely derive from /p/ since they become voiced [b] through the process of sequential voicing (*rendaku*) in compounds. For example, /tasuke/ 'help' + /pune/ 'boat' → [tasuke-bune] 'lifeboat' and not \*[tasuke-vune].

<sup>2</sup>Abbreviations: EDCD: Error-Driven Constraint Demotion; L: loser; OT: Optimality Theory; RCD: Recursive Constraint Demotion; ROTB: Richness of the Base; W: winner.

furthermore argue that indexed markedness constraints are only necessary when *exceptional triggering* (Pater 2010) occurs across a morpheme boundary, and that they should be restricted to just these circumstances. Using the data from Japanese, we go on to show that learning algorithms should be biased towards indexing faithfulness constraints where there is a choice, since doing so will create indexed markedness constraints in only the desired cases. We then discuss the implications this bias has for the debate over whether to allow the indexation of markedness constraints.

The Japanese phenomenon in question is the sporadic post-nasal voicing that occurs in Japanese counter words, which are combinations of a numeral and a classifier morpheme. Many classifier morphemes begin with a voiceless obstruent, and these morpheme-initial obstruents sometimes become voiced when preceded by the morpheme-final nasal consonant of certain numerals, but not in every case that we would expect. The initial obstruents of some classifier morphemes like *kai* ‘times’ and *soku* ‘footwear’ become voiced after the numerals *san* ‘three’, *sen* ‘thousand’, and *man* ‘ten thousand’, while the initial obstruents of other classifiers like *satu* ‘books’ and *kyaku* ‘chairs’ never become voiced at all (Itō and Mester 2003: 140). Interestingly, a small subset of classifier morphemes is intermediate between the other two subsets. They obligatorily become voiced after *san* ‘three’, but either completely resist becoming voiced after *sen* ‘thousand’ and *man* ‘ten thousand’ (e.g., the classifier *ken* ‘houses’) or else optionally voice after these two numerals (e.g., the classifier *hiki* ‘small animals’ (Itō and Mester 2003: 140). Even stranger is the fact that all classifier-initial obstruents fail to become voiced after the numeral *yon* ‘four’, despite it too ending in a nasal and thus being eligible to trigger post-nasal voicing (Itō and Mester 2003: 140). The above generalizations are summarized in Table 1 below (modified from Itō and Mester 2003: 140), where all instances of obligatory post-nasal voicing are presented in shaded cells.<sup>3</sup> Please note that this is not an exhaustive list of counter words in Japanese.

Item	Classifier	Numeral				
		2 /ni-/	3 /san-/	1000 /sen-/	10,000 /man-/	4 /yon-/
books	/-satu/	ni-satu	san-satu	sen-satu	man-satu	yon-satu
chairs	/-kyaku/	ni-kyaku	san-kyaku	sen-kyaku	man-kyaku	yon-kyaku
shots	/-patu/	ni-hatu	san-patu	sen-patu	man-patu	yon-patu
houses	/-ken/	ni-ken	san-gen	sen-ken	man-ken	yon-ken
animals	/-piki/	ni-hiki	san-biki	sen-hiki	man-hiki	yon-hiki
cups	/-pai/	ni-hai	san-bai	sen-bai	man-bai	yon-hai
footwear	/-soku/	ni-soku	san-zoku	sen-zoku	man-zoku	yon-soku
times	/-pen/	ni-hen	san-ben	sen-ben	man-ben	yon-hen

**Table 1.** Japanese counter words (modified from Itō and Mester 2003: 140)

<sup>3</sup>The form *san-hiki* ‘three animals’ and *sen-hiki* ‘thousand animals’ optionally voice the classifier’s initial obstruent, changing it to [b] instead of debuccalizing it to [h].

The crucial piece of information here is that not all nasal-final numerals are equal in terms of triggering post-nasal voicing. We will argue that without a bias toward indexing faithfulness constraints, the learner can be led to believe it is the classifier that exceptionally triggers voicing (rather than the numeral) - a hypothesis that turns out to be wrong and will need to be corrected.

Before moving on, we would like to acknowledge the comments of two anonymous reviewers on the relationship between [p], [b], and [h] in this data set. One reviewer points out that the presence of /p/ → [p] in this data set's column for *yon* is problematic for Itō and Mester's (2003) suggestion that the lack of voicing after *yon* may stem from earlier forms with a synonymous numeral *si* (see section 2). The other reviewer points out that we never see non-alternating [h] in this data set and asks how this fact would fit into our analysis of the voicing. We unfortunately could not further investigate these aspects of the data without straying from the focus of the article, although we suspect that they would be amenable to an analysis using indexed constraints.

The remainder of this article is structured as follows. Section 2 analyzes the Japanese counter word data within the context of the debate over whether to permit indexed markedness constraints. The section shows that indexed markedness constraints are necessary, but also argues that they should be limited to cases where the exceptional triggering of a process crosses a morpheme boundary. Section 3 demonstrates that Pater's (2004, 2006, 2010) Inconsistency Resolution algorithm for learning indexed constraints can handle the Japanese data, though it runs the risk of creating superfluous indexed markedness constraints that end up serving no purpose. Section 4 argues that the accidental proliferation of indexed constraints can be avoided by appealing to a preference for indexing faithfulness constraints over markedness constraints, and discusses the implications this bias has for the debate over whether to allow the indexation of markedness constraints. Finally, section 5 concludes and highlights directions for future research.

## 2. JAPANESE POST-NASAL VOICING AND CONSTRAINT INDEXATION

Indexed constraints were introduced to account for the common situation in which certain morphemes are treated differently than others, requiring a ranking of two (or possibly more) constraints that contradicts the ranking of those same constraints needed elsewhere in the language. This is precisely the case for the Japanese data above, where the presence and lack of post-nasal voicing require essentially opposite constraint rankings. In Optimality Theory (OT; Prince and Smolensky 2004) and other constraint-based phonological frameworks, the standard constraints that together produce post-nasal voicing are \*NC̰ and IDENT[voice]-IO, defined in (1) and (2) respectively

- (1) \*NC̰: Assign a violation for every sequence of a nasal stop immediately followed by a voiceless obstruent.
- (2) IDENT[voice]-IO:<sup>4</sup> Assign a violation for every input segment whose corresponding output segment does not have the same value for the feature [voice].

<sup>4</sup>We shorten this to Id[voi] in tableaux

If \*NC̥ is ranked above IDENT[voice]-IO, then any underlying voiceless obstruent will become voiced when it immediately follows a nasal obstruent in the output. This is as shown in (3a) for the counter word *san-zoku* ‘three pairs of footwear’. Underlying voiceless obstruents surface faithfully under the same ranking when they are not immediately preceded by a nasal stop, as shown in (3b) for the counter word *ni-soku* ‘two pairs of footwear’ (in fact they surface faithfully in this context no matter the constraint ranking).

(3) a.

	/san-soku/	*NC̥	ID[voi]
	[san-soku]	*!	
☞	[san-zoku]		*

b.

	/ni-soku/	*NC̥	ID[voi]
☞	[ni-soku]		
	[ni-zoku]		*!

This is of course not the whole story in the sub-lexicon of Japanese counter words. Take for instance the counter word *san-satu* ‘three volumes’, wherein the underlying /s/ does not voice to [z] despite immediately following [n]. To generate the mapping /san-satu/ → [san-satu], IDENT[voice]-IO must dominate \*NC̥, but it is impossible for \*NC̥ to both dominate and be dominated by IDENT[voice]-IO. Such ranking paradoxes are problematic and serve as the primary motivation for indexed constraints. Indexation enables constraints to occupy more than one position in the hierarchy, and as a result, ranking paradoxes are in many cases no longer an issue. The ranking paradox above can be resolved in one of two ways.

The first option, as illustrated in (4), is to create a copy of the faithfulness constraint IDENT[voice]-IO, indexing it such that it applies only to /satu/, and ranking it above \*NC̥. This high-ranking indexed constraint exempts /satu/ (and only /satu/) from the pressure to undergo post-nasal voicing and is akin to saying that faithfulness to underlying voice features is more important for /satu/ than for unindexed morphemes such as /soku/.

(4) a.

