Intermediate Markedness

Erin Olson*
MIT, ekolson@mit.edu

November 13, 2016

1 Introduction

Within the phonological acquisition literature, it has been observed that there are some children who acquire marked configurations within the target language in a two-step fashion. First, the marked configuration is acquired in some privileged position or positions within the word, such as in stressed syllables (Echols and Newport, 1992; Fikkert, 1994; Kehoe, 2000; Rose, 2000) or at word edges (Echols and Newport, 1992; Fikkert, 1994; Kehoe, 2000; Pater, 1997). Next, the marked configuration is acquired in all of the positions where it appears in the target language. In some cases (such as in English, (Kehoe, 2000)), the child may acquire a marked configuration in two privileged positions in sequence – for instance, they may acquire the marked configuration in stressed syllables first, and then acquire the marked configuration in both stressed syllables and at word edges, while finally arriving at the target language in a third stage.

A simple but instructive example of an intermediate stage comes from Rose (2000), who observed the following pattern in the acquisition of complex onsets for two children learning Québec French, Clara and Théo. Initially, neither child had complex onsets in any of their productions, as shown in (1).

(1) **Initial stage for the acquisition of complex onsets in Québec French** (Rose, 2000, p.132):

a. Clara, age 1;00.28 to age 1;09.01:

Target	*Child*	*Gloss*
/bʁiˈze/	[bœːˈçiː]	'broken'
/ˈflœʁ/	[βœː]	'flower'
/abʁiˈko/	[pupæˈko]	'apricot'

b. Théo, age 1;10.27 to age 2;05.29:

Target	*Child*	*Gloss*
/ˈklʊn/	[ˈkuɲ]	'clown'
/ˈtχɛ̃/	[ˈkɛ]	'train'
/bʁiˈze/	[piˈz̥e]	'broken'

*I am thankful to Edward Flemming, Bruce Hayes, Junko Ito, Elliott Moreton, Norvin Richards, Donca Steriade, the students in the 2015 fall semester of MIT's course 24.991, and the attendees of MIT's Phonology Circle, PhoNE 2016 @ NYU, NECPhon 2016 @ UMass Amherst, and AMP 2016 @ USC for helpful feedback. Special thanks to Adam Albright, Naomi Feldman, and Giorgio Magri for extensive feedback, comments, and suggestions. All remaining errors are my own.

In the following stage, both Clara and Théo have acquired complex onsets, but only in stressed syllables, as shown in (2).

(2) **Intermediate stage for the acquisition of complex onsets in Québec French** (Rose, 2000, pp.133-4):

a. Clara, age 1;09.29 to age 2;03.05:

Target	*Child*	*Gloss*
/biˈbʁɔ̃/	[paˈpχɔ]	'baby bottle'
/ˈglɪs/	[ˈklɪs]	'(s/he) slides'
/siˈtχʊj/	[θəˈtχuːj]	'pumpkin'
/ˈplœʁ/	[ˈplœχ]	'(s/he) cries'
/fʁiˈgo/	[buˈko]	'fridge'
/gliˈsad/	[kaˈsæd]	'(a) slide'
/tχuˈve/	[tʊˈve]	'found'

b. Théo, age 2;05.29 to age 2;11.29:

Target	*Child*	*Gloss*
/ˈgʁo/	[ˈgʁo]	'big'
/ˈtʁɛ̃/	[ˈkχɛ]	'train'
/ˈkle/	[ˈkxi]	'key'
/ˈplœʁ/	[ˈplœʊ̯]	'(s/he) cries'
/tχakˈtœʁ/	[taˈtœʊ̯]	'tractor'
/gʁyˈjo/	[kʰœˈjɔ]	'oatmeal'
/tχuˈve/	[kʊˈβi]	'found'

In the final stage, both children are adult-like in their productions of complex onsets. They are able to produce complex onsets in both stressed and unstressed syllables, as shown by the data in (3).

(3) **Final stage for the acquisition of complex onsets in Québec French** (Rose, 2000, p.135):

a. Clara, age 2;03.15:

Target	*Child*	*Gloss*
/ˈgʁo/	[ˈgʁo]	'big'
/tχuˈve/	[tχuˈve]	'found'
/plɑ̃ˈʃe/	[plɑ̃ˈʃe]	'floor'

b. Théo, age 3;00.07:

Target	*Child*	*Gloss*
/tχuˈve/	[kχaˈve]	'found'
/pχəˈne/	[pʁəˈne]	'(you pl.) take'
/plœˈʁe/	[pløˈʁe]	'to cry'
/gliˈsad/	[kiˈsad]	'slide'

What is of interest for these stages (hereafter generally referred to as **intermediate stages**) is that children who go through them *do not receive evidence from the target language that these positions should be privileged.* Since the target language has the marked structure in all possible positions, there will be no concrete evidence that it should be preferred in one or another. Rather, the existence of these stages must arise from some aspect of the child's phonological system.

One way of characterizing this child-specific phonological system is in the form of a specific ranking of constraints within Optimality Theory (Prince and Smolensky, 1993), one that is different from the adult language. For instance, in adult Québec French, a constraint disallowing complex onsets, *COMPLEX, must be dominated by a constraint prohibiting the deletion of segments, MAX. However, for the intermediate stage child grammar, it is often said to be the case that only a constraint prohibiting deletion of material from a stressed syllable, MAX/$\acute{\sigma}$, can outrank *COMPLEX – the more general MAX must be lower-ranked (Rose, 2000; Goad and Rose, 2004; Jesney and Tessier, 2007, 2008; Tessier, 2009). Under this view of modelling child grammar, the three stages that Clara and Théo go through can be characterized as shown in (4).

(4) **Learning path for Clara & Théo, using positional FAITHFULNESS:**

Initial stage grammar: *COMPLEX ≫ MAX/σ́, MAX

Intermediate stage grammar: MAX/σ́ ≫ *COMPLEX ≫ MAX

Final stage grammar: MAX/σ́, MAX ≫ *COMPLEX

Of particular note in the above analysis is the intermediate stage constraint ranking, which ranks the positional FAITHFULNESS constraint MAX/σ́ above the MARKEDNESS constraint *COMPLEX, which is in turn ranked above the more general FAITHFULNESS constraint MAX. According to Tessier (2009), *any* intermediate stage child grammar will follow this pattern, no matter what the particular FAITHFULNESS constraint is. Tessier's (2009) schema for an intermediate stage ranking is given in (5). Tessier refers to this stage as an **intermediate faithfulness** stage, and I will follow her in referring to intermediate stages as such when they are characterized in this way.

(5) **Ranking schema for intermediate faithfulness stages** (Tessier, 2009, p.7):
Positional FAITHFULNESS ≫ MARKEDNESS ≫ General FAITHFULNESS

While intermediate stages have in general been characterized as being due to high-ranking positional FAITHFULNESS constraints, Jesney and Tessier (2007, 2008) and Tessier (2009) have observed that modelling the learning path in (4) using standard assumptions about phonological learning with a gradual OT learner, such as Boersma's (1997) Gradual Learning Algorithm (GLA) or Magri's (2012) convergent variant (CGLA), proves to be impossible. While the target language is acquired successfully in all cases, these learning algorithms do not predict that children should *ever* go through an intermediate stage. Rather, the prediction is that their productions should go directly from disallowing the marked structure everywhere (as in the initial stage), to allowing the marked structure everywhere (as in the final stage). An overview of how gradual OT learners function (in particular, the CGLA) is provided in section 2, and a review of Jesney and Tessier's (2007, 2008) demonstration of their failure to predict intermediate faithfulness stages will be given in section 3.

This failure to predict an intermediate stage under standard learning assumptions has been used to argue that the GLA and other learners like it should be abandoned in favour of using other learning algorithms. Jesney and Tessier (2007, 2008) argue that a gradual learning algorithm based on Harmonic Grammar (Legendre et al., 1990; Smolensky and Legendre, 2006) should be used over an OT-based gradual learner, on the grounds that the intermediate stage can be characterized as a gang-up effect. While the general MARKEDNESS constraint (*COMPLEX) will outweigh both of the FAITHFULNESS constraints (MAX/σ́ and MAX) when considered individually, the combined weight of the FAITHFULNESS constraints will be greater than that of the MARKEDNESS constraint. Tessier (2009) also advocates for adapting a non-gradual learner, Hayes's (2004) Low Faithfulness Constraint Demotion algorithm (LFCD), into a gradual learner. Tessier opts for this learner since it is already biased towards ranking positional FAITHFULNESS constraints higher than their general counterparts. This learner does so by preferring to rank more "specific" constraints higher than more "general" constraints.

The aim of this paper is to provide yet another solution to the puzzle of the (C)GLA's failure to predict intermediate stages when using standard assumptions about phonological learning. I will show that intermediate stages *can* be predicted by gradual OT learners when they can be characterized by ranking positional MARKEDNESS constraints high in the grammar, as in (6). When characterized in this way, I will refer to the intermediate stages as **intermediate markedness** stages. Note that there is *no difference in data* between these two stages – rather, the only difference is in the particular set of OT constraints and their ranking that is used to differentiate them.

(6) **Ranking schema for intermediate markedness stages**:
Positional MARKEDNESS ≫ FAITHFULNESS ≫ General MARKEDNESS

It will be shown in section 4 that an analysis of Québec French that makes use of a positional MARKEDNESS constraint banning complex onsets in unstressed syllables, *COMPLEX/$\breve{\sigma}$, will correctly follow the attested order of complex onset acquisition. Under such an analysis, the learning path for Clara and Théo will be characterized as in (7).

(7) **Learning path for Clara & Théo, using positional MARKEDNESS**:

Initial stage grammar:	*COMPLEX/$\breve{\sigma}$, *COMPLEX ≫ MAX
Intermediate stage grammar:	*COMPLEX/$\breve{\sigma}$ ≫ MAX ≫ *COMPLEX
Final stage grammar:	MAX ≫ *COMPLEX/$\breve{\sigma}$, *COMPLEX

I will also show that when a particular positional FAITHFULNESS constraint cannot be supplanted by a positional MARKEDNESS variant, then it may still be the case that it can remain low-ranked in the hierarchy and still correctly predict children's intermediate stage grammars, as well as their acquisition order. This will be explored in more detail when discussing data from children acquiring unstressed syllables in English (Kehoe and Stoel-Gammon, 1997; Kehoe, 2000) in section 5.

Since, under gradual OT learners, the use of positional FAITHFULNESS constraints versus positional MARKEDNESS constraints has significant consequences for what paths of phonological acquisition are attested, it might be expected that adult grammars should also be affected by this choice of constraints. This is discussed further in section 6, along with alternate solutions to the issue identified by Jesney and Tessier (2007); **?** and Tessier (2009), and cases under which an HG-based learner might make different predictions. Section 7 concludes.

2 An overview of gradual OT learners

Gradual OT learners, such as the GLA (Boersma, 1997) and CGLA (Magri, 2012), differ from other kinds of OT learners in that they allow phonological learning to happen on an item-by-item basis, rather than by making inferences over a large body of stored data. In order to achieve this, OT constraint rankings are stored as numerical **ranking vectors**, rather than as simple ranking

statements (Boersma, 1997). A ranking vector can be generated from a ranking statement by assigning each constraint a numerical value, such that constraints that are highly ranked receive higher values than constraints that are ranked lower. It will thus be the case that many ranking vectors can be associated with the same ranking statement. An illustration is given in (8).

(8) **Ranking vectors associated with the ranking statement** MAX/*σ́* ≫ *COMPLEX ≫ MAX:

	MAX/*σ́*	*COMPLEX	MAX	
(	100,	99,	98	)
(	30,	5,	1	)
(	96,	62,	7	)
(		…		)

Under standard learning assumptions, all of the MARKEDNESS constraints begin ranked above all of the FAITHFULNESS constraints (Smolensky, 1996; Gnanadesikan, 2004). In other OT learning algorithms, this is usually achieved by positing that all MARKEDNESS constraints begin in the top constraint stratum, and all of the FAITHFULNESS constraints begin in the next stratum down (Tesar and Smolensky, 2000, *a.o.*). An alternate approach involves maintaining a persistent MARKEDNESS ≫ FAITHFULNESS bias throughout the learning process, such as in the learning algorithms devised by Hayes (2004) and Prince and Tesar (2004). For the gradual learners, this initial bias towards MARKEDNESS constraints is encoded by setting all of the MARKEDNESS constraints' initial ranking value to some high value, while having all of the FAITHFULNESS constraints' initial value set to 0 (Boersma, 1997; Magri, 2012).

From this initial point, learning is modelled in a gradual OT learner by adjusting the numerical values associated with each constraint slightly each time the child posits an error.[1] The constraints whose numerical values are to be altered are selected according to the following criteria and collected into an **update vector**. If a constraint **prefers the adult form** of the word, it will be **promoted** by some amount, which I will designate with the letter w. If that constraint **prefers the child's error** and **is not dominated** by some adult form-preferring constraint, then it will be **demoted** by some amount, which I will designate with the letter l. If the constraint prefers **neither** form, or if it prefers the error and is appropriately dominated, then its value will not be adjusted. Preference is calculated according to Prince's (2002) comparative tableaux conventions.

Hereafter, tableaux corresponding to children's posited errors will be given in the following format. The child's error will be selected as the **optimal candidate** in the tableau. Using Prince's (2002) conventions for constructing comparative tableaux, the error will be compared to the input, which represents the adult's form, under the assumption that the error is the **losing** candidate. The result of this comparison will be displayed in parentheses to the right of any violations. A schematic example is given in (9).

[1] For the remainder of this paper, I will make the assumption that the child's surface form and the error that they have posited are identical. However, there is substantial evidence that this is not necessarily the case, and that children's perception – and therefore their grammar – is much more advanced than what they are able to produce (Hayes, 2004; Fikkert, 2007; Fikkert and de Hoop, 2009, and references therein). I unfortunately do not have much more to say about how to resolve the matter. It is the hope that the CGLA is still able to function as outlined here, regardless of what evidence the child is using to posit errors.

(9) **Sample error tableau**:

/ɪmpʊt/	C_1	C_2	C_3
☞ a. ˈɛɹəɹ	(L)	∗∗ (W)	∗(*e*)
b. ˈtɑɹgət	∗!∗	∗	∗

Under the original GLA (Boersma, 1997), the promotion and demotion values, w and l, were held to be equivalent, and assumed to be quite small in value, usually either equal to or less than 1. It is thus the case that a child may need to posit many of the same kinds of error in order to successfully re-rank the constraints in question. However, it was shown by Pater (2008) that setting w and l to be equivalent would cause the GLA to fail to converge in certain cases. The CGLA (Magri, 2012) showed that this issue could be avoided if the total amount of promotion per row, W, was always less than the total amount of demotion, L. One simple way of ensuring that this is the case is to demote all error-preferring constraints by 1, but to promote all adult form-preferring constraints according to the formula $\frac{L}{W+1}$. A summary of how to construct the update vectors under the CGLA is given in (10).

(10) **Summary of how to construct an update vector in the CGLA**:

a. Any undominated error-preferring constraints (shown with an L) are **demoted** by 1.

b. Any constraints that prefer neither form (shown with an *e*), or prefer an error and are dominated by an adult form-preferring constraint, are assigned the value 0.

c. Any adult form-preferring constraints (shown with a W) are **promoted** according to the formula $\frac{L}{W+1}$.

