To appear in S. Lima, K. Mullin, B. Smith (eds) Proceedings of NELS39
February 2009draft

USELISTEDERROR: A grammatical account of lexical exceptions in phonological acquisition [*]

Anne-Michelle Tessier

University of Alberta

1. Introduction: error-driven, gradual learning and developmental exceptions

This paper attempts to provide an account for exceptionally-pronounced words in children's developing phonologies, situated within an Optimality-Theoretic, error-driven view of phonological learning. Two kinds of exceptions are discussed here, both of which involve a set of words that are in some way out of synch with the learner's current stage of development. The present proposal provides a way to keep exceptional forms beyond the reach of the 'core' grammar, while still using an independently-proposed approach to gradual OT learning to progress through and beyond exceptional stages.

To begin an error-driven learning account of exceptions: what's in an error? The particular form of errors assumed in this account, and how the learner makes them, is illustrated in (1) below. Learners use their grammar to try reproducing optimal forms heard in the ambient language; when their grammar provides an unfaithful result, they recognize this as an error (made as in 2) and represent it as in 3)*:*

1) observed, target English output: 'toast!' [tost]
   current grammar's output: 'to!' [to]

2) An error tableau:

/tost/	NOCODA	*COMPLEXCODA	*PHARYNGEAL	MAX-C
tost	*!	*!		
☞ to				*

[*] Thanks especially to Joe Pater for access to his version of Trevor's corpus and many helpful discussions and suggestions, and to Shelley Velleman for initially encouraging me to think about variation and exceptions in phonological development. Thanks also to Karen Jesney, Yvan Rose, Michael Becker, John McCarthy, Christopher Cox, and supportive audiences at the University of Western Ontario and NELS39. The full responsibility for all errors and omissions herein, despite all this help, remains my own.

3) Analyzing the error – an ERC row (see Prince 2002)

winner ~ loser	NoCoda	*ComplexCoda	*Pharyngeal	Max-C
tost ~ to	L	L	e	W

Learning a new grammar means building a new ranking of constraints so that for each error row like (3), winner forms like [tost] are preferred to their losers (like [to]). The logic of how re-ranking seeks to achieve this goal is described by the C/D Lemma of Prince and Smolensky (1993: 148), rephrased in Prince and Tesar (2004: 255) as follows (see also Tesar and Smolensky 2000):

4) If *every* L-prefering constraint is ranked below *some* W-prefering constraint, our grammar will prefer the Winner to the Loser.

While the lemma in 4) drives the core of constraint re-ranking, it does not provide a fully-deterministic algorithm for ranking constraints given a set of errors – at least because many different constraint rankings will choose the same optimal input for a given output. The present work follows especially Prince and Tesar (2004) and Hayes (2004) in building rankings with a *Biased* Constraint Demotion algorithm (hereafter BCD). These biases help the learner build the most ‘restrictive’ ranking compatible with a set of ERC rows as per the requirements of 4). For example, the error in (2) and (3) tells us nothing about the role of the markedness constraint *Pharyngeal – but since BCD imposes the bias that Markedness constraints be installed in the highest stratum possible, the grammar learned from 3) via BCD will retain *Pharyngeal at the top of the hierarchy, as in 5) below. (For much more rigorous detail, see Prince and Tesar 2004 and Hayes 2004.)

5) Grammar learned from 3) via BCD:
*Pharyngeal >> Max-C >> NoCoda, *ComplexCoda

Various works have argued that a learner using one-time ERC rows to build new rankings should or must also store these errors for later uses, in a library of errors referred to as the Support. Such stored errors have been used to discover hidden representational structure such as lexical stress (Tesar 1998, 2006; Tesar *et al* 2003) and to learn target grammar exceptions more generally (Pater, to appear), to learn restricted morpho-phonemics (McCarthy 2005), and to avoid learning traps caused by hidden structure (Tessier 2009: 29-34). Perhaps the most extensive use of such stored errors is Error-Selective Learning (ESL: Tessier 2007, 2009), which is also the starting point for the present work. ESL is proposed as a mechanism for gradual and restrictive learning within Optimality Theory. To summarize very briefly, the Error-Selective learner continually accumulates errors with its current ranking in a temporary error Cache, and periodically adds a few cached errors to the permanent Support; only then does the learner use their new Support to build a new ranking (see further explication in section 3 below).[1]

[1] Cf. the different approach to gradual constraint-based learning of the Gradual Learning Algorithm: e.g. Boersma and Hayes 2001, Boersma and Levelt 2003.