	/san-soku/	ID[voi] <sub>satu</sub>	*NC̥	ID[voi]
	[san-soku]		*!	
☞	[san-zoku]			*

b.

	/san-satu/	ID[voi] <sub>satu</sub>	*NC̥	ID[voi]
☞	[san-satu]		*	
	[san-zatu]	*!		*

Alternatively, as illustrated in (5), we could create a copy of the markedness constraint \*NC̥, indexing it such that it applies only to /soku/, and ranking it above IDENT[voice]-IO. This high-ranking constraint places /soku/ (and only /soku/) under greater

pressure to undergo post-nasal voicing. This is akin to saying that \*NC<sub>◌</sub> clusters are worse when the voiceless obstruent comes from the morpheme /soku/ rather than an unindexed morpheme.

(5) a.

	/san-soku/	*NC <sub>◌soku</sub>	ID[voi]	*NC <sub>◌</sub>
	[san-soku]	*!		*
☞	[san-zoku]		*	

b.

	/san-soku/	*NC <sub>◌soku</sub>	ID[voi]	*NC <sub>◌</sub>
☞	[san-satu]			*
	[san-zatu]		*!	

Either of these solutions is satisfactory if we limit ourselves to counter words with the numeral *san* ‘three’ and numerals without a final nasal obstruent like *ni* ‘two’. Once we exit this idealized sphere, it becomes apparent that the indexed faithfulness solution is the correct one for the data at hand. Before showing this, however, a brief note on the evaluation of indexed constraints is in order. The literature on indexed constraints generally adopts the notion of *local application*, defined in (6), whereby indexed constraints are only violated when the offending material is linked directly to underlying material in the indexed morpheme (Pater 2006, 2010).

- (6) Local application of \*X<sub>L</sub>: Assign a violation mark to any instance of the configuration X that contains a phonological exponent of a morpheme specified as L (Pater 2010).

In the case of faithfulness constraints like IDENT[voice]-IO, the offending material is an output segment that does not match its input correspondent along some feature. Accordingly, a copy of IDENT[voice]-IO indexed to a morpheme can only prevent a change of voicing in segments that are a part of that very morpheme. In contrast, markedness constraints like \*NC<sub>◌</sub> are violated by a specific configuration of material in an output string, so an indexed markedness constraint can be violated by a substring that crosses a morpheme boundary, as long as that substring contains the output correspondent of an input segment in the indexed morpheme. Accordingly, if a copy of \*NC<sub>◌</sub> is indexed to a nasal-final numeral (such as *san* ‘three’), then that numeral can induce post-nasal voicing in an immediately following classifier, since the nasal in the offending NC<sub>◌</sub> cluster was provided by the indexed numeral.

With local application in mind, we can now tackle the counter words with the numeral *yon* ‘four’, all of which categorically show no post-nasal voicing. Itō and Mester (2003) point out that the lack of voicing after *yon* ‘four’ may have arisen by historical accident. The language previously used the Sino-Japanese numeral *si* ‘four’ when counting, but since this numeral is homophonous with *si* ‘death’, it was replaced by the native Japanese numeral *yon* ‘four’. The original numeral *si* ‘four’ did not end in a nasal and so never induced post-nasal voicing, and it would seem that the lack of post-nasal voicing in modern *yon* ‘four’ forms arose through some kind of faithfulness to the original *si* ‘four’ forms (see OO-correspondence in Benua 1995). A synchronic OT analysis ideally should not have recourse to historical

faithfulness, however, and so we propose the following analysis. First, we rank generic IDENT[voice]-IO above generic  $*NC_{\circ}$ , which accounts for the lack of voicing in forms with *yon* ‘four’, as shown in (7a). Second, we index  $*NC_{\circ}$  to the numeral *san* ‘three’ and rank the copy above generic IDENT[voice]-IO, which accounts for the presence of voicing in some forms with *san* ‘three’, as shown in (7b). Finally, we index IDENT[voice]-IO to classifiers like /*satu*/ that categorically resist voicing, and rank these copies above  $*NC_{\circ san}$  to ensure the classifiers never voice, as shown in (7c).

(7) a.

	/yon-soku/	ID[voi] <sub>satu</sub>	$*NC_{\circ san}$	ID[voi]	$*NC_{\circ}$
☞	[yon-soku]				*
	[yon-zoku]			*!	

b.

	/san-soku/	ID[voi] <sub>satu</sub>	$*NC_{\circ san}$	ID[voi]	$*NC_{\circ}$
	[san-soku]		*!		*
☞	[san-zoku]			*	

c.

	/san-satu/	ID[voi] <sub>satu</sub>	$*NC_{\circ san}$	ID[voi]	$*NC_{\circ}$
☞	[san-satu]		*		*
	[san-zatu]	*!		*	

One final quirk of the Japanese data is the fact that the classifiers *ken* ‘houses’ and *hiki* ‘animals’ have post-nasal voicing after *san* ‘three’ but not after *sen* ‘thousand’ or *man* ‘ten thousand’. This can be interpreted to mean that such classifiers are less faithful to their underlying voicing features than classifiers like *satu* ‘volumes’, but at the same time are more faithful to their underlying voicing features than classifiers like *soku* ‘footwear’. This can be captured by inserting the sub-ranking IDENT[voice]-IO<sub>ken</sub> >  $*NC_{\circ man}$  below  $*NC_{\circ san}$  but above unindexed IDENT[voice]-IO. The tableaux in (8a) and (8b) show how *ken* ‘houses’ resists voicing after *yon* ‘four’ and *man* ‘ten thousand’. That *ken* nonetheless voices after *san* ‘three’ is shown in (8c).

(8) a.

	/yon-ken/	$*NC_{\circ san}$	ID[voi] <sub>ken</sub>	$*NC_{\circ man}$	ID[voi]	$*NC_{\circ}$
☞	[yon-ken]					*
	[yon-gen]		*!		*	

b.

	/man-ken/	$*NC_{\circ san}$	ID[voi] <sub>ken</sub>	$*NC_{\circ man}$	ID[voi]	$*NC_{\circ}$
☞	[man-ken]			*		*
	[man-gen]		*!		*	

c.

	/san-ken/	$*NC_{\circ san}$	ID[voi] <sub>ken</sub>	$*NC_{\circ man}$	ID[voi]	$*NC_{\circ}$
	[san-ken]	*!				*
☞	[san-gen]		*		*	

It is important to note that while this analysis can generate all of the observed surface forms and is thus observationally adequate, it does not reflect the intuition that the default behaviour of classifiers should be to remain voiceless no matter the numeral. Classifiers are part of the Sinitic stratum of the Japanese lexicon, and this stratum freely allows NC clusters (Itō and Mester 2003). The above analysis, however, predicts that post-nasal voicing after the numeral *san* ‘three’ is the default behaviour. To better reflect the limitation of voicing to a minority of classifiers, we could emulate Itō and Mester’s (2003) analysis of the counter word data by adopting Mascaró’s (2007) treatment of allomorphy. This approach posits that underlying forms can consist of a partially ordered set of allomorphs and includes a constraint PRIORITY that is violated when the selected allomorph is not the first in the partial order. Supposing that the voiceless allomorphs are preferred over the voiced allomorphs (which we denote by writing, for example, /soku > zoku/), we can generate the forms with *yon* by ranking PRIORITY above generic \*NC as shown in (9a). Note how IDENT[voice]-IO is not violated by either candidate, since the voiced and voiceless obstruent are part of the underlying form. The preference for the voiceless allomorph is overridden in forms with *san* ‘three’ because \*NC<sub>san</sub> dominates PRIORITY, as shown in (9b). Finally, the categorical lack of voicing in other classifiers (like *satu* ‘volumes’) results from their lack of a dual underlying form; the high-ranking generic IDENT[voice]-IO penalizes the candidate with voicing, as shown in (9c).<sup>5</sup>

(9) a. 

	/yon-soku>zoku/	ID[voi]	*NC <sub>san</sub>	PRIORITY	*NC
☞	[yon-soku]				*
	[yon-zoku]			*!	

b. 

	/san-soku>zoku/	ID[voi]	*NC <sub>san</sub>	PRIORITY	*NC
	[san-soku]		*!		*
☞	[san-zoku]			*	

c. 