For the remainder of the paper, the update vector generated from an error tableau will be shown to the right of that tableau. An example is given in (11).

(11) **Sample error tableau with update vector**:

/ɪmpʊt/	C_1	C_2	C_3	
☞ a. ˈɛɹəɹ	(L)	∗∗ (W)	∗(*e*)	$\rightarrow (-1, \frac{1}{2}, 0)$
b. ˈtɑɹgət	∗!∗	∗	∗	

Since the CGLA is guaranteed to converge in all cases, I will adopt it for the remainder of this paper, although it should be noted that the GLA is also incapable of predicting children's intermediate faithfulness stages, as shown by Jesney and Tessier (2007, 2008). An illustration of how the CGLA functions will be given in the following section, as well as a demonstration of how it fails to predict the Québec French intermediate faithfulness stage.

3 The CGLA fails to predict intermediate faithfulness stages

As observed by Jesney and Tessier (2007, 2008) and Tessier (2009), gradual OT learners like the GLA fail to predict the existence of intermediate faithfulness stages in child language acquisition under the standard assumption that all MARKEDNESS constraints started out ranked above

all FAITHFULNESS constraints. This diagnosis is also correct for the CGLA, and their observations will be extended to it in the following section. In order to do so, two of Théo's erroneous productions from (1b) and (2b) will be considered. These errors are repeated below.

(12) **Sample errors for Théo**:

Target	*Child*	*Gloss*
/ˈtχɛ̃/	[ˈkɛ]	'train', (1b)
/gʁyˈjo/	[kʰœˈjɔ]	'oatmeal', (2b)

The ranking that describes Théo's initial state grammar shown in (4) is compatible with the standard assumptions about initial state grammars – namely, that the MARKEDNESS constraint, *COMPLEX, outranks both of the FAITHFULNESS constraints, MAX/$\acute{\sigma}$ and MAX. It will therefore be assumed for this example that this is the initial state of the grammar for the CGLA learner, and that this ranking can be encoded by the ranking vector shown in (13).

(13) **Initial ranking vector for Théo**:

*COMPLEX	MAX/$\acute{\sigma}$	MAX
(12,	0,	0)

When the CGLA learner uses this ranking to predict a form for the adult Québec French word /ˈtχɛ̃/, 'train', it will instead produce the form [ˈkɛ], which is assumed to be identical to Théo's production at this stage.[2] After constructing this erroneous output, the CGLA learner will then compare it to the input form according to Prince's (2002) criteria, and generate an update vector according to the schema in (10). The comparison will provide the learner with evidence that *COMPLEX needs to be demoted, since it prefers the child's error and is not dominated by an adult form-preferring constraint. As such, it will be assigned the value −1 in the update vector. Similarly, the comparison provides evidence that both MAX/$\acute{\sigma}$ and MAX need to be promoted, since both of these constraints prefer the adult form. As such, they will each be assigned the value $\frac{1}{3}$ in the update vector, since there is one error-preferring constraint and two adult form-preferring constraints, and according to the formula in (10), the promotion amount should be equivalent to $\frac{L}{W+1} = \frac{1}{2+1} = \frac{1}{3}$.

(14) **Update vector generated by** /ˈtχɛ̃/ → [ˈkɛ]:

/ˈtχɛ̃/	*COMPLEX	MAX/$\acute{\sigma}$	MAX	
a. ˈtχɛ̃	∗!			→ (−1, $\frac{1}{3}$, $\frac{1}{3}$)
☞ b. ˈkɛ	(L)	∗ (W)	∗ (W)	

Similarly, when the CGLA learner uses this initial ranking to predict a form for the adult Québec French word /gʁyˈjo/, 'oatmeal', it will instead produce Théo's [kʰœˈjɔ]. Using this error to construct an update vector will provide the learner with evidence that *COMPLEX should be demoted and MAX promoted, but it will crucially not provide evidence that MAX/$\acute{\sigma}$ should be promoted.

[2] It could also be the case that the CGLA learner posits more faithful errors like [ˈtɛ̃], since the constraint set examined does not have any other MARKEDNESS constraints, namely, one that would ensure that the vowel [ɛ̃] must denazalize to [ɛ]. For the sake of simplicity in comparison, I will instead opt to show the child's error rather than any other error that could be posited by the CGLA.

(15) **Update vector generated by** /ɡʁyˈjo/ → [kʰœˈjɔ]:

/ɡʁyˈjo/	*Complex	Max/σ́	Max	
a. ɡʁyˈjo	∗!			→ (−1, 0, $\frac{1}{2}$)
☞ b. kʰœˈjɔ	(L)	(*e*)	∗ (W)	

Assuming that the learner cycles through only these two errors, it will go through the learning path outlined in (16). The error posited by the learner is given on the far left of each line, and the result of adding its associated update vector to the current ranking vector will be shown in the centre. If this update will cause the learner to change any rankings in the grammar, this will be indicated at the far right of each line by assigning each new grammar a number. The learning path followed will be summarized below the table.

(16) **Learning path for Québec French under intermediate faithfulness**:

a. *Learning steps triggered by* [ˈkɛ] *and* [kʰœˈjɔ]:

Error		*Complex	Max/σ́	Max		*Grammar*
	(	**12**,	**0**,	**0**	)	Initial
ˈkɛ	(	$12-1=\mathbf{11}$,	$0+\frac{1}{3}=\mathbf{\frac{1}{3}}$,	$0+\frac{1}{3}=\mathbf{\frac{1}{3}}$	)	
kʰœˈjɔ	(	$11-1=\mathbf{10}$,	$\frac{1}{3}+0=\mathbf{\frac{1}{3}}$,	$\frac{1}{3}+\frac{1}{2}=\mathbf{\frac{5}{6}}$	)	①
ˈkɛ	(	$10-1=\mathbf{9}$,	$\frac{1}{3}+\frac{1}{3}=\mathbf{\frac{2}{3}}$,	$\frac{5}{6}+\frac{1}{3}=\mathbf{1\frac{1}{6}}$	)	
kʰœˈjɔ	(	$9-1=\mathbf{8}$,	$\frac{2}{3}+0=\mathbf{\frac{2}{3}}$,	$1\frac{1}{6}+\frac{1}{2}=\mathbf{1\frac{2}{3}}$	)	
ˈkɛ	(	$8-1=\mathbf{7}$,	$\frac{2}{3}+\frac{1}{3}=\mathbf{1}$,	$1\frac{2}{3}+\frac{1}{3}=\mathbf{2}$	)	
kʰœˈjɔ	(	$7-1=\mathbf{6}$,	$1+0=\mathbf{1}$,	$2+\frac{1}{2}=\mathbf{2\frac{1}{2}}$	)	
ˈkɛ	(	$6-1=\mathbf{5}$,	$1+\frac{1}{3}=\mathbf{1\frac{1}{3}}$,	$2\frac{1}{2}+\frac{1}{3}=\mathbf{2\frac{5}{6}}$	)	
kʰœˈjɔ	(	$5-1=\mathbf{4}$,	$1\frac{1}{3}+0=\mathbf{1\frac{1}{3}}$,	$2\frac{5}{6}+\frac{1}{2}=\mathbf{3\frac{1}{3}}$	)	
ˈkɛ	(	$4-1=\mathbf{3}$,	$1\frac{1}{3}+\frac{1}{3}=\mathbf{1\frac{2}{3}}$,	$3\frac{1}{3}+\frac{1}{3}=\mathbf{3\frac{2}{3}}$	)	②

b. *Learning path followed:*

Initial grammar: *Complex ≫ Max/σ́, Max

↓

Grammar ①: *Complex ≫ Max ≫ Max/σ́

↓

Grammar ②: Max ≫ *Complex ≫ Max/σ́

At the last step of the learning path outlined above, Max has been promoted to be above *Complex, while Max/σ́ has been promoted, but is still dominated by *Complex. At this stage, the next time the CGLA learner generates an output for /ɡʁyˈjo/, it will do so *completely faithfully* – that is, the learner will *not* posit an error, and learning should stop at this stage. This ranking will predict the same output forms as the final stage ranking outlined in (4), where Max and Max/σ́ both dominate *Complex. The CGLA learner has therefore successfully acquired a grammar compatible with the adult language. However, it should be noted in the learning

path outlined above that there is no point at which MAX/*σ́* dominates *COMPLEX, which in turn dominates MAX – that is, there is *no intermediate faithfulness stage*.

The reason that this is so lies in the way the update vectors in (14) and (15) are generated. The update vector in (14), associated with the error [ˈkɛ], promotes both MAX and MAX/*σ́*. This is desirable, since eventually both constraints should be ranked above *COMPLEX. However, the update vector in (15), associated with the error [kʰœˈjɔ], only promotes MAX, meaning that MAX will be promoted at a faster rate than MAX/*σ́*. As such, it will outrank *COMPLEX first, leading to the lack of the intermediate faithfulness stage.

This will in fact be true of any two FAITHFULNESS constraints where one FAITHFULNESS constraint is a positional variant of the first. Since a positional FAITHFULNESS constraint will always receive a subset of the violations of a general FAITHFULNESS constraint, it will always be promoted at a slower rate than its more general counterpart (Jesney and Tessier, 2007, 2008; Tessier, 2009). As such, the general FAITHFULNESS constraint will always be re-ranked above a given MARKEDNESS constraint either at the same time as its positional variant (if both are always promoted), or before its positional variant. It is thus clear that intermediate faithfulness stages are *never* predicted to occur when using the CGLA in combination with positional FAITHFULNESS constraints to model language acquisition.

Various solutions have been proposed in order to derive intermediate faithfulness stages using gradual learning algorithms. Jesney and Tessier (2007, 2008) argue that a different, HG-based learning algorithm should be used to model these phenomena instead. In a similar vein, Tessier (2009) argues that a gradual version of Hayes's (2004) Low Faithfulness Constraint Demotion algorithm is needed in order to model these stages. In the following section, I will demonstrate that there is yet another solution to this issue, where what must be changed is not the learning algorithm used, but rather the characterization of the intermediate stage.

4 The CGLA is able to predict intermediate markedness stages

In this section, I will show that the intermediate stage in Québec French complex onset acquisition does *not* have to be characterized by ranking the positional FAITHFULNESS constraint MAX/*σ́* above the general MARKEDNESS constraint *COMPLEX, while leaving the general FAITHFULNESS constraint MAX low-ranked. Rather, I will show that these stages are also able to be characterized by ranking a positional MARKEDNESS constraint above a general FAITHFULNESS constraint, while leaving the more general MARKEDNESS constraint low-ranked. This is the schema outlined in (6), repeated in (17), below.

(17) **Ranking schema for intermediate markedness stages**:
Positional MARKEDNESS ≫ FAITHFULNESS ≫ General MARKEDNESS

In the following subsections, I will first demonstrate how the schema in (6/17) can be used to characterize the intermediate stage in Québec French. I will then demonstrate how the CGLA is able to predict the intermediate stage under this characterization. Finally, it will be shown that the CGLA will *always* be able to predict such a stage, regardless of the how often the learner is presented with the crucial evidence for demoting the positional MARKEDNESS constraint.

4.1 Positional markedness and intermediate stages

As originally observed by Beckman (1998), the positions targeted by positional FAITHFULNESS constraints are often those that are also targeted by positional MARKEDNESS constraints (Ito, 1986)[3]. That is, in many cases, the effects of positional FAITHFULNESS constraints can be expressed just as well with contextual MARKEDNESS constraints or positional licensing constraints. For instance, there is a natural re-formulation of MAX/$\acute{\sigma}$, a positional FAITHFULNESS constraint, as a positional MARKEDNESS constraint *COMPLEX/$\check{\sigma}$. This constraint is defined in (18).

(18) *COMPLEX/$\check{\sigma}$: Syllables that do not bear stress cannot license complex onsets.

Since *COMPLEX/$\check{\sigma}$ is the positional MARKEDNESS correlate of the positional FAITHFULNESS constraint MAX/$\acute{\sigma}$, the Québec French stages outlined in (1) through (3) should also be able to be characterized by an appropriate ranking of *COMPLEX/$\check{\sigma}$. The first stage, where neither Clara nor Théo has complex onsets, can be adequately captured by ranking both *COMPLEX/$\check{\sigma}$ and *COMPLEX above MAX.[4] Similarly, the final stage, where Clara and Théo both successfully produce complex onsets in all environments where they are seen in the adult language, can be characterized by ranking MAX above both *COMPLEX/$\check{\sigma}$ and *COMPLEX. The intermediate stage, on the other hand, must be characterized by the ranking in (19). This ranking is sufficient for characterizing the intermediate stage, as shown by the two tableaux in (20).

(19) **Intermediate markedness ranking for Québec French**:
*COMPLEX/$\check{\sigma}$ ≫ MAX ≫ *COMPLEX

(20) **Intermediate markedness tableaux for Québec French**:

a. *Théo correctly maps* /ˈtχɛ̃/ → [ˈkχɛ]:

/ˈtχɛ̃/	*COMPLEX/$\check{\sigma}$	MAX	*COMPLEX
a. ˈkɛ		*!	
☞ b. ˈkχɛ			*

b. *Théo incorrectly maps* /ɡʁyˈjo/ → [kʰœˈjɔ]:

/ɡʁyˈjo/	*COMPLEX/$\check{\sigma}$	MAX	*COMPLEX
a. kχœˈjɔ	*!		*
☞ b. kʰœˈjɔ		*	

Using this positional MARKEDNESS constraint will allow the learning path for Clara and Théo to be characterized as shown in (21), repeated from (7).

[3] I assume that positional licensing and contextual MARKEDNESS constraints are both subsumed under the label of positional MARKEDNESS

[4] What is crucial in this ranking is that *COMPLEX outranks MAX – it does not matter where *COMPLEX/$\check{\sigma}$ is ranked. As such, I will assume that it is ranked in the top stratum with *COMPLEX, in accordance with the assumption that the initial state grammar is one where all MARKEDNESS constraints outrank all FAITHFULNESS constraints.

(21) **Learning path for Clara & Théo in terms of positional markedness**:

Initial stage grammar:	*COMPLEX/$\breve{\sigma}$, *COMPLEX ≫ MAX
Intermediate stage grammar:	*COMPLEX/$\breve{\sigma}$ ≫ MAX ≫ *COMPLEX
Final stage grammar:	MAX ≫ *COMPLEX/$\breve{\sigma}$, *COMPLEX

In the following subsection, it will be shown that the CGLA will go through this *exact* learning path, thus successfully predicting an intermediate stage for Clara and Théo.

4.2 The CGLA and intermediate markedness

For this illustration, it will be assumed that the constraints in CON are exactly the same as they are in section 3, with the addition of the positional MARKEDNESS constraint *COMPLEX/$\breve{\sigma}$. This is done since, as claimed by Jesney (2011), classical OT must contain both positional MARKEDNESS and positional FAITHFULNESS constraints in order to adequately predict the typology of positional effects across languages, and also to show that including the positional FAITHFULNESS constraint in CON is unproblematic for the learner. Discussion of whether certain positional FAITHFULNESS constraints can be eliminated from CON will be given in section 6.2.

Under standard learning assumptions, both of the MARKEDNESS constraints, *COMPLEX/$\breve{\sigma}$ and *COMPLEX, will be ranked above both of the FAITHFULNESS constraints, MAX/$\acute{\sigma}$ and MAX. An initial ranking vector corresponding to this ranking is given in (22).

