Overapplication Opacity in Phonological Acquisition

Olga Urek
CASTL/University of Tromsø

Abstract

Phonological opacity is a challenge for parallel OT, which does not allow for intermediate levels of representation. Several modifications of the theory have been proposed over the years to incorporate opacity, all of them falling short of accounting for spontaneous opacity effects in developing grammars. In this paper I demonstrate that if certain independently motivated adjustments are made to the recent OT-based theory of opacity called *Optimality Theory with Candidate Chains* (OT-CC, see McCarthy 2007), it can successfully deal with spontaneous opacity effects.

1. Opacity in Optimality Theory

The term 'phonological opacity' refers to cases where a surface form of some language L appears to have exceptionally undergone (overapplication or counterbleeding opacity) or failed to undergo (underapplication or counterfeeding opacity) a certain phonological process active in language L.

(1)　　*Opacity (Kiparsky 1973)*
　　　　A phonological rule P of the form A → B / C_D is opaque if there are surface structures with either of the following characteristics:
　　　　a. instances of A in the environment C_D.
　　　　b. instances of B derived by P that occur in environments other than C_D.

Serialist accounts of opacity crucially rely on the existence of some intermediate form that deviates from the input and from the actual output. According to Prince & Smolensky (1993:6), in 'classic' Optimality Theory (OT) the "Input → Output map has no internal structure: all possible variants are produced by Gen in one step and evaluated in parallel". Precisely this property makes opacity a challenging issue for OT.

　　　There have been many attempts to incorporate opacity in OT, which can be roughly subdivided into two large (and partly overlapping) groups: those that rely on expanding some basic assumptions about the nature of underlying representations or the constraint component Con and those that make reference to forms distinct from the input and the actual output. Virtually any of those approaches is associated with a number of problems, which sometimes prove fatal under closer examination. The issue of acquisition is not the least of these problems (but see Bermúdez-Otero 2003 on the acquisition of target-like opacity in Stratal OT).

　　　Apart from being typologically adequate, any theory of grammar must be demonstrated to be learnable by means of an effective computable algorithm. Since OT was proposed in 1993, a number of learning algorithms have been put forward and claimed to solve the problem. The most widely acclaimed and thoroughly tested is, perhaps, the family of algorithms based on *constraint demotion* (CD; see Tesar & Smolensky 1993, Tesar 1995, Prince & Tesar 2004, Boersma 2008). In the course of time, CD algorithms have been shown to successfully account for the acquisition of transparent generalizations. Moreover, theoretical claims concerning the acquisition progression have been substantiated through empirical research.

　　　Unfortunately, OT-based solutions to the opacity problem are not always compatible with CD algorithms. The matter is further complicated by the fact that opaque generalizations are not limited to

fully-developed grammars: there is an extensive amount of early production data proving that non-target-like opacity effects spontaneously arise in developing grammars and are lost at the later stages of acquisition. Thus, acquisition and opacity intersect in more than one point: first, one has to account for the acquisition of target-like opacity effects; second, one has to deal with the spontaneous emergence and subsequent loss of non-target-like opaque generalizations in the course of acquisition.

Reranking of phonological constraints in the course of acquisition is generally assumed to happen in response to positive evidence. In the case of target-like opacity such positive evidence is readily available in the form of adult opaque utterances, while in the case of non-target-like opacity positive evidence is absent. If a theory of phonology postulates some special mechanism to deal with opacity, we would expect learners to employ this mechanism in the first case (under the pressure of positive evidence), but not in the second case (because non-target-like opaque forms are not obligatory in order to converge on the adult-like grammar). Non-target-like opaque forms are not the goal of acquisition process. Rather, they are epiphenomenal to the acquisition of transparent generalizations, and as such they are especially hard to model. In this paper I demonstrate that if certain independently motivated adjustments are made to the recent OT-based theory of opacity called *Optimality Theory with Candidate Chains* (OT-CC, see McCarthy 2007), it can successfully deal with spontaneous opacity effects[1].

This paper is organized as follows: in Section 2 I briefly introduce OT-CC; in Section 3 I review some evidence in favour of discarding the ranking metaconstraint; in Section 4, I propose some modifications to the original OT-CC formulation; in Section 5, I provide the analysis of spontaneous counterbleeding opacity with modified OT-CC; Section 6 contains discussion and conclusions.