	/san-satu/	ID[voi]	*NC <sub>san</sub>	PRIORITY	*NC
☞	[san-satu]		*		*
	[san-zatu]	*!			

All that being said, whether or not we make use of PRIORITY, we crucially need indexed markedness constraints (\*NC<sub>san</sub>, \*NC<sub>sen</sub>, and \*NC<sub>man</sub>) to account for the varying degrees to which Japanese numerals can trigger voicing in classifier morphemes. This is an interesting result considering the controversy over indexed markedness constraints in the literature. Some researchers restrict indexation to faithfulness constraints only (Fukazawa 1998, Fukazawa et al. 1998, Itō and

<sup>5</sup>Not shown here is that we need indexed \*NC<sub>sen</sub> and \*NC<sub>man</sub> as well as some instances of indexed PRIORITY to achieve the intermediate levels of post-nasal voicing exhibited by the classifiers *ken* ‘houses’ and *hiki* ‘animals’.

Mester 1999, 2001), while others allow indexation to apply to both faithfulness and markedness constraints (Pater 2000, 2004, 2006, 2010; Flack 2007, Gouskova 2007, 2012; Jurgec 2010, Jurgec and Bjorkman 2018). The two main arguments for permitting markedness indexation are that it allows us to analyze a number of patterns that would otherwise defy explanation (Pater 2000, Flack 2007, Gouskova 2007, 2012; Jurgec 2010, Jurgec and Bjorkman 2018) and that it also allows us to distinguish between the *exceptional blocking* of a process and the *exceptional triggering* of a process, which are the result of indexed faithfulness constraints and indexed markedness constraints, respectively (Pater 2004, 2006, 2010). The Japanese post-nasal voicing data are a rather interesting case with regards to this distinction, since they certainly contain exceptional triggering (*san* ‘three’ triggers voicing where *yon* ‘four’ does not) and, depending on our analytical choices, can be said to contain exceptional blocking (*soku* ‘footwear’ voices after *san* ‘three’ whereas *satu* ‘volumes’ does not).

On the other end of the debate, there are three main arguments against using indexed markedness constraints. The first argument is that indexed markedness constraints make it difficult to capture the asymmetrical implicational patterns in nativization processes and phoneme inventories (Itō and Mester 1999, 2001). For example, the cross-linguistic generalization that the presence of labial stops in an inventory implies the presence of velar stops but not vice versa (Maddieson 1984) requires a rigid ranking of \*LABIAL > \*VELAR (Itō and Mester 1999, 2001). The second argument is that in order to evaluate an indexed markedness constraint, we must allow it to somehow access underlying representations, even though markedness constraints are supposed to evaluate surface forms only (Itō and Mester 1999). The final argument is that by allowing both markedness and faithfulness constraints to be indexed, we create a “too many options” problem whereby, given an exceptional item, an indexed markedness constraint or an indexed faithfulness constraint can often achieve the same result, but there is no principled reason to choose one solution over the other (Inkelas and Zoll 2007). In light of these drawbacks, it is important to determine which types of patterns absolutely require indexed markedness constraints, and ideally find a principled way to limit the indexation of markedness constraints to just these cases.

We argue that the Japanese data necessarily require indexed markedness constraints precisely because the triggering of post-nasal voicing occurs across a morpheme boundary. When exceptional triggering does not cross a morpheme boundary, it is essentially preventing a marked configuration from appearing within a single morpheme, which can be re-analyzed as the exceptional blocking of an input-output discrepancy in the morphemes that do not undergo the process. In these cases, the choice between using indexed faithfulness constraints over indexed markedness constraints depends only on whether we view the undergoers or resisters of a process as the default (i.e., unindexed) items. When exceptional triggering *does* cross a morpheme boundary, however, it is preventing a marked configuration that is formed of material coming from the exceptional morpheme and some other morpheme, and when the marked configuration is avoided by changing material in the non-exceptional morpheme (as in 7b and 8c), these cases cannot be

re-analyzed as exceptional blocking. Only indexed markedness constraints can generate such exceptional behaviour because, unlike indexed faithfulness constraints, they can be violated by material that straddles a morpheme boundary, thus allowing one morpheme to cause changes to material in a different morpheme.

This observation presents a means to reconcile the two perspectives on indexed markedness constraints: we can acknowledge that they are necessary and should be permissible, while also acknowledging that they are powerful and should perhaps be limited to just those cases where exceptional triggering crosses a morpheme boundary. We will argue below that this restriction need not be mere stipulation and can be derived as the consequence of a learning bias. The argument comes in two parts. First, we will demonstrate in section 3 that current algorithms for learning indexed constraints are able to handle the Japanese data, although the order in which the data are presented can cause a learner to pick incorrect indexation targets, thus populating the hierarchy with duplicate constraints that serve no purpose. We then show in section 4 that these mistaken indexations happen because it is harder to identify the source of exceptional triggering than it is to identify the source of exceptional blocking. Furthermore, determining which morphemes are exceptional blockers makes it easier to correctly identify exceptional triggers. We accordingly propose that the search for exceptional blockers should be prioritized over the search for exceptional triggers, which translates into a preference for indexing faithfulness constraints over indexing markedness constraints. By treating indexed markedness constraints as a last resort in this way, a learner will only create indexed markedness constraints when it witnesses exceptional triggering across a morpheme boundary, exactly as desired.

### 3. LEARNING THE JAPANESE DATA

The current algorithm for learning indexed constraints is Pater's (2004, 2006, 2010) Inconsistency Resolution algorithm, which acts as a supplement to any of the Constraint Demotion learning algorithms for OT (Tesar and Smolensky 1993, 1995, 1998). For the purposes of this article, we use the Error-Driven Constraint Demotion algorithm (EDCD; Tesar and Smolensky 1993, 1995, 1998) as the base. The algorithm is error-driven in that the learner updates its current hypothesis when given a datum that it is unable to handle. How the learner determines that its current hypothesis is incorrect is rather simple. Every time the learner is presented with an input-output mapping, it runs the input through its current grammar. If the output produced by its current grammar matches the observed output, no learning takes place, since according to the evidence at hand, the learner has the correct grammar. If the output produced by its current grammar does not match the observed output, the learner's current grammar must be adjusted. In order to identify the constraints that must be re-ordered, the learner creates a mark-data pair. A mark-data pair is essentially a row from a comparative tableau (Prince 2002): it compares how well the winning (observed) candidate and the losing (predicted) candidate satisfy each constraint. If the winning candidate incurs fewer violations than the losing candidate on some constraint, then that constraint prefers the winner (marked in the tableau as

W). Similarly, if the winning candidate incurs more violations than the losing candidate on some constraints, then that constraint prefers the loser (marked in the tableau with L). If both candidates incur an equal number of violations on some constraint, then the constraint prefers neither (marked as an empty square in the tableau).

Every time the learner creates a mark-data pair, it will run the Constraint Demotion process against its accumulated set of pairs. This loop, described informally in (10), will adjust the constraint hierarchy such that it predicts the correct output.

- (10) i. Start at the highest level in the constraint hierarchy
- ii. Find constraints that prefer only winners (i.e., whose columns contain no L marks) and place these in the current level of the constraint hierarchy.
- iii. **Temporarily** delete the mark-data pairs (i.e., the rows) for which there was a W mark in those constraints' columns, then **temporarily** delete those constraints' columns.
- iv. If there are remaining mark-data pairs, move one level down in the hierarchy (creating said level if it does not already exist), then repeat (ii) and (iii).
- v. When no mark-data pairs remain, place the remaining constraints below the last-constructed level.

It is important that the deletion of mark-data pairs be temporary such that deletion lasts only while Constraint Demotion is being run. As soon as the Constraint Demotion loop terminates, the full database of mark-data pairs created up until the present is fully restored, and the loop will be run on this full set every time a new mark-data pair is added. By letting the learner keep track of all evidence that has led it to its current grammar hypothesis, it can avoid entering an infinite loop when given inconsistent data (Tesar 1997).