(22) **Initial ranking vector for Théo**:

*COMPLEX/$\breve{\sigma}$	*COMPLEX	MAX/$\acute{\sigma}$	MAX
(12,	12,	0,	0)

As in the intermediate faithfulness case in section 3, the learner will use this grammar to predict the form [ˈkɛ] for Québec French /ˈtχɛ̃/, 'train'. This form will then be compared to the input form, and it will generate the update vector shown in (23). This update vector will provide the learner with evidence that *COMPLEX needs to be demoted, while both MAX and MAX/$\acute{\sigma}$ need to be promoted (just as in section 3). It will not provide any evidence that *COMPLEX/$\breve{\sigma}$ needs to be demoted, however.

(23) **Update vector generated by** /ˈtχɛ̃/ → [ˈkɛ]:

/ˈtχɛ̃/	*COMPLEX/$\breve{\sigma}$	*COMPLEX	MAX/$\acute{\sigma}$	MAX	
a. ˈkχɛ		∗!			→ $(0, -1, \frac{1}{3}, \frac{1}{3})$
☞ b. ˈkɛ	(*e*)	(L)	∗ (W)	∗ (W)	

The CGLA learner will also use this initial ranking to predict the form [kʰœˈjɔ] for Québec French /gʁyˈjo/, 'oatmeal'. When comparing this form to the input form, the learner will generate the update vector shown in (24). This update vector provides evidence that both *COMPLEX

and *COMPLEX/$\check{\sigma}$ need to be demoted, and that MAX needs to be promoted. It will not provide evidence that MAX/$\acute{\sigma}$ should be promoted.[5]

(24) **Update vector generated by** /gʁyˈjo/ → [kʰœˈjɔ]:

/gʁyˈjo/	*COMPLEX/$\check{\sigma}$	*COMPLEX	MAX/$\acute{\sigma}$	MAX	
a. kχœˈjɔ	∗!	∗			→ (−1,−1,0,1)
☞ b. kʰœˈjɔ	(L)	(L)	(*e*)	∗ (W)	

Assuming that the learner cycles through only these two errors, it will to go through the learning path outlined in (25).

(25) **Learning path for Québec French: initial to intermediate markedness**:

a. *Learning steps triggered by* [ˈkɛ] *and* [kʰœˈjɔ]:

Error		*COMPLEX/$\check{\sigma}$	*COMPLEX	MAX/$\acute{\sigma}$	MAX		*Grammar*
	(	**12**,	**12**,	**0**,	**0**	)	Initial
ˈkɛ	(	$12+0=\mathbf{12}$,	$12-1=\mathbf{11}$,	$0+\frac{1}{3}=\mathbf{\frac{1}{3}}$,	$0+\frac{1}{3}=\mathbf{\frac{1}{3}}$	)	①
kʰœˈjɔ	(	$12-1=\mathbf{11}$,	$11-1=\mathbf{10}$,	$\frac{1}{3}+0=\mathbf{\frac{1}{3}}$,	$\frac{1}{3}+1=\mathbf{1\frac{1}{3}}$	)	②
ˈkɛ	(	$11+0=\mathbf{11}$,	$10-1=\mathbf{9}$,	$\frac{1}{3}+\frac{1}{3}=\mathbf{\frac{2}{3}}$,	$1\frac{1}{3}+\frac{1}{3}=\mathbf{1\frac{2}{3}}$	)	
kʰœˈjɔ	(	$11-1=\mathbf{10}$,	$9-1=\mathbf{8}$,	$\frac{2}{3}+0=\mathbf{\frac{2}{3}}$,	$1\frac{2}{3}+1=\mathbf{2\frac{2}{3}}$	)	
ˈkɛ	(	$10+0=\mathbf{10}$,	$8-1=\mathbf{7}$,	$\frac{2}{3}+\frac{1}{3}=\mathbf{1}$,	$2\frac{2}{3}+\frac{1}{3}=\mathbf{3}$	)	
kʰœˈjɔ	(	$10-1=\mathbf{9}$,	$7-1=\mathbf{6}$,	$1+0=\mathbf{1}$,	$3+1=\mathbf{4}$	)	
ˈkɛ	(	$9+0=\mathbf{9}$,	$6-1=\mathbf{5}$,	$1+\frac{1}{3}=\mathbf{1\frac{1}{3}}$,	$4+\frac{1}{3}=\mathbf{4\frac{1}{3}}$	)	
kʰœˈjɔ	(	$9-1=\mathbf{8}$,	$5-1=\mathbf{4}$,	$1\frac{1}{3}+0=\mathbf{1\frac{1}{3}}$,	$4\frac{1}{3}+1=\mathbf{5\frac{1}{3}}$	)	③

[5]It should be noted that the tableau in (24) is not a normal form tableau, where there is one L per relevant row, as required by the CGLA (Magri, 2012). However, decomposing this tableau into two separate normal form tableaux will give the same result in this case, as shown in (i).

(i) **Normal form tableaux give the same result as in (24)**:

a. *Tableau showing* *COMPLEX/$\check{\sigma}$ *must be dominated by* MAX:

/gʁyˈjo/	*COMPLEX/$\check{\sigma}$	*COMPLEX	MAX/$\acute{\sigma}$	MAX	
a. kχœˈjɔ	∗!	∗			→ $(-1,0,0,\frac{1}{2})$
☞ b. kʰœˈjɔ	(L)	(*e*)	(*e*)	∗ (W)	

b. *Tableau showing* *COMPLEX *must be dominated by* MAX:

/gʁyˈjo/	*COMPLEX/$\check{\sigma}$	*COMPLEX	MAX/$\acute{\sigma}$	MAX	
a. kχœˈjɔ	∗!	∗			→ $(0,-1,0,\frac{1}{2})$
☞ b. kʰœˈjɔ	(*e*)	(L)	(*e*)	∗ (W)	

c. *Sum of update vectors*:
$(-1,0,0,\frac{1}{2})+(0,-1,0,\frac{1}{2})=(-1,-1,0,1)$

I will therefore use the update vector in (24), since it conveys the same information in a more concise way.

b. *Learning path followed*:

Initial grammar: *COMPLEX/$\breve{\sigma}$, *COMPLEX ≫ MAX/$\acute{\sigma}$, MAX

↓

Grammar (1): *COMPLEX/$\breve{\sigma}$ ≫ *COMPLEX ≫ MAX/$\acute{\sigma}$, MAX

↓

Grammar (2): *COMPLEX/$\breve{\sigma}$ ≫ *COMPLEX ≫ MAX ≫ MAX/$\acute{\sigma}$

↓

Grammar (3): *COMPLEX/$\breve{\sigma}$ ≫ MAX ≫ *COMPLEX ≫ MAX/$\acute{\sigma}$

At the last step of the learning path outlined above, MAX has been promoted to be above *COMPLEX, *but it has not been promoted above* *COMPLEX/$\breve{\sigma}$. Rather, it sits in the middle of these two constraints, corresponding to the ranking that predicts the intermediate stage, as outlined in (20) in section 4.1. Thus, this CGLA learner has successfully arrived at an intermediate stage.

The reason why the intermediate markedness stage is predicted turns out to be parallel to the reason why the intermediate faithfulness stage fails to be predicted. To see why this is so, the update vectors in (23) and (24) can be examined. The update vector in (24), associated with the error [ˈkɛ], demotes both *COMPLEX and *COMPLEX/$\breve{\sigma}$. This is desirable, since eventually both of these constraints need to be dominated by MAX in Québec French. However, the update vector in (23), associated with the error [kʰœˈjɔ], *only demotes* *COMPLEX, meaning that it will be demoted at a faster rate than *COMPLEX/$\breve{\sigma}$. As such, it will outrank the topmost ranked FAITHFULNESS constraint (in this case, MAX) first, leading to the intermediate markedness stage.

This will be true of any two MARKEDNESS constraints, where one MARKEDNESS constraint is a positional version of the first. Since a positional MARKEDNESS constraint will always receive a subset of the violations of a general MARKEDNESS constraint, it will always be demoted at a slower rate than its more general counterpart. This should predict that, whenever a positional MARKEDNESS constraint is present in CON, it will lead to an intermediate stage.

Of course, now that an intermediate stage has been reached, it should also be shown that the CGLA learner can still arrive at the adult language from this stage. This is still possible, since the learner is still generating errors that will provide evidence that *COMPLEX/$\breve{\sigma}$ should be demoted and that MAX should be promoted. At the intermediate stage, the learner will correctly produce a complex onset in words like /ˈtχɛ̃/, 'train', parallel to Théo's production [ˈkχɛ]. Since these forms are consistent with the adult language ranking of *COMPLEX and MAX, they will not contribute any additional information about the adult ranking. However, the learner will still make errors on words like /gʁyˈjo/, 'oatmeal', which will still be produced like [kʰœˈjɔ]. Since the learner is still making an error, it will still be generating an update vector, as shown in (26).

(26) **Update vector generated by** /gʁyˈjo/ → [kʰœˈjɔ]:

/gʁyˈjo/	*COMPLEX/$\breve{\sigma}$	MAX	*COMPLEX	MAX/$\acute{\sigma}$	
a. kχœ'jɔ	∗!		∗		→ $(-1, \frac{1}{2}, 0, 0)$
☞ b. kʰœˈjɔ	(L)	∗ (W)	(L)	(*e*)	

While the comparative tableau is the same as the one in (24) with the constraint order switched, the update vector has changed, since *COMPLEX is now dominated by the adult form-preferring constraint MAX and does not need to be demoted further. Since there is now only

one active error-preferring constraint, *COMPLEX/$\breve{\sigma}$, the promotion amount will be raised to $\frac{1}{2}$. If the learner generates this error twice more, it will arrive at a constraint ranking compatible with the adult language, as shown in the learning path outlined in (27).

(27) **Learning path for Québec French: intermediate markedness to adult-like:**

a. *Learning steps triggered by* [kʰœˈjɔ]:

Error		*COMPLEX/$\breve{\sigma}$	MAX	*COMPLEX	MAX/$\acute{\sigma}$		*Grammar*
	(	**8**,	$\mathbf{5\frac{1}{3}}$,	**4**,	$\mathbf{1\frac{1}{3}}$	)	③
kʰœˈjɔ	(	$8-1=\mathbf{7}$,	$5\frac{1}{3}+\frac{1}{2}=\mathbf{5\frac{5}{6}}$,	$4+0=\mathbf{4}$,	$1\frac{1}{3}+0=\mathbf{1\frac{1}{3}}$	)	
kʰœˈjɔ	(	$7-1=\mathbf{6}$,	$5\frac{5}{6}+\frac{1}{2}=\mathbf{6\frac{1}{3}}$,	$4+0=\mathbf{4}$,	$1\frac{1}{3}+0=\mathbf{1\frac{1}{3}}$	)	④

b. *Learning path followed*:

Grammar ③: *COMPLEX/$\breve{\sigma}$ ≫ MAX ≫ *COMPLEX ≫ MAX/$\acute{\sigma}$

↓

Grammar ④: MAX ≫ *COMPLEX/$\breve{\sigma}$ ≫ *COMPLEX ≫ MAX/$\acute{\sigma}$

At this stage, the learner will be faithfully producing complex onsets wherever they appear in the adult language, and learning will cease. Thus, not only is the CGLA learner able to predict intermediate stages when making use of positional MARKEDNESS constraints, but it is also able to successfully learn a ranking compatible with the adult language. In the next subsection, it will be shown that this is the case no matter how often each kind of error is produced.

4.3 Intermediate stages always predicted

Because the update triggered by the error [ˈkɛ] always assigns a subset of the violations to *COMPLEX and *COMPLEX/$\breve{\sigma}$ as the error [kʰœˈjɔ], it will always be the case that [ˈkɛ] will predicted to be produced faithfully before [kʰœˈjɔ]. That is, there should *always* be an intermediate stage for complex onset acquisition in Québec French, no matter how often the learner attempts to generate each error.

This is shown more concretely by Figure 1. The graph in the figure is the result of running a modified version of Giorgio Magri's `EDRAexcel.py` script with varying frequencies of /ˈtχɛ̃/ and /gʁyˈjo/ input. The controlled variable was how often the learner attempted to produce the form /gʁyˈjo/, ranging from 10% of the data to 90% of the data, in steps of 10%. The form /ˈtχɛ̃/ made up the remainder of the data that the learner was trained on. The case where /gʁyˈjo/ was produced 0% of the time was left out, as the learner would not converge on the correct output grammar, since it would never receive evidence that *COMPLEX/$\breve{\sigma}$ should be demoted. The case where /gʁyˈjo/ was produced 100% of the time was also left out, as it would predict that both /ˈtχɛ̃/ and /gʁyˈjo/ would be produced faithfully at the same time.

The script was run five times for each input frequency with the following parameters: `theta_value = 100, iterations_per_stage_vector = [7000, 7000, 7000], plasticity_per_stage_vector = [0.1, 0.1, 0.1], noise_type = 'UNIFORM', sigma_per_stage_vector = [10.0, 2.0, 2.0], update_rule='CEDRA', num_repetitions_evaluation = 100000, sigma_evaluation`

= 2.0, printing_sparsity = 25. For each time the script was run, the step at which each input form was predicted to be produced faithfully was recorded. This was taken to be the first recorded step at which the relevant constraint rankings had been achieved – MAX ≫ *COMPLEX for /ˈtχɛ̃/, and MAX ≫ *COMPLEX/$\breve{\sigma}$ for /ɡʁyˈjo/. The average for each of the five results was then found and plotted.

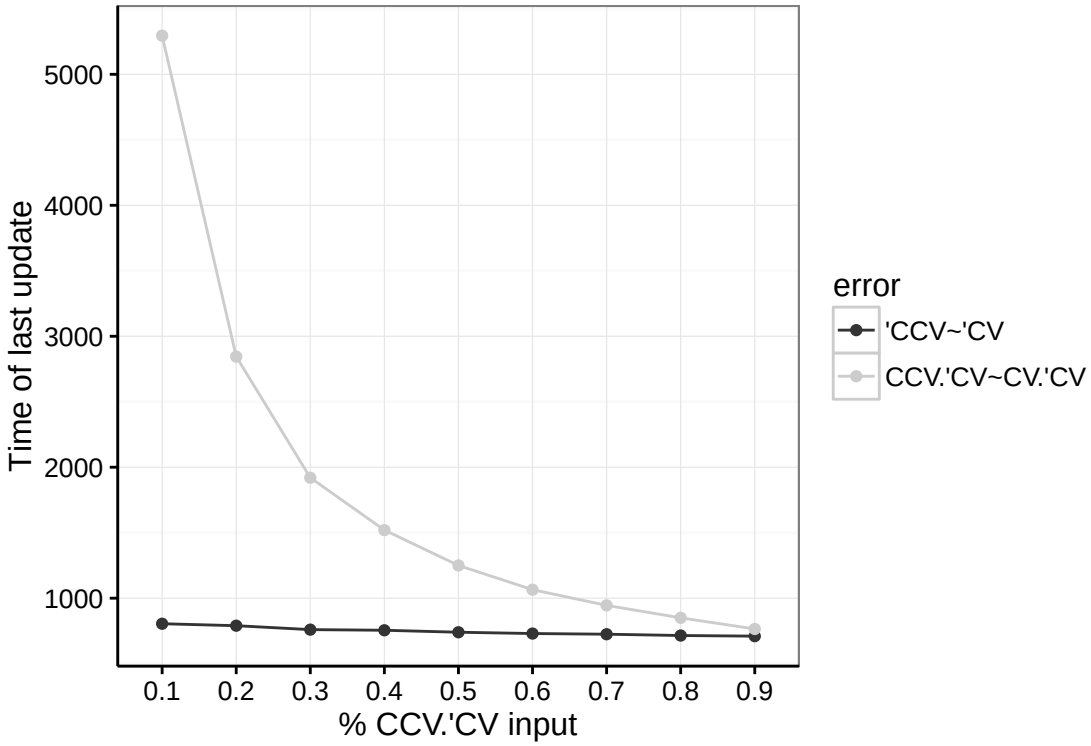

Figure 1: **Length of intermediate stage for Québec French**, as measured by the percentage of /ɡʁyˈjo/-like input and the time (learning step) at which each error was no longer produced

Two things are of note in the above graph. The first is that the time at which input /ˈtχɛ̃/ is expected to be produced faithfully is relatively constant. This is due to the fact that both errors provide evidence that *COMPLEX needs to be demoted and that MAX needs to be promoted. The second is that it is *always* the case that the time at which input /ɡʁyˈjo/ is produced faithfully is later than this constant time. That is, there will always be an intermediate stage. Furthermore, the fewer times the input /ɡʁyˈjo/ is attempted, the longer the intermediate stage will last.