2. Optimality Theory with Candidate Chains

Optimality Theory with Candidate Chains, or OT-CC, was proposed by John McCarthy (2007) and differs from classic OT primarily in its understanding of what constitutes an output candidate. In OT-CC, a candidate is a chain of forms, rather than a single form. Candidate chains are subject to a number of restrictions. First, chains should be faithfully initiated, meaning that the first member of every chain is a fully faithful parse of the input form. Second, forms in the chain should accumulate differences from the input gradually. That is, every successive candidate should add a single violation of a basic faithfulness constraint in a specific location in a form (localized unfaithful mapping or LUM). Third, candidate chains should improve harmonically, meaning that every successive form in the chain should be more harmonic than the preceding one relative to the given constraint hierarchy. In other words, every successive form in the chain should fare better on markedness constraints than its predecessor.

To be evaluated by the grammar, candidate chains are reduced to four crucial bits of information: input, output, L-set, which is the list of LUMs incurred by all forms of the chain, and LUMSeq, which is a sequentially ordered list of LUMs of each specific chain. The sequence of LUMs is projected from the chain. For chains consisting of only a faithful parse, L-set and LUMSeq are empty. For chains involving one faithfulness violation, L-set contains one LUM, while LUMSeq is empty, since no sequential ordering is possible.

In some cases, valid candidate chains constructed on the basis of some input form /in/ have identical output forms, but differ in the order of LUMs. According to McCarthy (2007:96), such situation arises every time there is no interaction among phonological processes. In OT-CC, such chains are referred to as 'convergent chains'. According to McCarthy (2007:96), "convergent chains are collapsed into a single candidate by merging their LUMSeqs, retaining all and only the LUM precedence relations

---

[1] Although OT-CC has been ambitiously claimed to be "the best theory of opacity - and of phonology generally" (see McCarthy 2007:3), a number of criticisms of the theory were proposed. Thus, Kavitskaya and Staroverov (2010) note that OT-CC in its original formulation cannot deal with so-called 'fed counterfeeding', where the two processes are both in feeding (transparent) and counterfeeding (opaque) relations. Kavitskaya and Staroverov (2010) propose a number of modifications to OT-CC that solve the problem. In turn, Kaplan (2011) presents Chamorro umlaut and Central Venetan metaphony and argues that OT-CC is incapable of accounting for these phenomena in a satisfactory way.

that are common to the set of convergent chains". The list of such precedence relations in the form of pairwise orders is referred to as a 'reduced LUMSeq' or just rLUMSeq. For non-convergent chains LUMSeq and rLUMSeq are identical.

According to McCarthy (2007:97), output is evaluated by markedness constraints, input-output relations encoded in L-set are evaluated by faithfulness constraints, while "evaluation of rLUMSeqs is the responsibility of **Prec** constraints, which favor certain precedence relations among the constituent LUMs of an rLUMSeq and penalize others". Following McCarthy (2007:98), Prec(A,B) (where A and B are faithfulness constraints), "demands that every B-violating LUM be preceded and not followed by an A-violating LUM in the rLUMSeq". Prec(A,B) can assign violation marks in two cases: first, "if there is a B-violating LUM in L-set, and this LUM is not preceded in the rLUMSeq by some A-violating LUM"; second, "if there is a B-violating LUM in L-set and it precedes some A-violating LUM".

Ranking of Precedence constraints is subject to certain restrictions. According to McCarthy (2007:98), "[a]lthough Prec constraints [...] are ranked and violable, their ranking is not entirely free". McCarthy proceeds with saying that "though Prec(A,B) obviously *depends on* whether B is violated, [it] must never *affect* whether B is violated" and "in OT, the only way to ensure this is to require that Prec(A,B) never dominate B"(italics are original). To this end, McCarthy(2007) postulates the ranking meta-constraint in the form B >> Prec(A,B).

3. Discarding the Metaconstraint

As any assumed universal ranking, the metaconstraint is a theory-internal stipulation, and as such it is undesirable. In this section I will first briefly introduce the reason for the postulation of the metaconstraint on the ranking of Prec, and then review some evidence in favour of discarding the metaconstraint.

3.1 *Reasoning behind the ranking metaconstraint*

According to McCarthy (2007), the reason for the metaconstraint comes from the analysis of counterbleeding interaction of palatalization and high-vowel syncope in Bedouin Arabic.