A reviewer asks why we use EDCD as opposed to Tesar and Smolensky's (1995) Recursive Constraint Demotion (RCD). The two algorithms are highly similar, the main difference being that RCD operates on a pre-determined set of mark-data pairs, typically constructed by the analyst. Our main reason for choosing EDCD over RCD is that we believe the former better reflects the notion that a learner should come equipped only with the minimum of information necessary to succeed. In the Japanese case, this turns out to be just the universal constraint set and a representative list of input-output pairs. Furthermore, as we show below, the constraint(s) selected for indexation by Pater's (2004, 2006, 2010) Inconsistency Resolution algorithm can vary according to the exact set of mark-data pairs. When the learner is working with an incomplete set of mark-data pairs, it might seem that two different indexed constraints can resolve the same ranking paradox, when in fact only one of these is a valid solution in the context of a more complete set of pairs. The goal of this and the next section is to demonstrate that the conditions under which a learner avoids adopting such "eventually incorrect" solutions are also the conditions under which a learner creates indexed markedness constraints only when there is exceptional triggering across a morpheme boundary.

With these reasons for choosing EDCD in mind, let us have it attempt to learn the Japanese data as an illustrative example. For reasons of space, we limit this learning demonstration to counter words with the numerals *san* ‘three’ and *yon* ‘four’; the remaining nasal-final numerals *sen* ‘thousand’ and *man* ‘ten thousand’ are not crucial to further discussion. Suppose we present the learner with the mappings /san-soku/ → [san-zoku] ‘three pairs of footwear’ and /san-satu/ → [san-satu] ‘three volumes’ in alternation. These mappings require opposite rankings of the constraints \*NC̣ and IDENT[voice]-IO, so if mark-data pairs are permanently deleted in the learning loop, the learner will perpetually need to create either the mark-data pair in (11a) or the one in (11b), and will have access to only one of these pairs at a time. When the learner creates (11a), it sees that the constraint \*NC̣ prefers only winners, and so the learner responds by ranking \*NC̣ above IDENT[voice]-IO. When the learner creates (11b), it sees that the constraint IDENT[voice]-IO prefers only winners, and so the learner responds by ranking IDENT[voice]-IO above \*NC̣. The learner will therefore constantly switch between two contradictory rankings.

- (11) a. 

	ID[voi]	*NC̣
/san-soku/ → [san-zoku] Loser: [san-soku]	L	W
- b. 

	ID[voi]	*NC̣
/san-satu/ → [san-satu] Loser: [san-zatu]	W	L

If, on the other hand, mark-data pairs are only temporarily deleted, the learner will eventually create both pairs and stall since neither constraint prefers only winners according to the current set. *Inconsistency detection* is the general name given to the ability of a learner to use a set of mark-data pairs to notice when no ranking exists that generates the currently observed data. This ability has found various uses in the literature on learning in OT in addition to the learning of exceptions (e.g., Prince 2002, Tesar et al. 2003, McCarthy 2005, Tesar and Prince 2007, Merchant 2008, Merchant and Tesar 2008, Akers 2012, Tesar 2012, 2014a, 2014b, 2016), though it is outside the scope of this article to discuss all such uses.

Either constraint, when indexed appropriately, is sufficient to resolve the inconsistency. Suppose we choose the first option and create a copy of \*NC̣ indexed to *soku*; The set of mark-data pairs that results from adding this indexed constraint is shown in (12a). The mark-data pair set is no longer inconsistent, so we can resume constraint demotion. The constraint \*NC̣<sub>soku</sub> is the only one that prefers only winners, so we place it at the top. This eliminates the first mark-data pair (row i), and now IDENT[voice]-IO prefers only winners, so we can place it beneath \*NC̣<sub>soku</sub>. Doing so eliminates the remaining mark-data pair (row ii), so we place the remaining constraint \*NC̣ at the bottom, which gives us the ranking \*NC̣<sub>soku</sub> > IDENT[voice]-IO > \*NC̣. The result of choosing the second option, creating a copy of IDENT[voice]-IO and indexing it to *satu*, is shown in (12b). We invite the reader

to verify that running the constraint demotion steps on this set of mark-data pairs will yield the ranking  $\text{IDENT}[\text{voice}]\text{-IO}_{\text{satu}} > *NC_{\circ} > \text{IDENT}[\text{voice}]\text{-IO}$ .

(12) a.

	ID[voi]	*NC <sub>◦</sub>	*NC <sub>◦soku</sub>
i. /san-soku/ → [san-zoku] Loser: [san-soku]	L	W	W
ii. /san-satu/ → [san-satu] Loser: [san-zatu]	W	L	

b.

	ID[voi]	*NC <sub>◦</sub>	ID[voi]-IO <sub>satu</sub>
i. /san-soku/ → [san-zoku] Loser: [san-soku]	L	W	
ii. /san-satu/ → [san-satu] Loser: [san-zatu]	W	L	W

Not only does this example illustrate how inconsistency resolution works, it is also an example of the “too many options” problem pointed out by Inkelas and Zoll (2007): either an indexed markedness constraint or an indexed faithfulness constraint can eliminate the inconsistency that the learner has found, and the data encountered so far offer no reason to choose one over the other. When more than one constraint is found to be eligible for indexation, Pater (2004, 2006, 2010) gives two additional search heuristics. First, the learner selects the constraint that prefers the winner for the smallest set of morphemes (Pater 2004, 2006, 2010). Second, if there is still more than one choice, the learner selects a markedness constraint rather than a faithfulness constraint (Pater 2004, 2006). If there is still more than one choice, the learner will select randomly (Pater 2004). We adopt these heuristics for the remainder of this section, though we will discuss them further in the next section, where we propose an alternative.

Both \*NC<sub>◦</sub> and IDENT[voice]-IO prefer the winner for all instances of one out of the three morphemes encountered thus far, and so the bias towards indexing markedness constraints is the deciding factor at this point. Accordingly, the algorithm opts for creating \*NC<sub>◦soku</sub> and will build the constraint hierarchy \*NC<sub>◦soku</sub> > IDENT[voice]-IO > \*NC<sub>◦</sub>. Recall, however, that the counter *yon* ‘four’ does not trigger voicing in any classifiers. This ranking therefore cannot handle the mapping /yon+soku/ → [yon-soku] ‘four pairs of footwear’. The learner instead predicts \*[yon-zoku] which, when added to the set of mark-data pairs, creates yet another inconsistent set, summarized in (13).

(13)

	ID[voi]	*NC <sub>◦</sub>	*NC <sub>◦soku</sub>
i. /san-soku/ → [san-zoku] Loser: [san-soku]	L	W	W
ii. /san-satu/ → [san-satu] Loser: [san-zatu]	W	L	
iii. /yon-soku/ → [yon-soku] Loser: [yon-zoku]	W	L	L

No markedness constraint prefers the winner for all instances of a given morpheme, but IDENT[voice]-IO prefers the winner for all instances of *satu* ‘volumes’. Creating indexed IDENT[voice]-IO<sub>satu</sub> and placing this constraint at the top of the hierarchy eliminates one row and one column of the data set as shown in (14) where the cells deleted on this round are marked in grey. The remaining rows and columns are inconsistent, so the learner stalls again.

(14)

	ID[voi]	*NC <sub>◦</sub>	*NC <sub>◦soku</sub>	ID[voi] <sub>satu</sub>
i. /san-soku/ → [san-zoku] Loser: [san-soku]	L	W	W	
ii. /san-satu/ → [san-satu] Loser: [san-zatu]	W	L		W
iii. /yon-soku/ → [yon-soku] Loser: [yon-zoku]	W	L	L	

Now \*NC<sub>◦</sub> prefers the winner for all instances of *san* ‘three’ so it may be indexed to this morpheme. The new constraint \*NC<sub>◦san</sub> is the only constraint that prefers only winners and the learner therefore adds it to the current level in the hierarchy, right below IDENT[voice]-IO<sub>satu</sub>. The resulting data set is shown in (15), where the cells deleted on this round are marked in grey.