This is a potentially troubling finding – documented intermediate stages are not a guarantee. For instance, Rose (2000) collected speech data from three Québec French-learning children, and only observed this intermediate stage for two of them. However, the graph above suggests a potential solution to this issue. If it is the case that, in Québec French (or French more generally), there are more /ɡʁyˈjo/-like words than /ˈtχɛ̃/-like words, then the intermediate stage would be predicted to be relatively short. This is shown by the right-hand side of

the graph, where the intermediate stage only lasts a few hundred steps. If the child has gone through this stage between recording sessions, then it would appear that that child had skipped the stage altogether. An examination of existing French corpora should be undertaken to provide an estimate of whether this is true for all children learning French.

While the CGLA learner successfully parallels the acquisition of complex onsets in Québec French as shown in (21) when allowed to make use of positional MARKEDNESS constraints, it remains to be seen whether this finding can be generalized to other phenomena which have been analyzed as intermediate FAITHFULNESS stages. This is the goal of the next section, which outlines an attempt at analyzing English unstressed syllable acquisition in terms of positional MARKEDNESS constraints.

5 Expanding the intermediate markedness analysis to English

As outlined in Tessier (2009), there is an extensive literature on the acquisition of English phonology that claims that children prefer to retain syllables in prominent positions, such as stressed syllables (Echols and Newport, 1992; Fikkert, 1994), right-edge syllables (Echols and Newport, 1992; Fikkert, 1994; Pater, 1997), left-edge syllables (Pater, 1997), and footed syllables (Gerken, 1996). The studies conducted by Kehoe and Stoel-Gammon (1997) and Kehoe (2000) show that, at least initially, children seem to prefer to retain syllables that are stressed, as well as unstressed syllables at the right edge of the word. Some data illustrating this stage are given in (28).[6]

(28) **First stage in unstressed syllable acquisition**, data from 34m3 (Kehoe, 2000, p.42):[7]

Target	*Child*	*Gloss*
/bəˈnænə/	[ˈnɑnʌ]	'banana'
/təˈmeɪ̯roʊ̯/	[ˈmedo]	'tomato'
/ˈɛləfənt/	[ˈʔʌfɪnt]	'elephant'
/ˈɑktəpəs/	[ˈɑʔpʊs]	'octopus'
/ˈtɛləˌfoʊ̯n/	[ˈkɑfo]	'telephone'
/ˌkæŋgəˈɹu/	[ˌkæŋˈruː]	'kangaroo'
/ˈælɪˌgeɪ̯ɾɚ/	[ˈæˌgejɚ]	'alligator'
/ˌɑvəˈkɑdo/	[ˌʌˈkɑgo]	'avocado'

Under Kehoe's (2000) analysis, this stage is characterized by the ranking in (29). It should be noted that there are two positional FAITHFULNESS constraints, MAX/$\acute{\sigma}$ and ANCHOR-R, that are ranked above a MARKEDNESS constraint, *$\breve{\text{V}}$ (Kehoe's *UNSTRESS), which is in turn ranked above a general FAITHFULNESS constraint, MAX. This ranking meets Tessier's (2009) criteria for being

[6] Kehoe (2000) observes an additional stage prior to the one outlined above, which she claims is characterized by left-alignment of primary stress. However, the data she provides are indistinguishable from those in (28), where stressed and right-edge unstressed syllables are retained, so I will begin my analysis from this stage instead.

[7] Kehoe (2000) does not provide transcriptions for the adult targets. I have therefore provided a broad transcription of my own speech in order to make these data more parallel to the Québec French data provided above. All transcriptions of child speech are from Kehoe (2000), and reflect the most common spontaneous production for each word when possible, although imitated productions were used if no spontaneous productions were observed.

an intermediate faithfulness stage, as outlined in (5). I will follow both Kehoe and Tessier in considering this stage to be intermediate as well, although I have chosen to make use of Kehoe's constraint set rather than Tessier's.

(29) **Constraint ranking for the first stage in unstressed syllable acquisition**:
ALIGN-$\acute{\sigma}$-L, MAX/$\acute{\sigma}$, ANCHOR-R $\gg$ *$\check{\text{V}}$ $\gg$ MAX

The next stage identified by Kehoe is one where children are able to produce unstressed medial syllables, provided that those syllables do not begin with a sonorant consonant. Unstressed medial syllables with a sonorant onset are still eliminated.

(30) **Second stage in unstressed syllable acquisition**, data primarily from 34f3 (Kehoe, 2000, p.43):[8]

Target	*Child*	*Gloss*
/bəˈnænə/	[ˈbænʌ]	'banana'
/təˈmeɪɾo/	[ˈmedo]	'tomato'
/ˈɛləfənt/	[ˈɑfɪn]	'elephant'
/ˈaktəpəs/	[ˈaʔtəpʊs]	'octopus'
/ˈtɛləˌfo͡ʊn/	[ˈtɛrˌfon]	'telephone'
/ˈkɹɑkəˌda͡ɪl/	[ˈʔɛkəˌdɑɪ]	'crocodile'
/ˌkæŋgəˈɹu/	[ˌkɛgɑˈruː]	'kangaroo'
/ˈælɪˌge͡ɪɾɚ/	[ˈeːˌgerə]	'alligator'

Kehoe analyzes this stage by saying that *$\check{\text{V}}$ has been ranked below MAX, but that there is an additional MARKEDNESS constraint ranked above MAX which prohibits sonorous onsets, *SONONS. Her full ranking is provided in (31), below. This ranking also corresponds to an intermediate faithfulness stage, since ANCHOR-R and MAX/$\acute{\sigma}$ are ranked above a MARKEDNESS constraint – this time, *SONONS – which is in turn ranked above MAX.

(31) **Constraint ranking for the second stage in unstressed syllable acquisition**:
ALIGN-$\acute{\sigma}$-L, ANCHOR-R, MAX/$\acute{\sigma}$ $\gg$ *SONONS $\gg$ MAX $\gg$ *$\check{\text{V}}$[9]

The third stage in the proposed acquisition path for unstressed syllables is one where initial syllables are produced faithfully, but where unstressed medial syllables with sonorant onsets are still missing.

[8]The first data point comes from another child who has been identified as belonging to the second stage, 28m3. Participant 34f3 instead most often produced the form [bənænʌ]. However, she still produced forms like [ˈmedo] for adult /təˈme͡ɪɾo/, 'tomato'. Since she does not reliably produce unstressed initial syllables, she is classified as belonging to the second stage, and I will follow Kehoe in this classification.

[9]Kehoe does not provide a ranking between MAX/$\acute{\sigma}$ and *SONONS – these are instead grouped in the same stratum. However, I believe that there is evidence that this ranking should exist – otherwise, it would be predicted that /təˈme͡ɪɾo/ should surface as something like [ˈtedo] rather than as [ˈmedo]. I have therefore included the ranking in this analysis.

(32) **Third stage in unstressed syllable acquisition**, data from 34f1 (Kehoe, 2000, p.44):

Target	*Child*	*Gloss*
/bəˈnænə/	[bʌˈnænʌ]	'banana'
/təˈme͡ɪɾo/	[təˈmedo]	'tomato'
/ˈɛləfənt/	[ˈɛlːfɪnt]	'elephant'
/ˈɑktəpəs/	[ˈɑktəpʊs]	'octopus'
/ˈtɛləˌfo͡ʊn/	[ˈtɛlfo]	'telephone'
/ˈkɹɑkəˌda͡ɪl/	[ˈkrɑkəˌdaɪl]	'crocodile'
/ˈælɪˌge͡ɪɾɚ/	[ˈælˌgeɾɚ]	'alligator'
/ˌsɪndɚˈɹɛlə/	[ˌsɪndəˈrɛlʌ]	'Cinderella'

At this stage, ALIGN-$\acute{\sigma}$-L has been demoted below MAX, while all other constraints retain their previous rankings. This is illustrated in (33). This stage also qualifies as an intermediate faithfulness stage because it retains the ANCHOR-R, MAX/$\acute{\sigma}$ ≫ *SONONS ≫ MAX ranking.

(33) **Constraint ranking for the third stage in unstressed syllable acquisition**:
ANCHOR-R, MAX/$\acute{\sigma}$ ≫ *SONONS ≫ MAX ≫ *$\check{\text{V}}$, ALIGN-$\acute{\sigma}$-L

After this stage, children are adult-like in their productions of unstressed syllables, and can be said to have successfully acquired the adult constraint ranking. This is reflected by re-ranking *SONONS and MAX. The full acquisition path is reiterated below.

(34) **Learning path for English children in terms of positional faithfulness**:

First stage grammar: ANCHOR-R, MAX/$\acute{\sigma}$, ALIGN-$\acute{\sigma}$-L ≫ *SONONS, *$\check{\text{V}}$ ≫ MAX

Second stage grammar: ANCHOR-R, MAX/$\acute{\sigma}$, ALIGN-$\acute{\sigma}$-L ≫ *SONONS ≫ MAX ≫ *$\check{\text{V}}$

Third stage grammar: ANCHOR-R, MAX/$\acute{\sigma}$ ≫ *SONONS ≫ MAX ≫ ALIGN-$\acute{\sigma}$-L, *$\check{\text{V}}$

Final stage grammar: ANCHOR-R, MAX/$\acute{\sigma}$, MAX ≫ *SONONS, ALIGN-$\acute{\sigma}$-L, *$\check{\text{V}}$

The remainder of this section is structured as follows. I will first attempt to construct an intermediate markedness account of English unstressed syllable acquisition. It will be shown that, while the effects of MAX/$\acute{\sigma}$ can successfully be ascribed to the effects of MARKEDNESS constraints, the same will not be possible for ANCHOR-R, and this particular positional FAITHFULNESS constraint *must* be referenced when describing the English children's grammars. However, even though ANCHOR-R must be included in CON, it will be shown that it can remain low-ranked in the hierarchy and still correctly predict what errors children will produce. It will also be shown that running a CGLA learner on the English data can correctly predict the order of acquisition of unstressed syllables, *if certain assumptions about the frequencies of each kind of input hold.* Discussion of what happens with other input frequencies is taken up at the end of this section.

5.1 Constraints active in the first stage

Before attempting to eliminate the two positional FAITHFULNESS constraints, some preliminary assumptions about what the child has learned about adult English by the first stage will be outlined. Since these are observations that should hold of the first stage, only the data from 34m3 will be analyzed in this subsection, with supplemental data from 24m1 and 38m3 when required.

I will first assume, along with Kehoe, that there are two FAITHFULNESS constraints relating to stressed syllables that are relevant for child grammars. One of them is MAX/*σ́*, which as already been used in the above analysis. The other is IDENT[stress], which ensures that stress does not change from input to output. While MAX/*σ́* is a positional FAITHFULNESS constraint, IDENT[stress] is not. I will therefore follow Kehoe in assuming that it can be ranked high enough in the children's constraint hierarchy for it to prevent stress shift in output forms. This is illustrated for the first stage in unstressed syllable acquisition in the tableau in (35).

(35) IDENT[stress] **is not violated**, based on data from 24m1 Kehoe (2000, p.48):

/d͡ʒɪˈɹæf/	IDENT[STRESS]	*V̆	MAX
a. dɪˈɹɑf		*!	
b. ˈdɪɹɑf	*!*	*	
☞ c. ˈdɑf			**
d. ˈdiː	*!		***

Contrary to Kehoe, however, I will assume a foot-free analysis of stress in English, using constraints from Gordon (2002). This will simplify the analysis, as it will not require that children have made hypotheses about foot structure or foot construction in order to produce a difference in stressed and unstressed syllables. As such, the constraint ALIGN-*σ́*-L will be calculated over syllables, rather than over feet, as in Kehoe (2000). It will be ranked in the same stratum as *V̆ for the time being, as this ranking is consistent with the data and will allow for a simpler modelling of the CGLA learner. This assumption requires that there is an additional constraint active in the grammar that prevents outputs from consisting of only the input stressed syllables, and requires that there is at least one unstressed syllable present in the output. There are two conceivable constraints that could enforce this – *CLASH and NONFIN. Since children in the first stage freely violate *CLASH in forms like [ˌʌˈkɑgo] for adult /ˌɑvəˈkɑdo/, 'avocado' and [ˈæˌgejɚ] for adult /ˈælɪˌgeɪɹɚ/, 'alligator', it is most likely that NONFIN is the constraint that ensures that at least one unstressed syllable is preserved in the output, and furthermore ensures that it should appear at the right edge of the word. It must be ranked below IDENT[stress] in order to ensure that stress does not shift in order to satisfy it. The ranking of the stress-referencing constraints referenced thus far is provided in (36), along with sample tableaux.

(36) **Ranking of stress-referencing constraints**:
IDENT[stress] ≫ NONFIN ≫ ALIGN-*σ́*-L, *V̆ ≫ MAX, *CLASH

a. *Tableau for* /d͡ʒɪˈɹæf/ → [ˈdɑf], based on data from 24m1 (Kehoe, 2000, p.48):

/d͡ʒɪˈɹæf/	IDENT[STRESS]	NONFIN	ALIGN-σ́-L	*V̌	MAX	*CLASH
a. dɪˈɹɑf		*	*!	*		
b. ˈdɪɹɑf	*!*			*		
☞ c. ˈdɑf		*			**	
d. ˈdiː	*!	*			***	

b. *Tableau for* /ˌɑvəˈkɑdo/ → [ˌʌˈkɑgo], based on data from 34m3 (Kehoe, 2000, p.42):

/ˌɑvəˈkɑdo/	IDENT[STRESS]	NONFIN	ALIGN-σ́-L	*V̌	MAX	*CLASH
a. ˌʌbʌˈkɑgo				**!		
☞ b. ˌʌˈkɑgo				*	**	*
c. ˌʌbʌˈkɑ		*!		*	**	
d. ˌʌˈkɑ		*!			****	*

Since it has been assumed that no child forms will be unfaithful to the stress of the adult form and that children will freely violate *CLASH, both IDENT[stress] and *CLASH will be left out of the remaining tableaux. Similarly, since it will also be assumed that no child forms will unnecessarily violate NONFIN, it will also be left out.