Error-Selective learning requires *two* error storage facilities, the Cache and the Support, Is there any independent use or benefit to storing all these errors? This paper argues yes: that stored errors are responsible for (at least some) lexical exceptions in phonological development. The connection between ESL and lexical exceptions is this: A learner that stores ERC rows in the Support has two ways to produce a word it has already heard. The first is to run the input through the current grammar and produce whatever it currently deems optimal; the second is to parrot back the memorized 'loser' form in its stored error. The core proposal of this paper, therefore, is a new constraint, USELISTEDERROR, which prefers the use of memorized stored errors. Within the ESL framework, this constraint can produce stages of both fossilized and precocious errors.

The rest of this short paper proceeds as follows. Section 2 introduces the first kind of exceptional data to be discussed – fossilized forms – and section 3 uses these forms to walk through the mechanics of the proposal, via ESL and USELISTEDERROR. Section 4 provides a second kind of exception, precocious forms, and discusses how they too might be captured. Section 5 raises some of the many open questions, and section 6 concludes.

2. Fossilized forms

From the current perspective, fossilization is a two-stage process that runs as follows. At some early stage, the learner's grammar does not allow a particular target structure (in 6a, codas). At a later stage, the learner has acquired the structure on the whole, but a few words continue to avoid it (6b), looking like fossils of the earlier stage. Schematically:

6)		a) Stage 1: *No Codas*	b) Stage 2: *Codas Preserved except...*	
toast	/tost/→	[to]	[tos]	
dog	/dag/→	[da]	[da]	***fossilized form***
bees	/biz/ →	[bi]	[biz]	

Examples of fossilization are fairly common in the acquisition literature (see especially discussions in work such as Macken 1979; Macken and Ferguson 1983; Menn 1976, 1983.) In Gierut and Dinnsen (1987)'s typology of 'kinds of phonological knowledge', fossilizations are deemed sufficiently common to get a separate entry on the list; three of the six delayed learners they discuss showed fossilizations, targeting /ʃ/, /t/, or /l/.

A concrete example is provided by one child's pronunciation of his own name – Trevor (Compton and Streeter 1977; Pater 1997.) At an early stage (roughly 1;4-1;8) nearly all of Trevor's /tɹ/ clusters are reduced to [t], or occasionally to [tʃ].[2] During this period, only four CC productions for /tɹ/ clusters are recorded ([tɹ] or [tw]):

[2] It is true that target English /tɹ/ clusters are also affricated; we cannot be sure how target-like the amount of affrication in tokens like /tʃi/ is, but it seems safe to assume that the rhotic is not perfectly target-like.

7) a) Usually /tɹ/ → [t] b) Occasionally /tɹ/ → [tʃ]

Child	Target	Age	Child	Target	Age
[ti]	tree	1;4.06	[tʃi]	tree	1;4.23
[teːn]	train	1;5.14	[tʃi]	tree	1;4.27
[tɛwə]	Trevor	1;7.11	[tʃɛːwɝ]	Trevor	1;8.02

By 2;1-2;2, most of Trevor's /tɹ/s are produced as clusters: [tɹ] or [tw], or [kɹ] via consonant harmony. However, this cluster in his name *'Trevor'* is still [t] or [tʃ]:

8) Representative sample of Trevor's stage of fossilization

a) Most words /tɹ/ → [tɹ], [tw] b) Fossilized 'Trevor' /tɹ/ → [t], sometimes [tʃ]

Child	Target	Age	Child	Target	Age
[twiː]	tree	2;1.0	[tæːvɚ iːtʃ eːg]	Trevor eats egg	2;1.0
[tɹʌk]	truck	2;1.05	[tævɚ]	Trevor	2;1.14
[kɹʌk]	truck	2;1.05	[tɛːvɚ]	Trevor	2;1.14
[giːn tɹʌk]	green truck	2;1.26	[tɛwɚ duː ɪt]	Trevor do it	2;1.14
[kaːr tɹʌk']	car, truck	2;1.26	[guːd ʒaːb tɛːvɚ]	good job Trevor	2;1.17
[tu twɹeːnθ]	two trains	2;1.26	[tɛːvɚ fɝːsti]	Trevor thirsty	2;2.23

As a quantitative example: for the month between 2;2 and 2;3, Trevor provides 25 tokens of his name. Of these, 18/25 are produced with an initial [t] or [tʃ], while 7/25 retain the full cluster [tɹ]. In contrast, the same period provides 13 tokens of other /tɹ/-initial words (*train, truck* and *tray*), of which 12/13 are produced faithfully as [tɹ].