(15)

	ID[voi]	*NC <sub>◦</sub>	*NC <sub>◦soku</sub>	*NC <sub>◦san</sub>
i. /san-soku/ → [san-zoku] Loser: [san-soku]	L	W	W	W
iii. /yon-soku/ → [yon-soku] Loser: [yon-zoku]	W	L	L	

This leaves us with a single mark-data pair, whose only W mark is in the column for general IDENT[voice]-IO. The learner thus places IDENT[voice]-IO below \*NC<sub>◦san</sub>, and since this leaves us with no more mark-data pairs, the remaining constraints \*NC<sub>◦</sub> and \*NC<sub>◦soku</sub> get placed at the bottom of the hierarchy below IDENT[voice]-IO. At this point, the learner has essentially found the target grammar and will only make an incorrect prediction when presented with a mapping in which *san* ‘three’ is specifically followed by a new “resisting” classifier that behaves like *satu*. Such mappings will render the mark-data pair set inconsistent, causing the learner to create an additional copy of IDENT[voice]-IO since this is the only constraint that prefers the winner for the newly observed morpheme. Placing this new copy at the top of the hierarchy alongside IDENT[voice]-IO<sub>satu</sub> eliminates the inconsistency introduced by the newly observed mapping.

The above shows that Pater’s (2004, 2006, 2010) Inconsistency Resolution algorithm can handle the Japanese counter word data, but the solution it finds for the data has an interesting quirk. Namely, if the learner has so far encountered only mappings without any voicing, presenting it with a mapping that contains both voicing and a new classifier will lead it to create a copy of the markedness constraint \*NC<sub>◦</sub> indexed to the classifier it just encountered. A copy of this markedness constraint

indexed to the classifier does indeed resolve the inconsistency at that exact time, but further learning eventually reveals that the classifier was not actually the trigger of the voicing. Consequently, further copies of \*N<sub>Ḷ</sub> are then created that correctly identify some of the numerals as the triggers of voicing, and any copies of \*N<sub>Ḷ</sub> indexed to classifiers get relegated to the bottom of the hierarchy, serving no ultimate purpose.

For reasons of economy and elegance it is worthwhile to ask if we can ensure that such superfluous constraints are never created in the first place. We argue in the next section that reversing Pater's markedness bias (i.e., preferring indexed faithfulness constraints when there is a choice) does just that. This is because, while it is possible to misidentify the source of exceptional triggering, it is impossible to misidentify the source of exceptional resistance. We also argue that implementing such a faithfulness bias has beneficial theoretical consequences. These include the fact that indexed markedness constraints are only created when there is exceptional triggering across a morpheme boundary, which is desirable in light of the controversy over permitting the indexation of markedness constraints.

#### 4. INDEXED MARKEDNESS AS A LAST RESORT ONLY

Recall from the previous section that Pater's (2004, 2006, 2010) Inconsistency Resolution algorithm includes two criteria for selecting an indexation target in the event that its main search criterion (find a constraint that prefers only winning candidates relative to some morpheme) produces more than one option. First the learner selects the constraint that prefers the winner for the smallest set of morphemes (Pater 2004, 2006, 2010). Second, if there is a tie for the smallest set of indexable morphemes, the learner selects a markedness constraint rather than a faithfulness constraint, when possible (Pater 2004, 2006). The latter heuristic is incompatible with our claim from section 2 that indexed markedness constraints should be limited to exactly those cases where an indexed faithfulness constraint is incapable of generating the observed mappings. While we could simply stipulate that Pater's bias should be replaced by a faithfulness bias in order to accommodate the above claim, this section provides two further reasons for considering a faithfulness bias. First, we show how a faithfulness bias can handle the cases that Pater (2004, 2006) argues require a markedness bias. Second, we show that the faithfulness bias circumvents the danger of incorrectly identifying exceptional triggers (as described in the previous section) and has several other positive theoretical consequences.

Before discussing the markedness bias, however, we would like to briefly comment on the preference for index classes with fewer members in Pater's (2004, 2006, 2010) algorithm. This preference is there to ensure that the learned constraint ranking reflects the language's overall phonology by reducing the number of indexed morphemes (i.e., as much of the phonology as possible is due to general constraints). It is, however, based on the fact that Pater's indexation operation will create a single copy of a constraint and index it to all morphemes for which it prefers only winners. We instead assume that a separate indexed constraint is made for each exceptional morpheme. This idea goes at least as far back as Itō and Mester (1999) who argue

in an appendix that large-scale stratal indices such as “foreign” or “Sino-Japanese” are likely a group of lexically specific indices occupying the same spot along the hierarchy. Other arguments for completely individualizing indexation include Coetzee and Pater’s (2011) demonstration that doing so allows us to match lexically specific rates of variation with near perfect accuracy, and Moore-Cantwell and Pater’s (2016) demonstration that doing so can account for gradience in phonotactic well-formedness while still capturing the more fixed pronunciations of individual words. We suggest another reason for treating indexation as an individualized operation: during online learning, it is not always guaranteed that all the words belonging to some necessary index class are present in the current mark-data set, and until the learner has seen a sufficiently large amount of data, it is possible for the exceptions to outnumber the regular items. We do, however, leave open the possibility of including a post-learning operation for conflating indexed constraints where possible. Such an operation is included in the learning algorithms of Itō and Mester (1999), Pater (2005), and Coetzee and Pater (2006).

Moving on to the markedness bias: Pater (2004, 2005) includes it in his algorithm in order to address the *subset problem* explored elsewhere in the Optimality Theoretic literature by Smolensky (1996) and Prince and Tesar (2004). Put simply, the subset problem arises when a learner adopts a constraint hierarchy that accepts all the output forms that the target grammar does, but also accepts more; if the learner continues to see only output forms produced by the target *subset* grammar, it will never encounter evidence that it has chosen an overly permissive *superset* grammar. Prince and Tesar (2004) consequently propose that learners should rank markedness constraints as high as possible and rank faithfulness constraints as low as possible, which ensures that the learned grammar permits the smallest set of output forms consistent with the observed data (i.e., the least permissive language). Pater (2004, 2006) suggests that his bias towards indexing markedness constraints reflects the spirit of this solution to the subset problem. We do not believe that this is necessarily the case. To see why, consider a single markedness constraint  $M$  and a single faithfulness constraint  $F$  that together can produce some alternation (i.e., when they are ranked as  $M > F$ ). A learner that is shown both undergoers and non-undergoers of this process will learn the ranking  $M_X > F > M$  under a bias towards indexing markedness constraints and will learn the ranking  $F_X > M > F$  under a bias towards indexing faithfulness constraints. Notice how the default ranking from the markedness bias ( $F > M$ ) is in fact more permissive than the default ranking from the faithfulness bias ( $M > F$ ).

That being said, Pater (2004, 2006) presents a case where a learner can seemingly be tricked into adopting a superset grammar unless said learner is biased towards indexing markedness constraints. In this special case, which Pater (2004, 2006) calls *exceptional blocking by markedness*, exceptional items fail to undergo a process demanded by a ranking  $M_1 > F$  in order to avoid violating a different markedness constraint  $M_2$ . This situation can be modelled by indexing exceptional items to a copy of  $M_2$ , but the same result is also achievable by indexing exceptional items to a copy of  $F$  instead. The faithfulness solution, however, misses the generalization that the exceptional morphemes are exceptional precisely because they are in some sense affected more strongly by  $M_2$  than are regular items. The problem, then, is that we want all exceptional items to potentially violate  $M_2$  by virtue of having a particular

phonological shape, but we can in theory create exceptions that do not have this profile. This is not necessarily an issue in our eyes. Exceptions are created only as needed in our algorithm and so exceptions with the wrong phonological profile will simply never arise. Nonetheless, for transparency's sake we discuss a concrete example from Pater (2004, 2006) more thoroughly in the paragraphs below.

To illustrate the problem, Pater (2006) presents a hypothetical language in which codas are deleted except in a handful of monosyllabic items. The ranking  $\text{NoCODA} > \text{MAX-IO}$  is responsible for the deletion of codas. Furthermore, placing  $\text{MINIMALWORD}$  (all words must be minimally bi-moraic) below  $\text{NoCODA}$  ensures that even monosyllabic words drop their codas. Tableaux for a monosyllable and disyllable are provided in (16a) and (16b) respectively.

(16) a.

	/tak/	NoCODA	MINIMALWORD	MAX-IO
☞	[ta]		*	*
	[tak]	*!		

b.

	/papak/	NoCODA	MINIMALWORD	MAX-IO
☞	[papa]			*
	[papak]	*!		