In addition to the constraints outlined above, there are two positional FAITHFULNESS constraints that lead to the emergence of intermediate faithfulness stages in the children's grammars in Kehoe's (2000) analysis – MAX/σ́ and ANCHOR-R. Both are ranked above *V̌ and *SONONS, which in turn are ranked above MAX, leading Tessier (2009) to classify this stage as an intermediate faithfulness stage. However, as has been shown previously, gradual OT learners like the CGLA are incapable of predicting such stages. In order for the correct order of unstressed syllable acquisition to be predicted by the CGLA, these positional FAITHFULNESS constraints must not be ranked as high as previously supposed (or should not be present in the grammar altogether – see discussion in section 6.2), and there should be positional MARKEDNESS constraints that enforce the same or similar restrictions on output forms that are active in children's grammars.

The constraint MAX/σ́ ensures that material in input stressed syllables is retained in the output. Therefore, a MARKEDNESS constraint that has a similar effect should penalize having material in unstressed syllables in the output. There is already such a constraint active in the hierarchy – namely, *V̌, which penalizes having unstressed vowels in the output. As such, I will assume that this is the proper MARKEDNESS analogue to MAX/σ́.

However, once MAX/σ́ is eliminated in favour of *V̌, there is no longer any way to prevent *SONONS from deleting stressed syllables with sonorous onsets, or from replacing sonorous onsets with other available obstruent onsets. That is, it should be predicted that children will produce forms such as /bəˈnænə/ as [ˈbænʌ] rather than [ˈnænʌ], or /təˈmeɪɹo/ as [ˈtedo] rather than [ˈmedo]. While some children in other stages do indeed do this (such as 28m3 in stage 2, (Kehoe, 2000, p.43)), other children do not, and their grammars should also be adequately accounted for. As such, I will assume that there is a positional variant of *SONONS that is active in these children's grammars, *SONONS/σ̌. Together with *V̌ and low-ranked CONTIG preventing discontinuous outputs like [ˈbænʌ], these constraints will give the correct child outputs without making reference to MAX/σ́. It may very well be the case that in other children's grammars,

*SONONS is satisfied (such as 28m3 in stage 2, who produces [ˈbænʌ] for adult /bəˈnænə/, 'banana'), but since this is not true of all children, I will instead opt to use the more restricted constraint.

(37) **The effects of** MAX/*ó* **can be ascribed to** *V̆ **and** *SONONS/*ŏ*, based on data from 34m3 (Kehoe, 2000, p.42):

/bəˈnænə/	ALIGN-*ó*-L	*SONONS/*ŏ*	*V̆	MAX	CONTIG
a. bəˈnɑnʌ	*!	*	**		
b. ˈbɑnʌ		*	*	**	*!
☞ c. ˈnɑnʌ		*	*	**	

While the effects of MAX/*ó* can be subsumed under two other MARKEDNESS constraints, an adequate MARKEDNESS correlate of ANCHOR-R cannot be found. Since ANCHOR-R penalizes output forms where the rightmost input syllable is no longer rightmost in the output (either by deletion, epenthesis, or metathesis), a conceivable MARKEDNESS constraint that could have the same effect would be a constraint that penalizes syllables that are not rightmost in the word. One such constraint is defined in (38).

(38) RIGHTMOST: Syllables are only licensed if they are rightmost in the word.[10]

The issue with the constraint defined above is that it is a MARKEDNESS constraint – it can only penalize output forms, without reference to input forms. It cannot decide between candidates where a choice needs to be made about which stressless syllable to retain, such as in the adult form /ˈænəməl/, 'animal'. No matter which input unstressed syllable is retained, it will always be rightmost in the output form, as shown in the tableau in (39).

(39) RIGHTMOST **fails to decide on a candidate**, based on data from 38m3 (Kehoe, 2000, p.43):

/ˈænəməl/	ALIGN-*ó*-L	RIGHTMOST	*SONONS/*ŏ*	*V̆	MAX
a. ˈænʌmɑ		**!	**	**	*
☞ b. ˈænʌ		*	*	*	***
☞ c. ˈæmɑ		*	*	*	***

However, once ANCHOR-R is introduced into the tableau – even when lowly ranked – it *will* be able to decide between the relevant candidates, as shown in (40).

(40) ANCHOR-R **decides on the correct output**, based on data from 38m3 (Kehoe, 2000, p.43):

/ˈænəməl/	ALIGN-*ó*-L	RIGHTMOST	*SONONS/*ŏ*	*V̆	MAX	ANCHOR-R
a. ˈænʌmɑ		**!	**	**	*	
b. ˈænʌ		*	*	*	***	*!
☞ c. ˈæmɑ		*	*	*	***	

[10] While this is a direct analogue of ANCHOR-R, this is not the most natural constraint, as it requires all outputs to be a single syllable in length. A somewhat more reasonable constraint could penalize having unstressed syllables in any position other than rightmost in the word, *ŏ/¬RIGHT (Thanks to Elliott Moreton for suggesting this constraint to me). However, since both constraints will behave similarly with respect to output forms, I will continue to use RIGHTMOST for the remainder of this analysis.

Since RIGHTMOST does not contribute to the analysis of English children's grammars, it will be left out of all remaining tableaux.

While not all positional FAITHFULNESS constraints have a (useful) corresponding MARKEDNESS constraint, those that do not can still have effects while remaining low-ranked in children's grammars. This means that these positional FAITHFULNESS constraints do not need to be ranked above any of the relevant MARKEDNESS constraints in order to correctly predict the children's output forms. If this is the case, they should not pose any problems for a gradual OT-based learning algorithm.

In the following sections, an intermediate markedness analysis will be outlined for the remaining intermediate stages of English unstressed syllable acquisition, and a full learning path will be outlined. It will then be demonstrated that the CGLA is able to follow this learning trajectory, given a certain set of data to work from.

5.2 English unstressed syllable acquisition as intermediate markedness stages

The first stage outlined in Kehoe (2000) is one where children preserve all stressed syllables and right-edge unstressed syllables. All other syllables, however, are deleted. An analysis of this stage that relies primarily on appropriate MARKEDNESS constraints has been proposed in section 5.1, and is repeated in (41).

(41) **Ranking of constraints for the first English intermediate markedness stage**:
ALIGN-$\acute{\sigma}$-L, *SONONS/$\breve{\sigma}$, *$\breve{\text{V}}$ ≫ MAX, ANCHOR-R

This stage, for the constraints involved, is equivalent to the initial stage of the grammar, where all MARKEDNESS constraints outrank all FAITHFULNESS constraints (IDENT[stress] and *CLASH excepted, as will be the case for the remainder of this paper). It does not seem to correspond to an intermediate markedness stage, per se, unlike Kehoe's analysis, where the first stage observed falls under the characterization of an intermediate faithfulness stage.

The second stage in unstressed syllable acquisition is one where unstressed medial syllables with obstruent onsets have been acquired, but where unstressed initial syllables and medial syllables with sonorous onsets are still deleted. This indicates that *$\breve{\text{V}}$ has been demoted below MAX, but that *SONONS/$\breve{\sigma}$ and ALIGN-$\acute{\sigma}$-L still dominate MAX. A ranking and sample tableaux are given in (42).

(42) **Ranking of constraints for the second English intermediate markedness stage**:
ALIGN-$\acute{\sigma}$-L, *SONONS/$\breve{\sigma}$ ≫ MAX ≫ *$\breve{\text{V}}$, ANCHOR-R

a. *Tableau for* /ˈɑktəpəs/ → [ˈaʔtəpʊs], based on data from 34f3 (Kehoe, 2000, p.43):

/ˈɑktəpəs/	ALIGN-$\acute{\sigma}$-L	*SONONS/$\breve{\sigma}$	MAX	*$\breve{\text{V}}$	ANCHOR-R
☞ a. ˈaʔtəpʊs				**	
b. ˈaʔpʊs			*!*	*	
c. ˈaʔtə			*!**	*	*

b. *Tableau for* /ˈtɛləˌfo͡ʊn/ → [ˈtɛrˌfon], based on data from 34f3 (Kehoe, 2000, p.43):

/ˈtɛləˌfo͡ʊn/	ALIGN-σ́-L	*SONONS/σ̆	MAX	*V̆	ANCHOR-R
a. ˈtɛrəˌfon		*!		*	
☞ b. ˈtɛrˌfon			*		

This stage, unlike the first stage examined, meets the criteria for being an intermediate markedness stage. There is a positional MARKEDNESS constraint, *SONONS/σ̆, which outranks a general FAITHFULNESS constraint, MAX, which in turn outranks a general MARKEDNESS constraint *V̆. Similarly, since ALIGN-σ́-L has the effect of ruling out initial unstressed syllables, it might also be considered a "positional" variant of *V̆, and therefore also meets the criteria for contributing to an intermediate markedness analysis.

The third stage is one where initial unstressed syllables are able to be produced, but where children still have difficulty with unstressed medial syllables with sonorant onsets. This indicates that the children have re-ranked ALIGN-σ́-L to be below MAX, but that *SONONS/σ̆ remains ranked above MAX. A ranking and sample tableaux are given in (43). This stage retains one of the rankings that leads to an intermediate markedness stage, *SONONS/σ̆ ≫ MAX ≫ *V̆, but has eliminated the ranking ALIGN-σ́-L ≫ MAX ≫ *V̆.

(43) **Ranking of constraints for the third English intermediate markedness stage**:
*SONONS/σ̆ ≫ MAX ≫ ALIGN-σ́-L, *V̆, ANCHOR-R

a. *Tableau for* /təˈme͡ɪɹo/ → [təˈmedo], based on data from 34f1 (Kehoe, 2000, p.44):

/təˈme͡ɪɹo/	*SONONS/σ̆	MAX	ALIGN-σ́-L	*V̆	ANCHOR-R
☞ a. təˈmedo			*	**	
b. ˈmedo		*!*		*	

b. *Tableau for* /ˈtɛləˌfo͡ʊn/ → [ˈtɛlfo], based on data from 34f1 (Kehoe, 2000, p.44):

/ˈtɛləˌfo͡ʊn/	*SONONS/σ̆	MAX	ALIGN-σ́-L	*V̆	ANCHOR-R
a. ˈtɛləfo	*!	*		**	
☞ b. ˈtɛlfo		**		*	

The final stage is one where the children have successfully acquired unstressed syllables in all positions – they are adult-like in their production of unstressed syllables. This indicates that all relevant MARKEDNESS constraints have been re-ranked below MAX, as illustrated in the ranking in (44).

(44) **Ranking of constraints for the final stage in English unstressed syllable acquisition**:
MAX ≫ *SONONS/σ̆, ALIGN-σ́-L, *V̆, ANCHOR-R

This section has attempted to show that the constraints outlined in section 5.1 are sufficient to characterize each of the intermediate stages observed in Kehoe (2000) by re-ranking MAX with respect to *V̆, ALIGN-σ́-L, and *SONONS/σ̆, respectively. The full expected learning path is outlined in (45).

(45) **Learning path for English children in terms of positional markedness**:

First stage grammar: *SONONS/$\check{\sigma}$, ALIGN-$\acute{\sigma}$-L, *$\check{\text{V}}$ ≫ MAX, ANCHOR-R

Second stage grammar: *SONONS/$\check{\sigma}$, ALIGN-$\acute{\sigma}$-L ≫ MAX ≫ *$\check{\text{V}}$, ANCHOR-R

Third stage grammar: *SONONS/$\check{\sigma}$ ≫ MAX ≫ ALIGN-$\acute{\sigma}$-L, *$\check{\text{V}}$, ANCHOR-R

Final stage grammar: MAX ≫ *SONONS/$\check{\sigma}$, ALIGN-$\acute{\sigma}$-L, *$\check{\text{V}}$, ANCHOR-R

In the following subsection, it will be shown that a CGLA learner is capable of following this learning path, under certain assumptions about the frequency of errors produced. Discussion of what is predicted using different assumptions about the frequency of errors will be undertaken in the final subsection.

5.3 The CGLA is able to follow the attested intermediate markedness learning path

It will be assumed that the learning algorithm is allowed to begin at the first stage grammar, since this is a grammar where all of the relevant MARKEDNESS constraints outrank all of the relevant FAITHFULNESS constraints. As such, the CGLA will begin with the initial ranking vector listed in (46).

(46) **Initial ranking vector for English**:

ALIGN-$\acute{\sigma}$-L	*SONONS/$\check{\sigma}$	*$\check{\text{V}}$	MAX	ANCHOR-R
(12,	12,	12,	0,	0)

When the learner uses this ranking to predict a form for the adult English word /ˈɑktəpəs/, 'octopus', it will instead produce the form [ˈɑʔpʊs], as produced by participant 34m3. When this error is compared to the adult form, it will produce the update shown in (47). This error will provide evidence that the undominated error-preferring constraint *$\check{\text{V}}$ should be demoted, while MAX should be promoted. ANCHOR-R will not be promoted, but will remain at the bottom of the hierarchy.

(47) **Update vector generated by** /ˈɑktəpəs/ → [ˈɑʔpʊs]:

/ˈɑktəpəs/	ALIGN-$\acute{\sigma}$-L	*SONONS/$\check{\sigma}$	*$\check{\text{V}}$	MAX	ANCHOR-R
a. ˈɑʔtʌpʊs			**!		
b. ˈɑʔtʌp			*	**	*!
☞ c. ˈɑʔpʊs	(*e*)	(*e*)	* (L)	** (W)	(*e*)

→ $(0, 0, -1, \frac{1}{2}, 0)$

Similarly, if the learner uses this ranking to predict a form for the adult English word /bəˈnænə/, 'banana', it will instead produce the form [ˈnɑnʌ], again, as produced by 34m3. Comparison of this error to the adult form will produce the update vector shown in (48). In addition to providing evidence that *$\check{\text{V}}$ should be demoted, this error will provide evidence that ALIGN-$\acute{\sigma}$-L should

be demoted as well. The same update vector will be produced when the learner predicts a form for the adult word /təˈmẽɪɾo/, 'tomato', producing [ˈmedo]. Since the update vector produced will be the same, only one of the relevant tableaux is shown below.

(48) **Update vector generated by** /bəˈnænə/ → [ˈnɑnʌ] **and** /təˈmẽɪɾo/ → [ˈmedo]:

/bəˈnænə/	ALIGN-*σ́*-L	*SONONS/*σ̆*	*V̆	MAX	ANCHOR-R
a. bəˈnɑnʌ	*!	*	**		
☞ b. ˈnɑnʌ	(L)	* (*e*)	* (L)	** (W)	(*e*)

→ (−1,0,−1,1,0)

Finally, say that the learner uses this ranking to predict a form for the adult English word /ˈɛləfənt/, 'elephant', producing 34m3's error [ˈʔʌfɪnt] instead. The comparative tableau and update vector associated with this error is given in (49). This error will provide evidence that *V̆ should be demoted, as the other errors do, and it also provides evidence that *SONONS/*σ̆* should be demoted.

(49) **Update vector generated by** /ˈɛləfənt/ → [ˈʔʌfɪnt]:

/ˈɛləfənt/	ALIGN-*σ́*-L	*SONONS/*σ̆*	*V̆	MAX	ANCHOR-R
a. ˈʔʌlʌfɪnt		*!	**		
b. ˈʔʌlʌf		*!	*	***	*
☞ c. ˈʔʌfɪnt	(*e*)	(L)	* (L)	** (W)	(*e*)

→ (0,−1,−1,1,0)

The CGLA will cycle through the four errors described above: [ˈɑʔpʊs], [ˈnɑnʌ], [ˈmedo], and [ˈʔʌfɪnt]. It will then successfully arrive at the second stage of English unstressed syllable acquisition, as re-analyzed in section 5.2 and shown in (50).