(2) *Velar palatalization in Bedouin Arabic (McCarthy 2007:99)*
    a. [ru:g]   'be calm'   [rawwig^j]   'do not make noise!'
    b. [gu:l]   'say'   [g^{j}i:l]   'it was said'

(3) *Syncope of /i/ in Bedouin Arabic (McCarthy 2007:187)*
    a. /kitib-at/   ['kitbat]   'it (m.) was written'
    b. /ʃarib-at/   ['ʃarbat]   'she drank'

(4) *Opaque interaction of palatalization and syncope (McCarthy 2007:100)*
    a. /ħa:kim-i:n/   [ħa:k^{j}mi:n]   'ruling (m. pl.)'
    b. /kitib-t/   [k^{j}tibt]   'you (m.sg.) were written'

According to McCarthy (2007:99), in Bedouin Arabic the velar stops /k/ and /g/ are palatalized to [k^j] and [g^j] when adjacent to the front vowel [i] as in (1), while short high vowels are deleted from non-final open syllables as in (3). Interaction of syncope and palatalization leads to counterbleeding opacity, where the velar palatalizes even when the high vowel conditioning the palatalization is deleted as in (4). To deal with these phenomena, McCarthy (2007) proposes the following constraints:

(5)   *Palatalization and syncope constraints (McCarthy 2007: 93-94)*
     a. *iCV (dominating Max)
        'Short high vowels are prohibited in open syllables'
     b. *ki (dominating Ident(back))
        'Sequences of a plain velar and a front vowel are prohibited'
     c. Prec(Ident(back), Max)
        'Max-violating LUM should be preceded and should not be followed by Ident(back)
        violating LUM'

Harmonically improving candidate chains for the input /ħa:kim-i:n/ constructed on the basis of the transparent ranking of the Bedouin Arabic are given below (note, that Prec constraints do not participate in the construction of candidate chains).

(6)   *Harmonically improving candidate chains for the input /ħa:kim-i:n/ and their LUMSeqs (McCarthy 2007:100)*
     a. <ħa:kimi:n>   <>
     b. <ħa:kimi:n, ħa:kʲimi:n>   <Ident(back)>
     c. <ħa:kimi:n, ħa:k.mi:n >   <Max>
     d. ☞ < ħa:kimi:n, ħa:kʲimi:n, ħa:kʲ.mi:n >   <Ident(back), Max>

(7)   *Harmonically improving candidate chains for the input /ʃarib-at/ and their LUMSeqs*
     a. <ʃa.ri.bat>   <>
     b. ☞ <ʃar.bat>   <Max>

(8)   *Tableau 1: Palatalization and syncope in Bedouin Arabic (adapted from McCarthy 2007:101)*

		*iCV	*ki	MAX	PREC(IDENT(back), MAX)	IDENT(back)
1.	**/ħa:kim-i:n/**					
a.	ħa:.ki.mi:n	*!	*			
b.	ħa:.kʲi.mi:n	*!				*
c.	ħa:k.mi:n			*	*!	
☞d.	ħa:kʲ.mi:n			*		*
2.	**/ʃarib-at/**					
a.	ʃa.ri.bat	*!				
☞b.	ʃar.bat			*	*	

In Tableau 1, candidates (1a) and (1b) are eliminated by the high-ranked markedness constraint against high vowels in open syllables. Candidates (1c) and (1d) tie on the faithfulness constraint Max, but (1c) incurs a fatal violation of the Precedence constraint, because Max-violating LUM in its LUMSeq is not preceded by Ident(back)-violating LUM. Thus, opaque candidate (1d) is correctly selected as optimal. In case of transparent application of syncope, the faithful candidate in (2a) is eliminated by the high-ranked *iCV constraint. This makes (2b) the winner, regardless its violation of Prec(Ident(back), Max).

     In OT, every legal permutation of constraints is supposed to yield an existing grammar. Therefore, if we discard the metaconstraint, the grammar where Prec(Ident(back), Max) dominates *iCV becomes theoretically possible. McCarthy (2007:101) stresses that ranking Prec(Ident(back), Max) above *iCV would yield a grammar where syncope is discouraged unless it counterbleeds palatalization (that is, front high vowel would only delete if it triggered palatalization). Such hypothetical situation is illustrated below.