It has been reasonably suggested (e.g. Menn 1976) that fossilization occurs when the learner lexicalizes their own pronunciation: in OT terms, when learners perform Lexicon Optimization on their own early outputs rather than on the ambient forms. In the case of a child's own name, this seems particularly appropriate – one's name is arguably whatever one *says* it is. What I present below is one method by which this effect could be achieved: a way in which the learner's earlier output for one lexical item might be maintained at later stage, regardless of whether its meaning is intrinsically tied to the speaker (as with a name) or not. As noted in section 5, more empirical evidence here is clearly needed.

3. Creating fossilized forms via error-selective learning

3.1 The constraint: USELISTEDERROR

The constraint proposed to drive stages of lexical exceptions is given in 9):

9) USELISTEDERROR (ULE) Assign a violation to any output form that is *non-identical to any of its stored loser forms.*

USELISTEDERROR starts at the top of the ranking, and is demoted in the usual way via evidence from ERC rows in the Support. With this addition, there are now more than just two stages by which a fully-unmarked grammar becomes fully-faithful: more than just M >> F and F >> M, there is also the position of ULE to consider. As the next section illustrates, a new intermediate stage created by USELISTEDERROR can derive error fossilization.

3.2 The pre-fossilization stage

We begin with the initial ranking as dictated by the biases of a modified BCD algorithm: modified to include a bias for USELISTEDERROR above all other constraints (10a). This example will use the NOCODA case from 3), so that the relevant initial state is (10b):

10)a) *Modified BCD biases* USELISTEDERROR >> MARKEDNESS >> FAITH
b) *Resulting fragment of initial state:* USELISTEDERROR >> NOCODA >> MAX

This initial state grammar will of course make errors that delete target codas. The first time that the learner attempts to produce the word 'dog', for example, it will make an error as in 11) below. Since this is the first time such an error has been made, there is no listed form anywhere, so USELISTEDERROR is vacuously satisfied:

11) First time processing 'dog': *no effect of USELISTED*

/dag/ **no stored loser**	USELISTEDERROR	NOCODA	MAX
a) dag		*!	
☞ b) da			*

This error gets added to the Cache, and similar errors pile up on new words – until the next time the learner attempts 'dog'.This time the Cache *does* have a stored loser for this word and this results in a slightly different error profile, as in 12):

12) Second time processing 'dog: *USELISTED violated*

/dag/ **stored loser**: [da]	USELISTEDERROR	NOCODA	MAX
a) dag	***!**	*!	
☞ b) da			*

Though the learner's grammar remains at the initial stage, the learner now has a Cache with at least these two stored errors:

13) *A fragment of the Cache, near the end of Stage One ...*

winner ~ loser	USELISTEDERROR	NOCODA	MAX
a) dag ~ da	**e**	**L**	W
b) dag ~ da **stored loser: da**	**L**	**L**	W

3.3 Creating fossilized forms via USELISTEDERROR

The error-selective learner that is piling up errors in a Cache like (13) must eventually move on; in ESL, learning occurs when a particular error is chosen from the Cache and added to the Support, so that it can be used to build a new ranking via its BCD algorithm. As already alluded to, the Error-Selective learner chooses an error that will make 'minimal changes' to the current grammar. To choose its error wisely, the ESL uses an error-selection algorithm, which takes into account all of the information in the Cache's ERC rows – comparing Ws and Ls, comparing violations of markedness vs. faithfulness – and picks one of the errors that will minimally change the current grammar. (This algorithm is not spelled out here for reasons of space – see Tessier 2007 ch 3, 2009.) Once an error has been chosen, the Cache is emptied, and the cycle begins.