(50) **Learning path for English: stage 1 to stage 2:**

a. *Learning steps triggered by* [ˈɑʔpʊs], [ˈnɑnʌ], [ˈmedo], *and* [ˈʔʌfɪnt]:

Error		ALIGN-*σ́*-L	*SONONS/*σ̆*	*V̆	MAX	ANCHOR		*Grammar*
	(	**12**,	**12**,	**12**,	**0**,	**0**	)	Initial
ˈɑʔpʊs	(	12+0=**12**,	12+0=**12**,	12−1=**11**,	$0+\frac{1}{2}=\mathbf{\frac{1}{2}}$,	0+0=**0**	)	①
ˈnɑnʌ	(	12−1=**11**,	12+0=**12**,	11−1=**10**,	$\frac{1}{2}+1=\mathbf{1\frac{1}{2}}$,	0+0=**0**	)	②
ˈmedo	(	11−1=**10**,	12+0=**12**,	10−1=**9**,	$1\frac{1}{2}+1=\mathbf{2\frac{1}{2}}$,	0+0=**0**	)	
ˈʔʌfɪnt	(	10+0=**10**,	12−1=**11**,	9−1=**8**,	$2\frac{1}{2}+1=\mathbf{3\frac{1}{2}}$,	0+0=**0**	)	
ˈɑʔpʊs	(	10+0=**10**,	11+0=**11**,	8−1=**7**,	$3\frac{1}{2}+\frac{1}{2}=\mathbf{4}$,	0+0=**0**	)	
ˈnɑnʌ	(	10−1=**9**,	11+0=**11**,	7−1=**6**,	4+1=**5**,	0+0=**0**	)	
ˈmedo	(	9−1=**8**,	11+0=**11**,	6−1=**5**,	5+1=**6**,	0+0=**0**	)	③

b. *Learning path followed*:

Initial grammar: ALIGN-$\acute{\sigma}$-L, *SONONS/$\breve{\sigma}$, *$\breve{\text{V}}$ ≫ MAX, ANCHOR-R

↓

Grammar (1): ALIGN-$\acute{\sigma}$-L, *SONONS/$\breve{\sigma}$ ≫ *$\breve{\text{V}}$ ≫ MAX ≫ ANCHOR-R

↓

Grammar (2): *SONONS/$\breve{\sigma}$ ≫ ALIGN-$\acute{\sigma}$-L ≫ *$\breve{\text{V}}$ ≫ MAX ≫ ANCHOR-R

↓

Grammar (3): *SONONS/$\breve{\sigma}$ ≫ ALIGN-$\acute{\sigma}$-L ≫ MAX ≫ *$\breve{\text{V}}$ ≫ ANCHOR-R

At the end of the learning path outlined above, the CGLA learner has correctly re-ranked *$\breve{\text{V}}$ below MAX, arriving at a grammar that is compatible with the second stage grammar outlined in section 5.2. It should be noted that while ANCHOR-R can be used by the grammar to correctly predict which syllable in this stage is selected for retention, *it will never be promoted*, since both the fully-faithful adult form and the child's error will obey this constraint, and will receive an *e* in the comparative tableaux. While this constraint does not move in the hierarchy, its effects are still observed in what the child is able to produce.

At this second stage, words such as adult /ˈɑktəpəs/, 'octopus', will be faithfully mapped to [ˈaʔtəpʊs], as illustrated by the productions of 34f3 and 28m3 in (30). However, forms such as /təˈm͡eɪɾo/, 'tomato', and /bəˈnænə/, 'banana', will still be mapped unfaithfully to output forms [ˈmedo] and [ˈbænʌ], respectively. This is due to the ranking of ALIGN-$\acute{\sigma}$-L above MAX. As such, this will continue to provide evidence that ALIGN-$\acute{\sigma}$-L should be demoted, while MAX should be promoted.

(51) **Update vector generated by** /bəˈnænə/ → [ˈbænʌ] **and** /təˈm͡eɪɾo/ → [ˈmedo]:

/bəˈnænə/	*SONONS/$\breve{\sigma}$	ALIGN-$\acute{\sigma}$-L	MAX	*$\breve{\text{V}}$	ANCHOR-R
a. bəˈnænʌ	*	*!		**	
☞ b. ˈbænʌ	* (*e*)	(L)	** (W)	* (L)	(*e*)

→ $(0,-1,\frac{1}{2},0,0)$

Similarly, this ranking will predict that adult forms such as /ˈɛləfənt/, 'elephant', should still be mapped to errors such as [ˈɑfɪn], since *SONONS/$\breve{\sigma}$ is still ranked above MAX. As such, this error will provide evidence that *SONONS/$\breve{\sigma}$ should continue to be demoted, while MAX should continue to be promoted.

(52) **Update vector generated by** /ˈɛləfənt/ → [ˈɑfɪn]:

/ˈɛləfənt/	*SONONS/$\breve{\sigma}$	ALIGN-$\acute{\sigma}$-L	MAX	*$\breve{\text{V}}$	ANCHOR-R
a. ˈɑlʌfɪn	*!		*	**	
b. ˈɑlʌf	*!		***	*	*
☞ c. ˈɑfɪn	(L)	(*e*)	*** (W)	* (L)	(*e*)

→ $(-1,0,\frac{1}{2},0,0)$

A CGLA learner beginning from the second stage grammar and cycling through the above three errors will successfully arrive at the third stage of English unstressed syllable acquisition.

This is shown in the learning diagram in (53). (It will be assumed that if there is an even ranking between constraints, then an error will still be produced on the next learning step.)

(53) **Learning path for English: stage 2 to stage 3:**

a. *Learning steps triggered by* [ˈbænʌ], [ˈmedo], *and* [ˈɑfɪn]:

Error		*SonOns/σ́	Align-σ́-L	Max	*V̌	Anchor		*Grammar*
	(	**11**,	**8**,	**6**,	**5**,	**0**	)	③
ˈɑfɪn	(	11 − 1 = **10**,	8 + 0 = **8**,	6 + ½ = **6½**,	5 + 0 = **5**,	0 + 0 = **0**	)	
ˈbænʌ	(	10 + 0 = **10**,	8 − 1 = **7**,	6½ + ½ = **7**,	5 + 0 = **5**,	0 + 0 = **0**	)	④
ˈmedo	(	10 + 0 = **10**,	7 − 1 = **6**,	7 + ½ = **7½**,	5 + 0 = **5**,	0 + 0 = **0**	)	⑤

b. *Learning path followed*:

Grammar ③: *SonOns/σ́ ≫ Align-σ́-L ≫ Max ≫ *V̌ ≫ Anchor-R

↓

Grammar ④: *SonOns/σ́ ≫ Align-σ́-L, Max ≫ *V̌ ≫ Anchor-R

↓

Grammar ⑤: *SonOns/σ́ ≫ Max ≫ Align-σ́-L ≫ *V̌ ≫ Anchor-R

At this point, the ranking will predict that both /bəˈnænə/ and /təˈmeɪ͡ɹo/ should be produced faithfully, but that /ˈɛləfənt/ should still be missing its medial syllable. This corresponds with the third stage ranking, as outlined in 5.2. Thus, the CGLA has correctly moved from the second stage observed by Kehoe into the third.

From this stage, the only error that will be predicted is that /ˈɛləfənt/, 'elephant', will be mapped to the error [ˈɛlːfɪnt], as is done by participant 34f1. The only remaining undominated loser-preferring constraint is *SonOns/σ́, and the update vector generated by the third stage grammar will be identical to the one generated for this error during the second stage grammar. The CGLA learner will only have to produce this error two more times in order to arrive at the proper adult grammar, as shown by the diagram in (54).

(54) **Learning path for English: stage 3 to stage 4**:

a. *Learning steps triggered by* [ˈɛlːfɪnt]:

Error		*SonOns/σ́	Max	Align-σ́-L	*V̌	Anchor		*Grammar*
	(	**10**,	**7½**,	**6**,	**5**,	**0**	)	⑤
ˈɛlːfɪnt	(	10 − 1 = **9**,	7½ + ½ = **8**,	6 + 0 = **6**,	5 + 0 = **5**,	0 + 0 = **0**	)	
ˈɛlːfɪnt	(	9 − 1 = **8**,	8 + ½ = **8½**,	6 + 0 = **6**,	5 + 0 = **5**,	0 + 0 = **0**	)	⑥

b. *Learning path followed*:

Grammar ⑤: *SonOns/σ́ ≫ Max ≫ Align-σ́-L ≫ *V̌ ≫ Anchor-R

↓

Grammar ⑥: Max ≫ *SonOns/σ́ ≫ Align-σ́-L ≫ *V̌ ≫ Anchor-R

Thus, it has been shown that a CGLA learner beginning from the first stage observed for English unstressed syllable acquisition can go through all of the stages attested by Kehoe (2000), provided that the learner is given an appropriate subset of the data to learn from and that the constraints used to characterize each stage are cast in terms of MARKEDNESS whenever possible. It has also been shown that even though reference to positional FAITHFULNESS constraints is sometimes inevitable, as in the case of ANCHOR-R, these constraints can remain low-ranked in the hierarchy and still correctly select the error produced by the child. In the following subsection, some discussion of the issues with this analysis are detailed, before a more general discussion is undertaken.

5.4 Local discussion

5.4.1 Learning path discrepancies

It should be noted that some very specific errors were selected when running the CGLA – namely, there were two adult forms with an initial unstressed syllable, /bəˈnænə/, 'banana', and /təˈm͡eɪɾo/, 'tomato', while only one each of an adult form that had a medial unstressed syllable with an obstruent onset (/ˈɑktəpəs/, 'octopus') and a sonorant onset (/ˈɛləfənt/, 'elephant'). This does not match the distribution of the words presented by Kehoe (2000), which is skewed towards words that have either unstressed medial or unstressed final syllables. If a CGLA learner were to run on the set of forms presented by Kehoe, it would be predicted that the first constraint to be demoted below MAX would be *V̆, as above, but that the second constraint to be demoted would be *SONONS/$\breve{\sigma}$, contrary to the learning path outlined in (45). This illustrates that the path taken by the CGLA, as a gradual OT learner, is heavily dependent upon not only the quality of the errors it posits, but the quantity of errors posited as well. Depending upon which words are attempted, the CGLA learner may not always predict the acquisition order of Kehoe (2000).

To show this more concretely, a series of simulations similar to those performed for Québec French was also performed using the English constraint set outlined in (41) and the inputs /ˈɛləfənt/ and /bəˈnænə/. For each simulation, the step at which each input form was predicted to be produced faithfully by the constraint ranking was recorded, in addition to the step at which the input form /ˈɑktəpəs/ was predicted to be produced faithfully. The results are shown in figure 2 (following page).

Similarly to the behaviour of /ˈtʁɛ̃/-like input in Québec French, the behaviour of /ˈɑktəpəs/-like words is constant across all input frequencies, since all update vectors calculated by the learner will provide evidence that *V̆ needs to be demoted and that MAX needs to be promoted. It should also be noted that the behaviour of the other two errors is similar to the behaviour of /gʁyˈjo/-like input in Québec French – the rate at which they are learned is dependent upon their frequency as input forms.

The right-hand side of the graph in figure 2 is consistent with the learning trajectory predicted by Kehoe (2000). However, the left-hand side of the graph predicts that there should also be a stage in which children are able to produce all unstressed syllables *except initial unstressed syllables*. This is not a stage observed by Kehoe, and furthermore, given the frequency of /bəˈnænə/-like words in her study, it is actually the stage that would be predicted if the learner

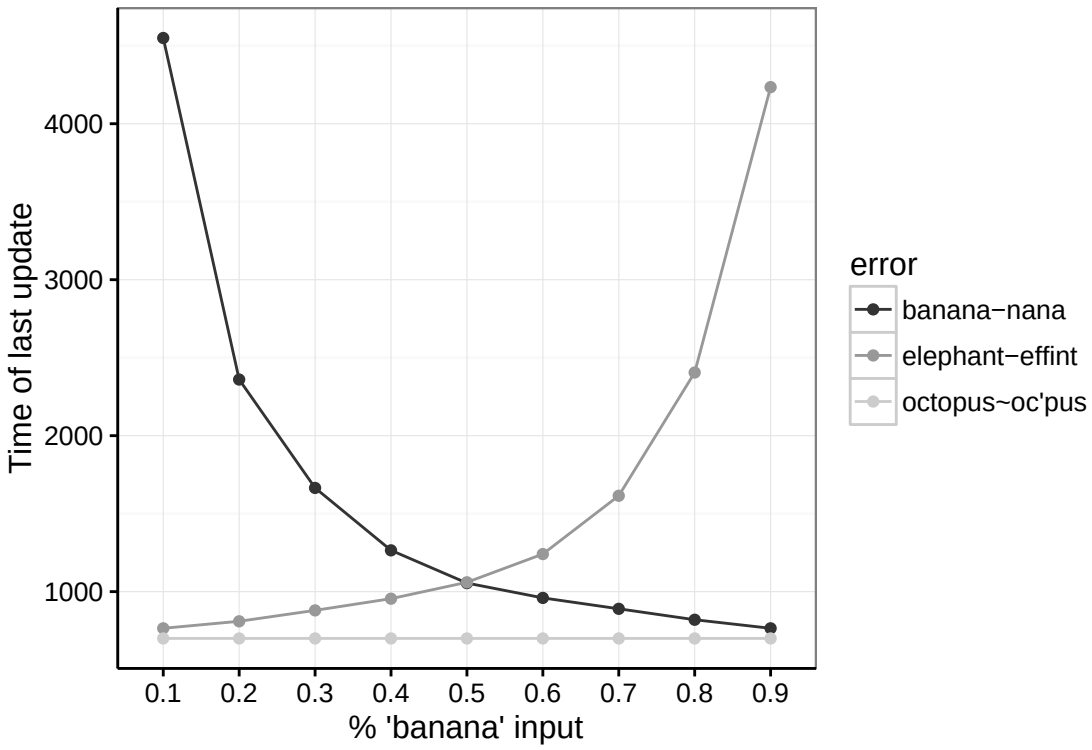

Figure 2: **Length of intermediate stages for English**, as measured by the percentage of /bəˈnænə/-like input at the time (learning step) at which an error was no longer produced

were given all of those words as inputs.

However, the set of words selected for inclusion in Kehoe's study are unusual in that they contain many longer words that children might be unlikely to attempt to produce or hear often in their environment. That is, they were selected precisely because they would be unlikely errors. It may be the case that the number and kinds of errors attempted by the children outside of the lab setting have a different distribution, such that they are more likely to hear or attempt to produce words with initial unstressed syllables rather than words with medial unstressed syllables with sonorant onsets.

It should also be noted that Kehoe's study was not a longitudinal study, as was observed by Kehoe herself (Kehoe, 2000, pp.62–3). Rather, children were grouped into stages based on what syllables they produced, and it was observed that these groups formed a series of subset relations, such that the set of syllables preserved in the first stage is a subset of those preserved in the second stage, and so on (Kehoe, 2000, p.53). It is therefore possible that individual children *did not go through this exact learning path*, but instead skipped stages or went through them in different orders.