Given this basic ranking, placing either the indexed faithfulness constraint  $\text{MAX-IO}_X$  or the indexed markedness constraint  $\text{MINIMALWORD}_X$  above  $\text{NoCODA}$  could generate exceptional monosyllabic words that keep their coda. However, the first option, an indexed faithfulness account, fails to capture the generalization that all exceptions are monosyllabic. A learner that opts for  $\text{MAX-IO}_X$  has chosen a superset grammar that can generate exceptional monosyllabic and polysyllabic items as opposed to a subset grammar that can generate only exceptional monosyllabic items. As shown in (17a), a disyllabic word indexed to  $\text{MAX-IO}_X$  would retain its coda. A disyllabic word indexed to  $\text{MINIMALWORD}_X$ , however, would still delete its coda because it has two moras with or without its coda, as shown in (17b).

(17) a.

	/papak <sub>X</sub> /	MAX-IO <sub>X</sub>	NoCODA
	[papa]	*!	
☹	[papak]		*

b.

	/papak <sub>X</sub> /	MINIMALWORD <sub>X</sub>	NoCODA
☞	[papa]		
	[papak]		*!

Pater (2004, 2006) therefore argues that, when there is a choice, markedness constraints should be chosen as indexation targets over faithfulness constraints, but his analysis hinges on letting Richness of the Base (ROTB; Prince and Smolensky 2004) applies to indices. According to this assumption, the rich base contains inputs of the

shape /X/<sub>Y</sub> where X ranges over an infinitude of phonological structures and Y is either empty (i.e., the input is unindexed) or is one of an infinitude of indices. We contend, however, that ROTB does not apply to indices in the same way that it does to the phonological content of underlying forms. While there is no restriction on which and how many indices can exist, indices are not present *a priori* in the universal rich base and arise only where necessary through learning.<sup>6</sup> For example the rich base contains the input /papak/ (and an infinitude of other unindexed input strings) but does not also contain inputs of the shape /papak/<sub>X</sub> where X stands in for an infinitude of possible indices. In the absence of exceptional items larger than one syllable, a learner that has opted to copy MAX-IO will not create copies for any disyllabic or larger items since there is no evidence that such copies are necessary. This means that the learner’s grammar will effectively be the desired subset grammar, whether or not the learner is aware of the fact that all exceptions are monosyllables. Knowledge of this shared property among exceptions is therefore not essential to learning appropriate indexed constraints, and furthermore the speaker could still be aware of the shared monosyllabicity without it being encoded directly into the grammar. Such knowledge would resemble the ability to find words that rhyme. In this way, cases of blocking by markedness can also be analyzed as blocking by faithfulness, and consequently we do not require a bias towards indexing markedness constraints.

To summarize the argumentation thus far, we have provided a theoretical reason to prefer faithfulness indexation over markedness indexation (indexed markedness constraints are powerful and need restriction) and have shown that the arguments given for the reverse preference do not categorically point to its necessity. We can go a bit further now and show that the faithfulness bias has other positive consequences, thus solidifying its beneficial status. Recall from the last section that the first inconsistency our learner encountered was the set of mark-data pairs in (18). This set was created after seeing both an undergoer and a non-undergoer of post-nasal voicing that contain the same numeral. The set is inconsistent because no constraint prefers only winners (they both have an L in their column somewhere).

(18)

	ID[voi]	*NC
i. /san-soku/ → [san-zoku] Loser: [san-soku]	L	W
ii. /san-satu/ → [san-satu] Loser: [san-zatu]	W	L

According to the main search criterion of the Inconsistency Resolution algorithm, this learning “dead end” can be escaped in two ways. One way is to copy

<sup>6</sup>A reviewer asks whether abandoning indexical ROTB also abandons one of Indexed Constraint Theory’s fundamental aims, which is to differentiate exceptional mappings from truly impossible mappings within a grammar. We believe that it does not, since the final grammar adopted by our learner will accept/produce only those exceptional mappings that result from the indexed constraints that were induced. Constraints that were not copied and then indexed will categorically exclude other mappings as impossible.

IDENT[voice]-IO and index it to *satu* ‘volumes’ since the constraint prefers only winners for all instances of that morpheme. The other way is to copy  $*N\underset{\circ}{C}$  and index it to *soku* ‘footwear’, again since the constraint prefers only winners for all instances of that morpheme. The first option correctly identifies *satu* as an exceptional blocker of voicing, but the latter option incorrectly identifies *soku* as an exceptional undergoer of voicing. With its preference for indexing markedness constraints, Pater’s (2004, 2006, 2010) algorithm will opt for the latter escape hatch, but this will eventually be corrected once the learner has enough data to realize that the voicing was actually exceptionally triggered by *san* ‘three’. We contend that due to the mechanics of local application (indexed constraints are only violated when the offending material is linked to the indexed morpheme: (Pater 2006, 2010)), a bias towards indexing faithfulness constraints reduces the risk of such misidentification.

The misidentification occurs precisely because indexed markedness constraints are readily violated by phonological material that crosses a morpheme boundary. For example,  $*N\underset{\circ}{C}_X$  is violated by a sequence of a nasal and a voiceless obstruent when at least either the nasal or the obstruent is linked back to the indexed morpheme. Indexed faithfulness constraints, on the other hand, are violated only when the entirety of some “offensive configuration” is linked back to the indexed morpheme. For example, IDENT[voice]-IO<sub>X</sub> is only violated when an input element from the indexed morpheme changes its voicing value in the output. Accordingly, when creating indexed faithfulness constraints, it is typically clear which morpheme is resisting the regularly imposed phonological change,<sup>7</sup> but when creating indexed markedness constraints, it is often unclear which morpheme is triggering the regularly absent process. Biasing indexation to faithfulness constraints is essentially assuming, as much as possible, that exceptional items are exceptional resisters. Such a strategy is consistent with attempting to keep relations between markedness constraints identical for all morphemes in a language and thus ensuring that the learned hierarchy reflects the language’s overall phonotactics as much as possible. By seeking out opportunities for indexing faithfulness constraints first, the learner starts by looking for all items in its current data set that resist a process, which makes it easier to find the exceptional triggers of that same process, if any exist.

Aside from this small gain in learning efficacy, the bias towards faithfulness indexation gives us two further advantages. First of all, the bias alleviates the “too many options” problem that was raised by Inkelas and Zoll (2007), according to which an indexed markedness constraint or an indexed faithfulness constraint can often achieve the same result, but there is no principled reason to choose one solution over the other. Since any of the possible indexed markedness solutions will only be considered by the learner when there is no indexed faithfulness solution available, the

---

<sup>7</sup>Hypothetical cases for which it would not be apparent which morpheme exceptionally resists a phonological change can be constructed. For example, a language could enforce a process of degemination or vowel shortening when identical segments are concatenated across a morpheme boundary. If the process failed to occur when a certain morpheme was involved, it would not be clear from a single data point which of the two morphemes should be indexed. We are unaware of any such processes being attested cross-linguistically.

learner inherently deals with fewer decisions that have ambiguous answers. Second, by considering indexed markedness only as a last resort, we more easily generate the asymmetrical implicational patterns in nativization processes and phoneme inventories that are discussed by Itō and Mester (1999, 2001).