The important upshot is that at early stages, when many different errors are possible, the error-selection algorithm might well choose *either* of the errors from 13), among any of the other Cached options. The interesting result will come about if the error chosen from the Cache is (13a). Suppose that this error is added to the Support, and this updated Support is fed to the BCD algorithm to produce a new ranking.

14) *The error in 13a), added to the Support*

winner ~ loser	USELISTEDERROR	NOCODA	MAX
dag ~ da **(stored loser)**	**e**	**L**	**W**

What will our Biased Constraint Demotion do, based on this single error? Recall the biases from 10), which tell us that in the absence of errors, the ranking must be USELISTED >> Markedness >> Faith. The error in 14) shows that in this case, one M >> F ranking must be reversed, preserving codas in the target grammar. But since the Support's ERC rows give no evidence as to the ranking of USELISTED, it remains where the biases want it, at the top of the ranking. Thus:

15) *The Stage Two grammar, built off the Support in 12)*
USELISTEDERROR >> MAX >> NOCODA

This ranking in 15) has now fossilized the error in 14). When this grammar processes any other word, its coda will be preserved by MAX:

16) *MAX >> NOCODA usually preserves codas...*

/kaet/ **no stored loser**	USELISTEDERROR	MAX	NOCODA
☞ a) kaet			*
b) kae		*!	

But for 'dog', things are different. Stage 2 was brought about by adding 14) to the Support, while all the other traces of the old Cache have been emptied and forgotten. This stored error caused the learner to acquire faithful codas in most word, like 'cat', but it is now also responsible for the *failure* to produce a coda in ‘dog’. Since ‘dog’ has a stored loser form, and since USELISTEDERROR is undominated, the grammar must use it:

17) *USELISTEDERROR >> MAX retains the old error*

/dag/ **stored loser [da]**	USELISTEDERROR	MAX	NOCODA
a) dag	*!		*
☞ b) da		*	

To summarize: if the learner happens to acquire faithful codas via an error that does *not* contain a ULE violation, then that particular form will be fossilized without its coda at the next stage.

3.4 Moving beyond fossilization

In a sense, the grammar in the previous section has gotten halfway to the target grammar – it has demoted NOCODA beneath MAX, but it has not yet demoted ULE as well. To reach the target, an ERC row for the fossilized word ‘dog’ needs to get into the Support. Such errors will certainly be recorded in the Cache – because at this stage ‘dog’ is still being reduced, creating errors like 18):[3]

18) *ERC row that must be added to the Support to reach the target grammar*

winner ~ loser	USELISTEDERROR	NOCODA	MAX
dag ~ **da (stored loser)**	**L**	L	W

Once this ERC row is added to the Support, the learner’s next ranking will be one in which *both* constraints rank beneath MAX. This grammar fragment is the target, in which codas are preserved regardless of stored forms:

19) *Ranking built from 18)*: MAX >> USELISTEDERROR >> NOCODA

20) The end of fossilization

/dag/ **stored loser**: [da]	MAX	USELISTEDERROR	NOCODA
☞ a) dag		*	*
b) da	*!		

We have now seen how adding ULE to the error-selective BCD learner can result in three stages in the acquisition of a marked structure – and that the new intermediate stage is one that fossilizes old pronunciations of that marked structure. To summarize:

[3] Some technical work, however, is necessary to ensure that ERC rows which overcome fossilization get into the *Support*. Given the details of ESL in Tessier (2007), this will probably require one of the following two tweaks: getting the Violation Threshold down to one (which seems unlikely), or treating ERC rows in which USELISTED assigns a L as special (which might be best.) Further spelling-out is required.

21) Stage 1: USELISTEDERROR >> NOCODA >> MAX
Stage 2: USELISTEDERROR >> MAX >> NOCODA *fossilization stage*
Stage 3: MAX >> USELISTEDERROR >> NOCODA[4]

4. Precocious forms

4.1 An example of precocious forms, and their possible cause

Precocious forms (as I will call them here) are the opposite of fossilized ones: words that faithfully produce a target structure which is otherwise avoided in the child's grammar via some process. The process to be discussed here is Velar Fronting (VF), whereby velar segments are fronted in some contexts to coronal (alveolar) place. VF appears to be a quite common process among English-learning children (see e.g. Chiat 1983, Brett, Chiat & Pilcher 1987, Dinnsen 2002, Ingram 1974, Inkelas and Rose 2007, Stoel-Gammon 1996.)