5.4.2 Choice of constraints

This particular analysis of the phenomenon of unstressed syllable acquisition in English relies heavily on the idea that *SONONS/$\breve{\sigma}$ is a constraint that can be considered to be a "positional variant" of *V̆. However, it appears at first glance to be an arbitrary constraint, and should ideally find some justification in the phonetic properties of English or other languages. As a first pass, it might be said that this constraint is part of some hierarchy of constraints, such that sonorants – especially those in unstressed syllables – prefer to be in the nucleus or coda of a syllable, rather than in the onset. However, a full-scale phonetic study of this phenomenon is currently outside the scope of this paper.

This particular analysis also relies on the idea that ANCHOR-R is active in child grammars, and crucially, that its counterpart ANCHOR-L is ranked below it in the hierarchy. While this is not reflected in the ranking vectors provided in section 5.3, it must nevertheless be true, since there are no instances of children preferring to preserve initial unstressed syllables over final ones. However, if this constraint were to be included in CON, it would be expected that there would be children who are learning North American English who *do* prefer to preserve initial unstressed syllables. The fact that this pattern is not attested could be due the fact that in child language, ANCHOR-R is universally ranked above ANCHOR-L. This, however, appears to be undesirable, since it is often ANCHOR-L that is preferentially ranked high in adult grammars. If the ranking is universal for children, why might the opposite ranking be perferred in adult grammars?

An alternate approach to this issue lies in the particular constraint that is used to preserve right-edge syllables. If this constraint is not ANCHOR-R, but is instead some constraint that preferentially preserves syllables that are longer in duration than normal – without inherent reference to word edges – then the same behaviour can be adequately modelled without having to posit a ranking between ANCHOR-R and ANCHOR-L. For instance, say that there is some constraint MAX-V/[long], which preserves input vowels that are longer in duration.[11] If the child's input is exactly the same as the adult's output – a standard assumption made during learning – then those vowels that are long will be those that are prosodically prominent, either by virtue of being stressed, or by virtue of being at the right edge of a prosodic constituent. If the child has recourse to MAX-V/[long], they will preferentially select those vowels that are long in the adult's output – namely, the stressed and right-edge syllables, consistent with the first stage in unstressed syllable acquisition. Similarly to ANCHOR-R, this constraint may remain ranked at the bottom of the constraint hierarchy and still select the appropriate syllables for inclusion in the output. This constraint, while apparently adequate to replace ANCHOR-R in the analysis above, was not chosen for inclusion for two reasons. First, since Kehoe's initial analysis made use of ANCHOR-R, inclusion of ANCHOR-R makes the alternate analysis more similar to the one that was previously entertained. Second, since MAX-V/[long] must make an arbitrary distinction between what durations count as "long" and which count as "short", children may interpret this constraint differently than adults. That is, they may draw the dividing line along the duration continuum in a different place, which would lead to much more variable behaviour with respect to MAX-V/[long].

[11] Thanks to Sam Zukoff for suggesting this particular constraint to me.

While there are some outstanding questions about this particular analysis of North American english, this section has shown that these children's acquisition of unstressed syllables can be successfully predicted by a gradual OT learner, provided that their analysis can be cast in terms of positional MARKEDNESS constraints, and that a particular subset of errors is produced. Recourse to positional FAITHFULNESS constraints has been shown to be necessary in the case of ANCHOR-R, but it has also been shown that ANCHOR-R (or a similar constraint, MAX-V/[long]) can remain low-ranked and still predict the attested errors. Since it is not necessary that any stage meets the criteria for inclusion as an intermediate faithfulness stage, as was originally hypothesized by Kehoe (2000), this set of data is no longer problematic for gradual OT learners. In the following section, discussion of this general finding will be undertaken.

6 Discussion

While the majority of this paper has been focused on demonstrating what conditions on the OT constraint set are necessary for children's learning trajectories to be unproblematic for gradual OT learners, the analysis proposed does raise many questions about the consequences of these conditions. What might this mean for child grammars and children's learning trajectories in general? What does this mean for the analysis of adult grammars that make use of positional FAITHFULNESS constraints? What consequences might this have for use of HG-based learners? What other approaches are possible? These questions will be addressed in turn in the following sections.

6.1 Predictions for children's grammars

One main consequence of the analysis proposed above is that intermediate markedness stages should be fairly well-attested in children's learning trajectories, as all MARKEDNESS constraints will begin highly-ranked and their effects will be seen until a FAITHFULNESS constraint outranks them. Since positional MARKEDNESS constraints will drop in the hierarchy much slower than their general counterparts, their effects should be seen for much longer in child grammars in general.

However, the literature on intermediate stages (characterized as intermediate faithfulness stages) is fairly limited – only a few children ever show direct evidence of a proposed intermediate faithfulness stage (Rose, 2000), while much of the other patterns are inferred (Fikkert, 1994; Kehoe and Stoel-Gammon, 1997; Kehoe, 2000). What might account for this apparent lack of intermediate stages in most other children?

There are three potential answers to this question. First, it may be the case that other phenomena, while not classified as intermediate faithfulness stages, can still be characterized under the umbrella of intermediate markedness. These phenomena may already be characterized in terms of high-ranking MARKEDNESS constraints, positional or otherwise, and can thus be counted as intermediate markedness stages. Second, it may also be the case that the intermediate stage occurs before the child is necessarily producing words. Children's perception has been shown to be more adult-like than their production at any one stage in development

(Hayes, 2004; Fikkert, 2007; Fikkert and de Hoop, 2009, and references therein), so it may be that there are children who go through such a stage when learning to perceive words, rather than when they are learning to produce words. Finally, as outlined in sections 4.3 and 5.4.1, it may be that the intermediate stage only lasts a very small amount of time, so that the researcher does not have a chance to document it.

The above analysis has also made the assumption that the set of constraints that are present in children's grammars is equivalent to the set of constraints present in adult grammars. However, this analysis is not incompatible with the addition of child-specific constraints into CON, provided that those constraints are MARKEDNESS constraints rather than FAITHFULNESS constraints. Proposed child-specific constraints, such as McAlister Byun's (2012) MOVE-AS-UNIT constraints, seem to fit this generalization, since they are cast as MARKEDNESS constraints rather than FAITHFULNESS constraints. They should therefore pose no issues for gradual OT learners if included in CON, and are compatible with the above approach.

Finally, while the above analysis has suggested a re-analysis of intermediate faithfulness stages that have been posited for Québec French (Rose, 2000) and American English (Kehoe and Stoel-Gammon, 1997; Kehoe, 2000), it remains to be seen whether such an analysis can be proposed for every intermediate faithfulness stage in the literature. One case of particular interest is the case of Hebrew word truncation, as observed by Adam (2002) and Bat-El (2007), and analyzed as an intermediate faithfulness stage by Jesney and Tessier (2008).

Children learning Hebrew, like the children learning English discussed earlier, go through a series of intermediate stages when acquiring unstressed syllables. The first of these stages has been documented by Adam (2002), and is referred to as the trochaic stage. In this stage, children produce only word-final syllables if those syllables are stressed in the adult form, or only the final two syllables, if the penultimate syllable is stressed.

(55) **Hebrew truncation stage**, data from Adam (2002), as presented in Jesney and Tessier (2008):

Target	*Child*	*Gloss*
/ˈsefer/	[ˈfefe]	'book'
/ʃarˈʃeret/	[ˈetet]	'necklace'
/laˈredet/	[ˈdedet]	'to get down'
/lifˈtoax/	[ˈgoax]	'to open'
/kaˈdur/	[ˈdur]	'ball'
/saˈbon/	[ˈbon]	'soap'
/lisˈgor/	[ˈgor]	'to close'
/neʃɾˈka/	[ˈka]	'kiss' (n.)

Under the analysis proposed by Jesney and Tessier (2008), this stage is consistent with an initial stage where all MARKEDNESS constraints (in their analysis, ALIGN-$\acute{\sigma}$-L as calculated over feet and TROCHEE) outrank all FAITHFULNESS constraints (in their analysis, MAX).

The next stage observed in the sequence is one where children are able to produce a single unstressed syllable to the left of a word-final stressed syllable, but only in nouns – verbs still adhere to the truncation pattern outlined in (55) (Bat-El, 2007).

(56) **Hebrew intermediate stage**, data from Bat-El (2007), as presented in Jesney and Tessier (2008):[12]

Target	*Child*	*Gloss*
/bakˈbuk/	[bukˈbuk]	'bottle'
/arˈje/	[aˈje]	'lion'
/sufganˈd͡ʒa/	[inˈd͡ʒa]	'donut'
/livjaˈtan/	[iˈtan]	'whale'
/maʃˈpic/	[ˈpic]	'squirt'
/hoˈrid/	[ˈjit]	'put down'
/koˈfec/	[ˈfec]	'jump'
/naˈfal/	[ˈfal]	'fell'

According to Jesney and Tessier (2008), this stage is properly characterized as an intermediate faithfulness stage, and must make reference to a positional FAITHFULNESS constraint MAX/NOUN, which ensures that any material in nouns is preferentially retained. In an OT grammar, this constraint must be ranked above the constraint TROCHEE, which in turn must outrank the more general MAX.

The final stage outlined in this portion of the learning path is one where verbs are also allowed to contain two syllables, regardless of where the stress in those two syllables falls. Some data from Adam (2002) are provided below.

(57) **Hebrew disyllabic stage**, data from Adam (2002), as presented in Jesney and Tessier (2008):

Target	*Child*	*Gloss*
/ˈsefer/	[ˈsefer]	'book'
/ʃarˈʃeret/	[ˈetet]	'necklace'
/laˈredet/	[ˈdedet]	'to get down'
/laˈkaxat/	[ˈkaxat]	'to take'
/kaˈdur/	[aˈdur]	'ball'
/xaˈtul/	[aˈtul]	'cat'
/liʃˈtot/	[iˈtot]	'to drink'
/neʃɪˈka/	[ʃiˈka]	'kiss' (n.)

Under an OT analysis, this stage is characterized as one where both MAX/NOUN and MAX outrank TROCHEE, but do not outrank ALIGN-$\acute{\sigma}$-L.

The difficulty in translating this analysis to being one of intermediate markedness lies in the fact that there does not seem to be any positional MARKEDNESS constraint parallel to MAX/NOUN. However, while there may not be a direct MARKEDNESS parallel, there may nevertheless be a MARKEDNESS constraint or set of MARKEDNESS constraints that collectively may yield the same results, as was proposed for the analysis of English, above. One such analysis involves a limit on utterance length. If whole utterances at the intermediate stage outlined in (56) are limited

[12] Jesney and Tessier (2008) list the fourth data point with the gloss 'crocodile', but I have been told by two native speakers of Hebrew that a more appropriate gloss is 'whale', so I have glossed it as such.

to two syllables, and if utterances can consist of either a single noun or a noun and a verb, then the data from the intermediate stage are expected. Of course, this should be examined with reference to the data in Adam (2002) and Bat-El (2007) in much more detail to see how plausible this analysis is, but it does bring the at first puzzling Hebrew data under the umbrella of intermediate markedness.

6.2 The status of positional faithfulness in adult grammars

The resulting grammar for the children examined in this paper and for the learning algorithms run is one where the more marked adult grammar is successfully acquired. All MARKEDNESS constraints banning marked structures have been demoted below the relevant FAITHFULNESS constraints. However, there are conceivably adult languages that look like the children's intermediate stages. For example, Goad and Rose (2004) claim that Brazilian Portuguese displays the same pattern of complex onset licensing as the intermediate stage in Québec French, outlined above. If this is the case, should they be analyzed in terms of positional FAITHFULNESS, as done by Goad and Rose (2004), or should they be analyzed in terms of positional MARKEDNESS?

If the analysis pursued above is correct, it would suggest that they should be analyzed in terms of positional MARKEDNESS. Assuming that Goad and Rose's (2004) analysis of Brazilian Portuguese is correct, complex onsets are only observed in stressed syllables. A child learning Brazilian Portuguese will therefore never receive any evidence that a constraint like *COMPLEX/$\breve{\sigma}$ should be demoted, and it should be the constraint that is responsible for the limitation on where complex onsets are licensed. This then raises the question of whether the positional FAITHFULNESS counterpart of *COMPLEX/$\breve{\sigma}$, MAX/$\acute{\sigma}$, is necessary at all. It may be the case that positional FAITHFULNESS constraints can be eliminated from the grammar altogether, if a MARKEDNESS counterpart can be found.

Some research indicates that this should not be the case. First, according to Jesney (2011), both positional FAITHFULNESS constraints and positional MARKEDNESS constraints are necessary in OT in order to adequately capture the typology of marked structure limitations in adult languages, provided that conjunctive and disjunctive constraints are not allowed in CON. Second, a positional FAITHFULNESS constraint whose behaviour can be paralleled by a positional MARKEDNESS constraint in one environment can prove to be a valuable tiebreaker in another.

As an example, there could be a child that acquires voiced obstruents in an intermediate stage fashion.[13] In the first stage, this child is unable to produce voiced obstruents, and converts them all to voiceless obstruents. In the intermediate stage, this child is able to produce voiced obstruents *only in onset position*; coda obstruents are still devoiced. Finally, the child is able to faithfully produce voiced obstruents in all positions. This learning path is given using the schema in (58).

[13] I am indebted to Giorgio Magri for the following two examples.

(58) **Learning path for child acquiring obstruent voicing**:

Initial stage grammar:	/za/ → [sa]; /Caz/ → [Cas]
	↓
Intermediate stage grammar:	/za/ → [za]; /Caz/ → [Cas]
	↓
Final stage grammar:	/za/ → [za]; /Caz/ → [Caz]

In this example, the constraints *[+voice] and IDENT[voice] are in no doubt active, since in the initial stage grammar *[+voice] outranks IDENT[voice], and in the final stage grammar this ranking is switched. However, the intermediate stage could in principle be due to two constraints being highly ranked in the grammar – the positional MARKEDNESS constraint *[+voice]/CODA, or the positional FAITHFULNESS constraint IDENT[voice]/__V. Under the account proposed above, the intermediate stage behaviour should be due to *[+voice]/CODA. Since there are two constraints that could in theory do the same work in the grammar, and since one will, by virtue of being a MARKEDNESS constraint, be more active in the grammar, it may be beneficial to eliminate the unnecessary FAITHFULNESS constraint altogether.

However, as another example, say that there is a child who is attempting to learn a language that allows heterogenic coda-onset clusters. In the initial stage of their language, they simplify each cluster by changing the voicing of the coda consonant to match that of the onset consonant. Individual coda consonants remain faithful to the voicing of the adult coda consonant, indicating that there is no universal ban on coda consonants being [+voice]. There are no cases where these children preserve the voicing of the coda consonant and change the voicing of the onset consonant. From this stage, they then progress directly to one that is faithful to the adult inputs. This is shown in (59).

(59) **Learning path for child acquiring heterogenic coda-onset clusters**:

Initial stage grammar:	/azka/ → [aska], *[azga]; /asga/ → [azga], *[aska] /az/ → [az]; /ak/ → [ak]
	↓
Final stage grammar:	/azka/ → [azka]; /asga/ → [asga] /az/ → [az]; /ak/ → [ak]

Under this analysis, the initial stage grammar is one where a constraint such as AGREE[voice] is active, and causes adjacent consonants to agree in their value for the feature [voice]. However, the child always chooses to resolve any violations of this constraint by preserving the [voice] feature of the onset, regardless of what that place feature is. The faithful mappings /az/ → [az] and /ak/ → [ak] indicate that the child has correctly learned that coda consonants must retain their [voice] features, so the MARKEDNESS constraint *[+voice]/CODA should not be highly-ranked in the child's grammar. It thus appears that the constraint that is responsible for the child's selection of onset consonant place over coda consonant place is due to a positional FAITHFULNESS constraint, IDENT[voice]/__V.