Itō and Mester (1999, 2001) show that where  $n$  changes are applicable to a given word, there are  $2^n$  logically possible outputs, yet languages often permit only up to  $n + 1$  different outputs. This is because, rather than being independent, the changes enter into an asymmetrical implicational relationship. If we assign a numerical order to the changes, with a higher number meaning less importance to the grammar, we typically observe that applying change  $x$  cannot be done without also applying changes up to and including  $x - 1$  (i.e., all the changes that are more important than change  $x$ ). Conversely, one can apply the changes up to and including  $x - 1$  without also having to apply change  $x$ . For example the French word [ʒɔ̃g|œ:ʁ] ‘juggler’ can undergo four changes when being nativized by a German speaker:  $ʒ \rightarrow j$ ,  $\tilde{\text{ɔ}} \rightarrow \text{ɔ}$ ,  $\text{œ} \rightarrow \text{ø}$ , and  $\text{ʁ} \rightarrow \text{r}$  (Itō Mester, 2001). This would give  $2^4 = 16$  logically possible pronunciations, although only *five* are actually produced by German speakers, giving us the following implicational hierarchy: if  $ʒ \rightarrow j$  then  $\text{ʁ} \rightarrow \text{r}$ , if  $\text{ʁ} \rightarrow \text{r}$  then  $\text{œ} \rightarrow \text{ø}$ , and if  $\text{œ} \rightarrow \text{ø}$  then  $\tilde{\text{ɔ}} \rightarrow \text{ɔ}$  (Itō and Mester 2001). Furthermore, Itō and Mester (1999, 2001) point out that phoneme inventories tend to be constructed according to implicational scales such that the presence of a more marked sound along a given scale implies the presence of all less marked sounds on that scale. For example, Maddieson (1984: 35) found that “an implicational hierarchy can be set up such that the presence of /p/ implies the strong likelihood of the presence of /k/, which similarly implies presence of /t/”. Itō and Mester (1999, 2001) analyze the above asymmetrical implications by having a rigid hierarchy of markedness constraints. By allowing indexed markedness constraints only as a last resort in our learning algorithm, we maximally preserve a language’s markedness hierarchy, making it more likely for the observed asymmetrical implications to emerge.

## 5. DISCUSSION AND CONCLUSION

A synchronic analysis of post-nasal voicing in the Japanese counter word data requires both indexed faithfulness constraints and indexed markedness constraints, thus adding to the evidence in favour of allowing markedness constraints to be indexed. We furthermore argued that indexed markedness constraints are only necessary when exceptional triggering occurs across a morpheme boundary (like in the Japanese data), and in light of the various arguments that have been raised against indexed markedness constraints, we argue that they should be restricted to just these circumstances. Current learning algorithms can handle the Japanese data, but they run the risk of generating redundant constraints that serve no purpose in the final grammar, precisely because they are biased to choose markedness constraints over faithfulness constraints when there is more than one option for indexation. An algorithm biased towards indexing faithfulness constraints does not encounter this problem, and in fact will only create indexed markedness when there is exceptional triggering across a morpheme

boundary, exactly as desired. Furthermore, such a faithfulness bias can mitigate the “too many options” problem (Inkelas and Zoll 2007) faced by theories that permit the indexation of both faithfulness and markedness constraints, and can more easily generate the asymmetrical implicational patterns present in nativization processes and phoneme inventories (Itō and Mester 1999, 2001).

Like any analysis, though, our proposal that indexation is biased to faithfulness constraints comes with its limitations, the two most important of which are that we assume the learner starts with a full constraint set and already knows the underlying form of all morphemes. Recent work has investigated potential methods for inducing the constraints themselves (e.g., Hayes and Wilson 2008) without relying on the standard assumption that the constraint set is innate or universal. Furthermore, it is obviously false that learners come equipped with knowledge of underlying representations, and recent work has shown that under the assumption that the learning target is an output-driven map, inconsistency detection makes it possible to learn constraint rankings and underlying forms side by side (Merchant 2008, Merchant and Tesar 2008, Akers 2012, Tesar 2012, 2014a, 2014b, 2016). It is interesting that a continually updated set of mark-data pairs is versatile enough to signal that a word might be an exception or that a word may need a different underlying form, and future work may be able to create a hybrid approach to resolving inconsistency that both updates underlying forms and indexes constraints as appropriate. Three further limitations of the current proposal are its incorrect predictions with respect to the productivity of a phonological alternation, its reliance on the strict dominance relations between constraints in standard OT, and its reliance on exceptionality being an all-or-nothing property of a morpheme, each of which are discussed in the paragraphs below.

The proposed algorithm will converge on an indexed faithfulness grammar where one is available, a bias which we have argued is desirable. That being said, there are still cases in which we would prefer the indexed markedness solution, and these have to do with the proper prediction of default phonology. Cases of exceptional triggering that do not cross a morpheme boundary can be generated with indexed markedness ( $M_X > F > M$ ), but they can easily be re-analyzed as cases of exceptional blocking ( $F_X > M > F$ ) if the exceptions are re-labelled as regular items and vice versa. The only diagnostic between the two analyses would be the number of items that undergo versus resist the alternation. The proposed algorithm, however, always picks the indexed faithfulness solution no matter the ratio between undergoers and resisters, which makes the rather unintuitive prediction that in cases where all but one word in a language resists an alternation, the one undergoer will be labelled the default by virtue of never causing the creation of an indexed copy of the faithfulness constraint in question, and every other morpheme will be labelled an exception by virtue of triggering the creation of such an indexed constraint. If this were truly how default phonology was determined, we would expect speakers in Wug tests to always apply an alternation wherever they can to nonce words if at least one word of the language exhibits it, whereas we normally expect them to extend productive alternations. As a tentative solution, we could implement the proposal of Burnes (2016) who suggests that when an alternation is discovered to be unproductive, the exceptional undergoers should store all of their allomorphs in

their underlying forms. Storing the allomorphs in the underlying form exempts the exceptional undergoers from the relevant faithfulness constraint(s), thus allowing the faithfulness constraint(s) to be ranked higher than would be the case otherwise, letting the constraint(s) reflect the behaviour of regular morphemes. This of course raises the question of how an algorithm would create multiple underlying forms for a single morpheme. Multiple underlying forms are perhaps not the only possible remedy, but the incorrect predictions vis-à-vis productivity are a crucial weakness of the current proposal and will require further research.

On the subject of productivity, a reviewer asks how many counter words exist in Japanese, and whether the numeral-classifier combinations that exhibit exceptional phonology are simply stored in the lexicon. Work by Yang (2016) proposes a specific threshold number of exceptions past which a rule becomes unproductive and should be abandoned. This threshold is equal to the number  $N$  of potential undergoers divided by the natural log of  $N$ . For example, a rule applicable to 100 items can tolerate up to about 22 exceptions before it becomes more worthwhile to simply memorize what happens to each item individually. Does the rule of post-nasal voicing after *san* ‘three’ meet this standard for productivity? This is difficult to answer since counter words are generally written in *kanji* (the logographic component of the Japanese writing system) which do not indicate voicing. One means to investigate the rule’s productivity (or lack thereof) would be to conduct a wug test where native Japanese speakers are taught new objects that require a nonce classifier and then are asked to count them. We leave such an investigation for future research.

The Error-Driven Constraint Demotion algorithm (Tesar and Smolensky 1998) and the Inconsistency Resolution algorithm (Pater 2004, 2006, 2010) both assume that constraints are related to each other through a relation of strict dominance. Recently, however, the field has been moving from frameworks where constraints are in strict dominance relationships towards frameworks where constraints are assigned numerical weights. Such weight-based frameworks are able to model variable rates of application as well as gradient acceptability and include Stochastic OT (Boersma 1997), Harmonic Grammar (Legendre et al. 1990; see Pater 2009 for a recent overview), Noisy Harmonic Grammar (Coetzee and Pater 2011), and Maximum Entropy grammars (Goldwater and Johnson 2003, Hayes and Wilson 2008). It is therefore worthwhile to ask how indexed constraints would be learned in a weight-based framework. The learning algorithms for the above weight-based frameworks are relatively similar, all using iterative procedures that converge on the optimal constraint weights. As they are currently formulated, however, these algorithms have no way of distinguishing cases in which a morpheme is inconsistent with itself (e.g., /pad/ variably surfaces as [pat] or [pad]) from cases in which a morpheme is inconsistent with another morpheme (e.g., /pad/ generally surfaces as [pat] whereas /kad/ generally surfaces as [kad]). The former *intra-lexical inconsistency* can be captured by noisy evaluation and arguably does not necessitate constraint indexing, but the latter *cross-lexical inconsistency* cannot always be resolved with noisy evaluation and very likely requires one or more indexed constraints. Further research is needed to determine how a learning algorithm for frameworks using constraint

weights could find exceptional morphemes when the cross-lexical inconsistency caused by these exceptions is obscured by intra-lexical inconsistency.