22) Velar Fronting in E's grammar (Inkelas and Rose 2007)

'kiss'	/kɪs/	→	[tɪs]
'again'	/əgɪn/	→	[ədɪn]

The relevant example of precocious forms comes from K, a child discussed in Bleile and Tomblin (1991). For the first six recording sessions of this study, K's grammar generally applied VF to word-initial or stressed velars (e.g. 23a), but also allowed a small set of relevant words to retain their velars (23b):

23) Precocious Forms in K's grammar (from Bleile and Tomblin 1991)

a) Regular VF pattern:

	Target	*Child*
'candy'	['kaendi]	['taendi]
	(inferred example)	

b) Precocious forms without VFs:[5]

	Target	*Child*
'clown'	[klaʊn]	[kaʊn]
'cookie'	['kʊki]	['kʊki]
'okay'	[o'ke]	[o'ke]

While velar place is marked compared to coronals, Inkelas and Rose (2008) argue that the phonological contexts for children's VF do not mirror attested velar-to-coronal processes in adult grammars. As a result, Inkelas and Rose argue that child VF is driven by child-specific articulatory pressures. In what follows, I will assume that these pressures cause the child learner to construct a child-specific markedness constraint and add it to their grammar. (For a somewhat similar assumption, see Pater 1997; for a discussion of constructing constraints from experience of articulatory difficulty, see Hayes 1999.). This constraint must be a positionally-sensitive ban on velars, relativized

[4] Note that if the learner had added the other error to the Support at Stage 1, they would have not fossilized any errors and instead acquired all codas at the same time, moving straight to Stage 3.
[5] The full set of precocious words reported: *clown, okay, kitty cat, cow, Cookie Monster, (ice) cream, Gumbi*. Note that this study did not report any influence of vowel quality on the pattern.

according to Inkelas and Rose (2008) to stressed and word-initial onsets. To save space, I will simply refer to this constraint in tableaus as *K.

4.2 An approach to precocious forms via USELISTEDERROR

The idea pursued here is that precocious forms are able to escape the current grammar's unfaithful mappings like velar fronting precisely because the markedness constraint driving VF is child-constructed. Adding errors to the Support *before* constructing the velar fronting constraint can create precocious forms later on.

The first relevant stage of learning is one at which no *K constraint has been constructed. Here, ERC rows like the one in (24) below can be added to the Support, without VF:

24) Fragment of the Support at an early stage (USELISTEDERROR not yet relevant)

winner ~ loser	*COMPLEXONSET	MAX	IDENT-PLACE
klaʊn ~ **kaʊn (stored loser)**	L	W	e

At some later point, the child is overcome by increasing articulatory demands (perhaps under the pressure of a growing lexicon), and so decides (by admittedly unspecified means) to add a *K constraint to their grammar. A necessary assumption is that learners use their ranking biases to add any newly-minted constraints to the ranking in its initial state position. In this case, a new markedness constraint like *K must be ranked above IO faithfulness constraints but *below* ULE:

25) *Stage 2:* USELISTEDERROR >> *K >> IDENT-PLACE

This grammar in 25) is now one that generally applies velar fronting. In particular, any new words that do *not* have stored error forms in the Support will undergo fronting:

26) The new *K constraint causes velar fronting

/kændi/	USELISTEDERROR	*K	IDENT-PLACE
a) kændi	(*)	*!	
☞ b) tændi			*

However, the fact that *K is ranked beneath ULE means that there is a possibility for precocious forms to escape velar fronting. If any errors with VF targets are already stored in the Support – like 'clown' in 24) – the undominated ULE constraint will force the production of their stored loser form and thereby block VF:

27) Errors in the Support created before *K do not undergo VF

/klaʊn/ **stored loser:** [kaʊn]	USELISTEDERROR	*K	IDENT-PLACE
a) klaʊn	*!	*	
☞ b) kaʊn		*	
c) taʊn	*!		*

In fact, this treatment of precociousness also encompasses some fossilization too. To retain the velar in a precocious word like ‘clown’, it must have been added to the Support and so learned from, meaning that e.g. the current grammar normally retains onset clusters like [kl]. On the other hand, these precocious words must be pronounced as their stored error form – meaning that with respect to onset clusters, ‘clown’ as [kaʊn] also shows a degree of fossilization.