Since it appears that IDENT[voice]/__V is necessary for describing at least some grammars, it should not be eliminated entirely from the constraints available in CON, and should instead be retained.

6.3 Comparison with HG-based learners

This paper has demonstrated that, contrary to what has been asserted by Jesney and Tessier (2007, 2008) and Tessier (2009), gradual OT learners are indeed capable of modelling intermediate faithfulness stages in the course of phonological acquisition under standard MARKEDNESS ≫ FAITHFULNESS learning assumptions. As such, their existence should not be used as an argument in favour of one algorithm over another, or as an argument to modify existing algorithms. This is not to claim that gradual OT learning algorithms do not need to be modified for other reasons, but simply that this particular phenomenon in child phonological acquisition is not an adequate reason to do so.

However, this does not mean that other learners are incapable of modelling such intermediate stages. As demonstrated in Jesney and Tessier (2007, 2008), an HG-based gradual learner is perfectly capable of modelling these stages while making reference only to positional and general FAITHFULNESS constraints. In addition, it may be the case that HG-based learners are able to model these stages while including only positional FAITHFULNESS or only positional MARKEDNESS constraints in CON, without needing recourse to both varieties (Jesney, 2011).

While both kinds of algorithms are able to model the phenomena discussed here, they do make different predictions about what specific rankings of constraints are possible when modelling intermediate stages. Under the assumption that all FAITHFULNESS constraints begin at the bottom of the hierarchy, the analysis proposed in this paper predicts that there should never be any stages in phonological acquisition that are true intermediate faithfulness stages, where the ranking positional FAITHFULNESS ≫ MARKEDNESS ≫ general FAITHFULNESS is enforced. Since both FAITHFULNESS constraints will begin with a ranking value of 0, and since the general FAITHFULNESS constraint will always receive at least as much evidence for its promotion as its positional counterpart, there should *never* be a situation where the positional variant must be ranked higher than the general variant. Thus, it is almost always the positional MARKEDNESS constraints that are responsible for intermediate stage behaviour, and positional FAITHFULNESS constraints, when active, must be ranked lower in the hierarchy than previously assumed.

However, this is not the case for an HG-based learner. Although the starting assumptions are the same, where all FAITHFULNESS constraints begin at 0 and all MARKEDNESS constraints begin higher-ranked, it is still possible for an HG-based learner to arrive at a true intermediate faithfulness stage grammar where both the FAITHFULNESS constraints are ranked low and no reference to a positional MARKEDNESS constraint is made. All that is required is that the weight of both FAITHFULNESS constraints, when considered individually, do not exceed the weight of the general MARKEDNESS constraint, while their combined weight does exceed the weight of the MARKEDNESS constraint. Thus, the positional and general FAITHFULNESS constraints can both remain low-ranked, and cause the grammar to behave as if it had the strict ranking of positional FAITHFULNESS over general MARKEDNESS over general FAITHFULNESS Jesney and Tessier (2007, 2008).

An HG-based learner is therefore able to model a true intermediate faithfulness stage without recourse to positional MARKEDNESS constraints. If a stage in acquisition can be found that must irrevocably be analyzed as an intermediate faithfulness stage, then it would provide some evidence in favour of the HG-based learner. As of the present time, I do not know of such a

stage, but I do not believe that it can be ruled out entirely.

6.4 Alternate approaches

It should be stressed that while the analysis proposed in this paper is one particular way of modelling children's intermediate stages using gradual OT learners, it is by no means the only method of doing so. This account has been based around standard assumptions about the initial stage of the child's grammar, where all MARKEDNESS constraints begin highly-ranked and all FAITHFULNESS constraints begin ranked as a whole at the bottom of the grammar. As noted by Tessier (2009, p.17), however, if the usual assumptions about how the FAITHFULNESS constraints are initially ranked are dispensed with, gradual OT learners will be able to model these stages as intermediate faithfulness stages. For instance, if all of the positional FAITHFULNESS constraints are initially much higher ranked than their general counterparts (but not as highly ranked as the MARKEDNESS constraints), then it will be more likely that the MARKEDNESS constraints will be ranked below the positional FAITHFULNESS constraints sooner than their general counterparts, even though they will not rise in the hierarchy as quickly.

Intermediate faithfulness stages can also be appropriately modelled using gradual OT learners if it is assumed that, rather than containing positional FAITHFULNESS constraints and their general counterparts, CON contains instead only positional FAITHFULNESS constraints, with a fixed ranking between them, in accordance with the PMap of Steriade (2002). As an example, in Québec French, the appropriate ranking that describes the intermediate faithfulness stage would be as in (60), where MAX/$\breve{\sigma}$, a constraint that enforces the retention of any material in unstressed syllables, replaces MAX.

(60) **Ranking for the intermediate stage of Québec French under the PMap**:
MAX/$\acute{\sigma}$ ≫ *COMPLEX ≫ MAX/$\breve{\sigma}$

If a fixed ranking can be determined between the constraints MAX/$\acute{\sigma}$ and MAX/$\breve{\sigma}$, such that MAX/$\acute{\sigma}$ ≫ MAX/$\breve{\sigma}$, then it will always be the case that MAX/$\acute{\sigma}$ will be the first to cross any MARKEDNESS constraints, and an intermediate faithfulness stage can thus be reached. A brief consideration of how the PMap functions shows that such a ranking is plausible. The difference between a segment being present in a stressed syllable vs. it being absent is presumably larger than the difference between a segment be present vs. being absent in an unstressed syllable. Since segments overall tend to be longer in stressed syllables, deleting such a long segment is presumably more salient than deleting a shorter one. This perceptual difference can then be used to infer a universal ranking such that MAX/$\acute{\sigma}$ will always outrank MAX/$\breve{\sigma}$.

While this is an attractive solution, it is nevertheless still problematic for gradual OT learners, as these learning algorithms are currently unable to preserve innate hierarchical rankings of constraints. Until these algorithms can be modified to preserve such hierarchies, this particular solution will not be tenable.

7 Conclusion

This paper has examined the phenomenon of intermediate stages in child language acquisition, whereby children are able to produce a particular marked structure only in a subset of the environments where it is found in the adult language. Within the early literature on these stages, it was established that these intermediate stages should be characterized by ranking certain positional FAITHFULNESS constraints high in the children's grammar, such as MAX/$\acute{\sigma}$ in Québec French (Rose, 2000; Goad and Rose, 2004; Jesney and Tessier, 2007, 2008; Tessier, 2009), or MAX/$\acute{\sigma}$ and ANCHOR-R in English (Kehoe and Stoel-Gammon, 1997; Kehoe, 2000). Jesney and Tessier (2007, 2008) and Tessier (2009) generalized the schema for characterizing these stages, showing that an intermediate stage should arise whenever a positional FAITHFULNESS constraint outranks a general MARKEDNESS constraint, which in turn outranks a general version of that same FAITHFULNESS constraint. Jesney and Tessier (2007, 2008) and Tessier (2009) have, however, also shown that gradual OT learners as they are formulated will never arrive at this constraint ranking, as they will always promote the general FAITHFULNESS constraint more than the positional FAITHFULNESS constraint. They use this fact to argue for the adoption of an HG-based learning algorithm over an OT-based one (or for adopting a gradual version of an existing OT-based algorithm, in the case of Tessier (2009)), since the mechanism for selecting the optimal candidate in HG allows for the two kinds of FAITHFULNESS constraints to gang up on the MARKEDNESS constraint in the appropriate contexts.

This paper has offered another solution to this issue, and shows that gradual OT learners are indeed capable of modelling the intermediate stages outlined in the literature, provided that they are able to use positional MARKEDNESS constraints as well as positional FAITHFULNESS constraints. If allowed to do so, intermediate stages can be characterized as intermediate markedness stages, where a positional MARKEDNESS constraint is ranked above a FAITHFULNESS constraint, which is in turn ranked above a more general MARKEDNESS constraint. Furthermore, using an intermediate markedness analysis with gradual OT learners works for the very reason that using intermediate faithfulness accounts do not – since general MARKEDNESS constraints will always be demoted more than their positional counterparts, they will be outranked by FAITHFULNESS constraints much sooner, leaving the positional MARKEDNESS constraints highly ranked and active in the children's grammar. This was shown to correctly model the acquisition path for the Clara and Théo, two children acquiring complex onsets in Québec French. As such, the fact that these intermediate stages exist in the acquisition of marked structures for children should not be used as an argument against the use of gradual OT learners.

The analysis provided in this paper was then tested on some additional data from the acquisition of unstressed syllables in English (Kehoe and Stoel-Gammon, 1997; Kehoe, 2000). It was shown that while positional FAITHFULNESS constraints cannot be entirely eliminated from the grammar – as was the case for ANCHOR-R – they can remain low-ranked in the hierarchy of constraints and still select the observed output. Furthermore, allowing them to do so will not disrupt the course of learning in any significant way, which was shown by the successful running of a gradual OT learner on the English data.

This analysis has, however, opened up other questions for future research. If it is a viable solution, it should also be checked if adult languages that appear to be consistent with chil-

dren's intermediate stages can also follow the same generalizations. That is, is it always the case that a suitable positional MARKEDNESS constraint can be used to achieve the same results as a positional FAITHFULNESS constraint? If not, can it be shown that this positional FAITHFULNESS constraint can remain low-ranked in the grammar and still do work?

In addition, due to the sparseness of the data from phonological acquisition, it may also be the case that the test cases for demonstrating a difference between positional FAITHFULNESS and positional MARKEDNESS accounts are missing. There may be intermediate stages that must be analyzed as intermediate faithfulness stages, and cannot be analyzed as intermediate markedness stages. If these stages are shown to exist, it will provide evidence that the above account is incorrect, and that some alteration to the learning algorithms used must be undertaken. It may be the case that an HG-based gradual learner should be adopted, as advocated for in Jesney and Tessier (2007, 2008), or that some modification should be made to the gradual OT learner in order to make it compatible with a PMap analysis, as outlined in section 6.4. The options available are numerous, and should be considered in more detail before coming to a decision on the usefulness of available learning algorithms.

References

Adam, Galit. 2002. From variable to optimal grammar: Evidence fom language acquisition and language change. Doctoral Dissertation, Tel-Aviv University, Tel-Aviv.

Bat-El, Outi. 2007. Category-specific phonology in adult and child hebrew. Handout from OCP4, Rhodes.

Beckman, Jill. 1998. Positional faithfulness. Doctoral Dissertation, Universy of Massachusetts, Amherst.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21:43–58.

Echols, Catharine H., and Elissa L. Newport. 1992. The role of stress and position in determining first words. *Language Acquisition* 2:189–220.

Fikkert, Paula. 1994. On the acquisition of prosodic structure. Ph.d. dissertation, University of Leiden.

Fikkert, Paula. 2007. Acquiring phonology. In *Handbook of phonological theory*, ed. P. de Lacy, chapter 23, 537–554. Cambridge, MA: Cambridge University Press.

Fikkert, Paula, and Helen de Hoop. 2009. Language acquisition in optimality theory. *Linguistics* 47:311–357.

Gerken, LouAnn. 1996. Prosodic structure in young children's language production. *Language* 72:683–712.

Gnanadesikan, Amalia. 2004. Markedness and faithfulness constraints in child phonology. In *Constraints in phonological acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, chapter 3, 73–108. Cambridge: Cambridge University Press.

Goad, Heather, and Yvan Rose. 2004. Input elaboration, head faithfulness, and evidence for representation in the acquisition of left-edge clusters in west germanic. In *Constraints in phonological acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, chapter 4, 109–157. Cambridge: Cambridge University Press.

Gordon, Matthew. 2002. A factorial typology of quantity-insensitive stress. *Natural Language & Linguistic Theory* 20:491–552.

Hayes, Bruce. 2004. Phonological acquisiton in optimality theory: the early stages. In *Constraints in phonological acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, chapter 5, 158–203. Cambridge: Cambridge University Press.

Ito, Junko. 1986. Syllable theory in prosodic phonology. Doctoral Dissertation, University of Massachusetts, Amherst.

Jesney, Karen. 2011. Cumulative constraint interaction in phonological acquisition and typology. Doctoral Dissertation, University of Massachusetts, Amherst.

Jesney, Karen, and Anne-Michelle Tessier. 2007. Re-evaluating learning biases in harmonic grammar. *University of Massachusetts Occasional Papers in Linguistics* 36:69–110.

Jesney, Karen, and Anne-Michelle Tessier. 2008. Gradual learning and faithfulness: Consequences of ranked vs. weighted constraints. In *Proceedings of NELS 28*, ed. Muhammad Abdurrahman, Anisa Schardl, and Martin Walkow. Amherst: GLSA publications.

Kehoe, Margaret M. 2000. Truncation without shape constraints: the latter stages of prosodic acquisition. *Language Acquisition* 8:23–67.

Kehoe, Margaret M., and C. Stoel-Gammon. 1997. Truncation patterns in english-speaking children's word productions. *Journal of Speech, Language, and Hearing Research* 40:526–541.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic grammar – a formal multi-level connectionist theory of linguistic wellformedness: an application. In *Proceedings of the Twelfth Annual Conference on the Cognitive Science Society*, 884–89. Cambridge, MA: Lawrence Erlbaum.

Magri, Giorgio. 2012. Convergence of error-driven ranking algorithms. *Phonology* 29:213–269.

McAlister Byun, Tara. 2012. Positional velar fronting: An updated articulatory account. *Journal of Child Language* 39:1043–1976.

Pater, Joe. 1997. Minimal violation and phonological development. *Language Acquisition* 6:201–253.

Pater, Joe. 2008. Gradual learning and convergence. *Linguistic Inquiry* 39:334–345.

Prince, Alan. 2002. Entailed ranking arguments. Ms, Rutgers University. Available as ROA-500 from the Rutgers Optimality Archive.

Prince, Alan S., and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical Report 2, Rutgers University Center for Cognitive Science.

Prince, Alan S., and Bruce Tesar. 2004. Learning phonotactic distributions. In *Constraints in phonological acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, chapter 8, 245–291. Cambridge: Cambridge University Press.

Rose, Yvan. 2000. Headedness and prosodic licensing in the l1 acquisition of phonology. Doctoral Dissertation, McGill University, Montréal, Québec.

Smolensky, Paul. 1996. The initial state and 'richness of the base' in optimality theory. Ms, Johns Hopkins University. Available as ROA-154 from the Rutgers Optimality Archive.

Smolensky, Paul, and Géraldine Legendre. 2006. *The harmonic mind: From neural computation to optimality-theoretic grammar*. Cambridge, MA: MIT Press.

Steriade, Donca. 2002. The phonology of perceptibility effects: the p-map and its consequences for constraint organization. Ms., November 2002.

Tesar, Bruce, and Paul Smolensky. 2000. *Learnability in optimality theory*. Cambridge, MA: MIT Press.

Tessier, Anne-Michelle. 2009. Frequency of violation and constraint-based phonological learning. *Lingua* 119:6–38.