Finally, the learning simulations above assume that exceptionality is an all-or-nothing property of a morpheme, but recent work has found that the exceptional status of a morpheme can be “revoked” in the presence of other morphemes (Jurgec 2014, Gouskova and Linzen 2015, Jurgec and Bjorkman 2018). For example, some Dutch speakers produce [ɹ] in recent loans from English like the name Op[ɹ]ah, but when derivational morphology is added to those loanwords, these same speakers produce the native segment [ʁ], like in the diminutive Op[ʁ]ah-tje (Jurgec 2014, Jurgec and Bjorkman 2018). Jurgec and Bjorkman (2018) analyze these data in OT and propose that indexed constraints specify a *property* and a *domain* such that they apply only when all morphemes within the specified domain have the specified property. We suspect that it will not be too difficult to modify current constraint indexing algorithms such that, after determining that an indexed constraint is necessary, they can determine the domain within which that indexed constraint must apply. Whether our proposed faithfulness bias will be necessary in such a modified algorithm remains to be seen.

## REFERENCES

- Akers, Crystal. 2012. Commitment-based learning of hidden linguistic structures. Doctoral Dissertation, Rutgers University, New Brunswick.
- Benua, Laura. 1995. Identity effects in morphological truncation. In *UMOP 18: Papers in Optimality Theory*, ed. Jill Beckman, Laura Walsh Dickey, Suzanne Urbanczyk, 77–136. Amherst, Massachusetts: GLSA.
- Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, vol. 21, 43–58.
- Burness, Phillip. 2016. Unproductive alternations and allomorph storage: The case of Sino-Japanese. *Journal of Japanese Linguistics* 32: 74–94.
- Coetzee, Andries and Joe Pater. 2006. Lexically ranked OCP-Place constraints in Muna. [ROA-842].
- Coetzee, Andries and Joe Pater. 2011. The place of variation in phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, Jason Yu, 401–431. Oxford, 2nd.
- Flack, Kathryn. 2007. Templatic morphology and indexed markedness constraints. *Linguistic Inquiry* 38: 749–758.
- Fukazawa, Haruka. 1998. Multiple input-output faithfulness relations in Japanese. Unpublished manuscript.
- Fukazawa, Haruka, Mafuyu Kitahara and Hajime Ono. 1998. Lexical stratification and ranking invariance in constraint-based grammars. *Proceedings of the Chicago Linguistics Society* 34(2), ed. M. Catherine Gruber, Derrick Higgins, Kenneth Olson, and Tamra Wysocki, 47–62. Chicago: University of Chicago.
- Goldwater, Sharon and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. *Proceedings of the Workshop on variation within Optimality Theory*, 111–120.
- Gouskova, Maria. 2007. The reduplicative template in Tonkawa. *Phonology* 24: 367–396.

- Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30: 79–133.
- Gouskova, Maria and Tal Linzen. 2015. Morphological conditioning of phonological regularization. *The Linguistic Review* 32(3): 427–473.
- Hayes, Bruce and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonological learning. *Linguistic Inquiry* 39: 379–440.
- Inkelas, Sharon and Cheryll Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically condition phonology. Unpublished manuscript.
- Itō, Junko and Armin Mester. 1999. The phonological lexicon. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 62–100. Oxford: Blackwell.
- Itō, Junko and Armin Mester. 2001. Covert generalizations in Optimality Theory: The role of stratal faithfulness constraints. In *Proceedings of the 2001 International Conference on Phonology and Morphology*, ed. The Phonology-Morphology Circle of Korea, 3–33.
- Itō, Junko and Armin Mester. 2003. *Japanese morphophonemics: Markedness and word structure*. Linguistic Inquiry Monographs no. 41. Cambridge: MIT Press.
- Jurģec, Peter. 2010. Disjunctive lexical stratification. *Linguistic Inquiry* 41: 149–161.
- Jurģec, Peter. 2014. Morphology affects loanword phonology. In *Proceedings of NELS 43*, ed. Hsin-Lun Huang, Ethan Poole, and Amanda Rysling, vol. 1, 191–202. Amherst, MA: GLSA.
- Jurģec, Peter and Bronwyn Bjorkman. 2018. Indexation to stems and words. *Phonology* 35: 577–615.
- Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic Grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the twelfth annual conference of the Cognitive Science Society*, 388–395. Cambridge: Lawrence Erlbaum.
- Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- Mascaró, Joan. 2007. External allomorphy and lexical representation. *Linguistic Inquiry* 38(4): 715–735.
- McCarthy, John. 2005. Taking a free ride in morphophonemic learning. *Catalan Journal of Linguistics* 4: 19–56.
- Merchant, Nazarré. 2008. Discovering underlying forms: Contrast pairs and ranking. Doctoral Dissertation, Rutgers University, New Brunswick.
- Merchant, Nazarré and Bruce Tesar. 2008. Learning underlying forms by searching restricted lexical subspaces. In *Proceedings of the Forty-First Conference of the Chicago Linguistics Society (2005), vol. II: The Panels*, 33–47.
- Moore-Cantwell, Claire and Joe Pater. 2016. Gradient Exceptionality in Maximum Entropy Grammar with Lexically Specific Constraints. *Catalan Journal of Linguistics* 15: 53–66.
- Pater, Joe. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17: 237–274.
- Pater, Joe. 2004. Exceptions in Optimality Theory: Typology and learnability. Presented at the Conference on Redefining Elicitation: Novel Data in Phonological Theory. New York University.
- Pater, Joe. 2005. Learning a stratified grammar. In *Proceedings of the 29th Boston university conference on language development*, ed. Alejna Brugos, Clark-Cotton Manuella, and Seungwan Ha, 482–492. Somerville: Cascadilla Press.
- Pater, Joe. 2006. The Locus of Exceptionality: Morpheme-Specific Phonology as Constraint Indexation. In *University of Massachusetts Occasional Papers in Linguistics*, ed. Steve

- Parker, *Papers in Optimality Theory*, vol. 32. Amherst: University of Massachusetts GLSA.
- Pater, Joe. 2009. Weighted Constraints in Generative Grammar. *Cognitive Science* 33: 999–1035.
- Pater, Joe. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In *Phonological argumentation: essays on evidence and motivation*, 123–154. London: Equinox.
- Prince, Alan. 2002. Arguing Optimality. In *Papers in Optimality Theory II*, ed. Andries Coetzee, Angela Carpenter, and Paul de Lacy. University of Massachusetts, Amherst: GSLA.
- Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.
- Prince, Alan and Bruce Tesar. 2004. Learning phonotactic distributions. In *Fixing priorities: Constraints in phonological acquisition*, ed. Rene Kager, Joe Pater, Wim Zonneveld. Cambridge: Cambridge University Press.
- Smolensky, Paul. 1996. The initial state and ‘richness of the base’ in Optimality Theory. Unpublished manuscript.
- Tesar, Bruce. 1997. Multi-recursive constraint demotion. Unpublished manuscript.
- Tesar, Bruce. 2012. Learning phonological grammars for output-driven maps. *Proceedings of NELS 39*.
- Tesar, Bruce. 2014a. *Output-driven phonology*. Cambridge University Press.
- Tesar, Bruce. 2014b. When worst is best: Grammar construction in phonological learning. In *Proceedings of NELS 43*, vol. 2, 167–180.
- Tesar, Bruce. 2016. Phonological learning with output-driven maps. *Language Acquisition* 24(2): 148–167.
- Tesar, Bruce, John Alderete, Graham Hornwood, Nazarré Merchant, Koichi Nishitani, and Alan Prince. 2003. Surgery in language learning. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, ed. Gina Garding and Mimu Tsujimura, 477–490. Somerville, MA: Cascadilla Press.
- Tesar, Bruce and Alan Prince. 2007. Using phonotactics to learn phonological alternations. In *Proceedings of the thirty-ninth conference of the Chicago Linguistics Society (2003)*, 209–237.
- Tesar, Bruce and Paul Smolensky. 1993. The learnability of Optimality Theory: An algorithm and some basic complexity results. Technical report (CU-CS-678-93).
- Tesar, Bruce and Paul Smolensky. 1995. The learnability of Optimality Theory. In *Proceedings of the 13th West Coast Conference on Formal Linguistics*, 122–137. Stanford, CA: CSLI Publications.
- Tesar, Bruce and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29: 229–268.
- Yang, Charles. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press.