4.3 Moving beyond precocious stages

Compared to fossilization, it appears that the treatments of precocious forms at later stages may be more complicated. For the child K discussed above, the next stage was one in which velar fronting was *uniformly* applied, even in the previously-precocious forms, as presented below:

28) Rates of Velar Fronting among Precocious Words

	Sessions	*# VF tokens*	*# Unfronted tokens*	
Stage 2	1-6	3	**16**	*precocious stage* (§4.2)
Stage 3	7-22	**58**	3	*uniform VF*

To complete the learning trajectory: a post-test, administered two weeks after the final session 22, revealed no further VF in any words including ‘clown’. This suggests that the child’s Stage 4 was on the way to mastery of the target grammar.

The question is how this progression to Stages 3 and 4 might have come about. To derive uniform velar fronting in Stage 3, *K must have come to be ranked *above* ULE, so that even old errors with stored forms would undergo fronting:

29) Stage 3: VF across the board, no more precocious forms

/klaʊn/ **stored:** [kaʊn]	*K	USELISTEDERROR	IDENT-PLACE
a) klaʊn	*!	*	
b) kaʊn	*!		
☞ c) taʊn		*	*

Once Stage 3 has been reached, the learner at least has a straightforward route to the final grammar. From errors based on unfaithful mappings like 29) (and others), the learner will eventually add evidence to the Support that will demote both *K and USELISTED beneath faith constraints such as IDENT-PLACE. This will be the final stage, in which velars are faithfully preserved regardless of position or the stored errors of earlier stages.

30) Stage 4: no velar fronting, no precocious forms

/klaʊn/ **stored:** [kaʊn]	IDENT-PLACE	USELISTEDERROR	*K
☞ a) klaʊn		*!	*
b) taun	*!		

The biggest question raised by this progression is how the learner might get from stages 2 to 3. Comparing the rankings in 25) and 29), it must be the case that USELISTEDERROR and *K have changed places – and it is not clear how any ERC rows could provide the evidence to prompt such a ranking reversal.

The only answer to be offered here, of a purely speculative nature, is that because *K is a specially-created constraint, its specialness also carries over to its treatment by the BCD algorithm. Just as a learner can invent this constraint under articulatory pressure, perhaps they can also decide that it is important enough to be promoted above other constraints *despite* the Support's evidence to the contrary. One empirical question is how common this child's developmental sequence is – that is, whether it is typical for precocious forms to gradually give in (or 'regress') to a child-specific process before it is eventually overcome. Either way, this issue remains formally unresolved.

5. Discussion and Open Questions

The previous three sections have demonstrated, with two toy examples, how an Error-Selective learner might acquire phonological grammars which allow developmental lexical exceptions. In the first case, USELISTEDERROR can retain an old pronunciation, producing a fossilized form. In the second case, when a new constraint has been added to the grammar, USELISTEDERROR can in fact license old structure which has only recently been deemed marked, and so produce a precocious form.

One general question about this approach is whether USELISTEDERROR is really the monolithic constraint used in the tableaus above, or whether it is parameterized in some way(s). In the case of fossilization: surely multiple old processes, such as coda deletion, might be fossilized by a single learner, and surely not every one is always overcome at the same time and speed. It might be that ULE is in fact a family of constraints, each relativized to either an unfaithful mapping or a markedness pressure, which can be independently demoted by adding ERC rows to the grammar.

Another big question is the nature of child-specific constraints, such as the one assumed here to be driving velar fronting. How are such constraints constructed? On what grounds and from what primitives? These are important questions that cannot be properly addressed here. It will merely be noted that if patterns like velar fronting are not potential members of target adult grammars then we must allow for some child-specific method of incorporating them into phonological grammars.

The poster child for child-specific processes is clearly Consonant Harmony: e.g. Goad (1997); Pater and Werle (2001); also Fikkert and Levelt (2008). While adult languages exhibit long-distance consonant harmonies of many kinds (see recently Hansson 2001, Rose and Walker 2004), it has often been noted that they do *not* impose long-distance harmony among major places of articulation, as children such as Trevor very often do:

31) Examples of Trevor's Consonant Harmony (from Pater and Werle 2001)

'dog'	/dɑg/	→	[gɑg]	'cup'	/kʌp/	→	[kʌk]
'coat'	/kot/	→	[kok]	'cat'	/kæt/	→	[kæk]

As such, consonant harmony seems the best place to look for empirical confirmation of the claims made about precocious forms in section 4. If consonant harmony is the result of a child-constructed constraint, it should frequently show precocious exceptions. As a first step, Moskovitz (1980) reports a fairly common anecdotal pattern: a child whose earliest pronunciations of ‘truck’ did not undergo harmony, but which later harmonized, to [**k**ak]. The prediction here is that often some early words will *not* harmonize, having been stored in the Support before harmony was instantiated in the grammar.

A different issue for child-specific constraints is what precisely it means to add a constraint in its ‘initial state position’, to a grammar that has already moved away from its purely-biased rankings on the basis of the Support. How should *K be ranked with respect to the full panoply of faithfulness constraints, interleaved as they will be with the other markedness constraints moving down the hierarchy via errors already made? This too remains a question for further research.

6. Conclusions

The primary goal of this paper has been to argue that at least some lexical exceptions in phonological development can be derived as part of normal constraint-based grammatical learning. A secondary goal has been to suggest a new use of stored errors, as extensively required by Error-Selective Learning, beyond the original issues of gradualness and restrictiveness in learning. This approach to lexical exceptions in phonological development is certainly preliminary in nature. Systematic corpus data of the nature, frequency and typology of exceptional words during acquisition is somewhat sparse, and both in-depth case studies and cross-linguistic comparisons should be used to delve deeper into the empirical basis of this research. Furthermore, a computational implementation of error-selective learning and its use of USELISTEDERROR to create fossilized and precocious forms will be necessary to best probe the extent to which this approach ‘scales up’. But building one formal view of these phenomena may hopefully provide a clearer lens with which to examine the range, scope, and implications of exceptional words in phonological acquisition.

References

Bernhardt, Barbara. H., and Joseph P. Stemberger. 1998. *Handbook of Phonological Development: From the Perspective of constraint-based nonlinear phonology*, San Diego, Academic Press.

Bleile, Ken and J. B. Tomblin. 1991. Regressions in the Phonological Development of Two Children. *Journal of Psycholinguistic Research 20* 6: 483- 499.

Boersma, Paul and Bruce Hayes. 2001. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry 32*: 45-86.

Boersma, Paul and Claartje Levelt. 2003. Optimality Theory and phonological acquisition. *Annual Review of Language Acquisition* 3: 1-50.

Brett, L., S. Chiat & C. Pilcher. 1987. Stages and Units in Output Processing: Some Evidence from Voicing and Fronting Processes in Children. *Language and Cognitive Processes* 2:165-177.

Chiat, S. 1983. Why Mikey's Right and My Key's Wrong: The Significance of Stress and Word Boundaries in a Child's Output System. *Cognition* 14:275-300.

Compton, A.J. and M. Streeter. 1977. 'Child Phonology: Data Collection and Preliminary Analyses'. In *Papers and Reports on Child Language Development 7*. Stanford University, Stanford, California.

Dinnsen, Daniel. 2002. A reconsideration of children's phonological representations. In B. Skarabela, S. Fish, and A. H.-J. Do (eds.) *Proceedings of the Boston University Conference on Language Development* 26, Somerville: Cascadilla: 1-23.

Fikkert, Paula & Clara C. Levelt. 2008. 'How does place fall into place? The lexicon and emergent constraints in the developing phonological grammar'. In P. Avery, B. Elan Dresher & K. Rice (eds.), *Contrast in phonology: Perception and Acquisition.* Berlin: Mouton.

Goad, Heather. 1997. Consonant harmony in child language: an optimality-theoretic account. In S. J. Hannahs and M. Young-Sholten (eds.) *Focus on phonological acquisition*. Amsterdam: John Benjamins. 113-142.

Gierut, Judith and Daniel Dinnsen. 1987. On predicting ease of phonological learning. *Applied Linguistics* 8: 241-263.

Hansson, Gunnar. 2001. *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. dissertation, University of California, Berkeley.

Hayes, Bruce. 1999. Phonetically-Driven Phonology: The Role of Optimality Theory and Inductive Grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer, and K. Wheatly (eds)., *Functionalism and Formalism in Linguistics, Volume I: General Papers.* John Benjamins, Amsterdam: 243-285.

Hayes, Bruce. 2004. Phonological Acquisition in Optimality Theory: the early stages. In R. Kager, J. Pater & W. Zonneveld (eds.), *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: CUP.

Ingram, David. 1974. Fronting in child phonology. *Journal of Child Language*1:233-241.

Inkelas, Sharon and Yvan Rose. 2008. Positional Neutralization: A Case study from Child Language. *Language* 83: 707-736.

Kiparsky, Paul and Lise Menn. 1977. On the acquisition of phonology. In John Macnamara (ed.), *Language Learning and Thought.* New York: Academic Press: 47-78.

Macken, Marlys. 1980. The child's lexical representation: The 'puzzle-puddle-pickle' evidence. *Journal of Linguistics* 16: 1-17.

Macken, Marlys and Charles Ferguson. 1983. Cognitive aspects of phonological development: Model, evidence and issues. In K. Nelson (ed.) *Children's Language.* vol 3. Hillsdale, NJ: Erlbaum.

McCarthy, John J. 2005. Taking a Free Ride in morphophonemic learning. *Catalan Journal of Linguistics* 4: 19-56.

Menn, Lise. 1971. Phonotactic rules in beginning speech: a study in the development of English discourse. *Lingua* 26: 225-251.

Menn, Lise. 1976. *Pattern, control, and contrast in beginning speech: A case study in the development of word form and word function.* Ph.D. dissertation, UI Urbana.

Menn, Lise. 1983. ‘Development of articulatory, phonetic, and phonological categories. In B. Butterworth (ed.) *Language production.* New York: Academic Press.
Moskovitz, B. A. 1980. Idioms in phonology acquisition and phonological change. *Journal of Phonetics* 8: 69-83.
Pater, Joe. 1997. ’Minimal violation and ‘Minimal Violation and Phonological Development’. *Language Acquisition* 6 (1).
Pater, Joe. To appear. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Steve Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation.* London: Equinox.
Pater, Joe and Adam Werle. 2001. Typology and variation in child consonant harmony. In C. Féry, A. D. Green and R. van de Vijver (eds.), *Proceedings of HILP 5*, University of Potsdam.
Prince, Alan. 2002. Entailed Ranking Arguments. Ms, Rutgers University. ROA-500.
Prince, Alan and Bruce Tesar. 2004. Learning Phonotactic Distributions. In R. Kager, J. Pater & W. Zonneveld,(eds.), *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: CUP.
Rose, Sharon and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80, 475-531.
Stoel-Gammon, Carol. 1996. On the Acquisition of Velars in English. In B. Bernhardt et al. (eds.) *Proceedings of the UBC International Conference on Phonological Acquisition*. Somerville, MA: Cascadilla, 201-214.
Tesar, Bruce. 2006. Learning from Paradigmatic Information. *Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince*. Paper 14. Linguistics Research Center. Available from http://repositories.cdlib.org/lrc/prince/14
Tesar, Bruce., John Alderete, Graham Horwood, Nazarré Merchant, Koichi Nishitani, and Alan Prince. 2003. Surgery in language learning. In G. Garding and M. Tsujimura (eds.) *Proceedings of the Twenty- Second West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla, 477-490.
Tesar, Bruce and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.
Tessier, Anne-Michelle. 2007. *Biases and Stages in Phonological Acquisition.* Ph.D. dissertation, University of Massachusetts Amherst. ROA #883-1106.
Tessier, Anne-Michelle. 2009. Frequency of violation and constraint-based phonological learning. *Lingua* 119 (1): 6-38.
Zuraw, Kie. 2000. *Patterned Exceptions in Phonology.* Ph.D. dissertation, UCLA. Available at [http://www.linguistics.ucla.edu/people/zuraw/dnldpprs/diss.pdf]

Anne-Michelle Tessier
Department of Linguistics
4-22 Assiniboia Hall
University of Alberta
Edmonton, AB T6G 2E7
CANADA

amtessier@ualberta.ca