(9)   *Tableau 2: Effect of violating the ranking meta-constraint (adapted from McCarthy 2007:102)*

/ʃarib-at/	PREC(IDENT(back),MAX)	*ki	*iCV	MAX	IDENT(back)
☞a. ʃa.ri.bat			*		
☹ b. ʃar.bat	*!			*	

In Tableau 2, candidate (b) incurs a fatal violation of Prec(Ident(back),Max) because Max-violating LUM in its LUMSeq is not preceded by Ident(back)-violating LUM. According to McCarthy (ibid), the state of affairs illustrated in Tableau 2 "goes beyond anything encountered in attested opaque phonology" and "no real phonological system does or could work this way, so we need a language-independent explanation for why this never happens". In order to rule out the unwanted scenario McCarthy proposes to postulate meta-constraint on the ranking of Prec that would ensure that Prec (Ident(back), Max) can never dominate Max.

3.2   *Obligatorily counterbleeding processes*

Matthew Wolf (2008:350), however, argues that the scenario modelled in Tableau 2 is, in fact, attested. He presents a case of Chimwi:ni (Swahili:Somalia), where vowel shortening is blocked when there is no spirantization for it to counterbleed (see also Hyman 1993). Wolf calls the phenomenon of Chimwi:ni "an obligatorily counterbleeding process".

(10)   *Obligatorily counterbleeding processes (Wolf 2008:350)*
        A B-violating process occurs just in case it would counterbleed an A-violating process. If no A-violating process occurs, the B-violating process does not occur.

According to Wolf (2008:350-351), in Chimwi:ni the perfective suffix /-i:ɬe/ triggers spirantization of certain preceding consonants as in (11). When spirantization occurs, the initial vowel of the suffix shortens as in (12). However, the suffix vowel fails to shorten when it is not preceded by consonant that undergoes spirantization as in (13). Some stems with final consonants other than [p, t, ʈ, k, ɬ] exceptionally undergo spirantization, and in such cases shortening of the suffix vowel does occur as in (14).

(11)   *Chimwi:ni spirantization (Wolf 2008:351)*
        a.   /p, t, ʈ/   →   [s]
        b.   /k/   →   [ʃ]
        c.   /ɬ/   →   [z]

(12)   *Chimwi:ni vowel shortening (Wolf 2008:351)*
        a.   [ku-ɬipa]   'to pay'   /ɬip-i:ɬe/   →   [ɬis-iɬe]   'he paid'
        b.   [ku-laʈa]   'to let go'   /laʈ-i:ɬe/   →   [las-iɬe]   'he let go'
        c.   [x-ʃi:ka]   'to hold'   /ʃi:k-i:ɬe/   →   [ʃi:ʃiɬe]   'he held'
        d.   [x-kuɬa]   'to grow'   /kuɬ-i:ɬe/   →   [kuziɬe]   'he grew'

(13)   *Vowel fails to shorten if not preceded by a spirantized consonant (Wolf 2008:351)*
        a.   /pamb-i:ɬe/   →   [pamb-i:ɬe]   'he decorated'
        b.   /kos-i:ɬe/   →   [kos-e:ze]   'he made a mistake'
        c.   /set-i:ɬe/   →   [set-e:ɬe]   'he stamped on'

(14)   *Exceptional spirantization (Wolf 2008:351)*
        a.   /big-i:ɬe/   →   [biš-iɬe]   'he hit'
        b.   /law-i:ɬe/   →   [laz-iɬe]   'he went out'

On the basis of these data, Wolf (ibid) concludes the following: first, assuming that the length of the suffix vowel triggers the spirantization of the preceding consonant, Chimwi:ni presents a case of

counterbleeding interaction between spirantization and shortening, since shortening removes the context for spirantization, but spirantization occurs nevertheless; second, shortening of the perfective suffix vowel is blocked just in case there is no spirantization for it to counterbleed. Therefore, Chimwi:ni exhibits exactly the kind of scenario McCarthy (2007) wanted to rule out by means of meta-constraint on the ranking of Prec. The existence of a natural language violating the metaconstraint is, probably, the most important argument in favour of discarding the metaconstraint.

3.3   The Metaconstraint and Spontaneous Opacity

Another reason for discarding a metaconstraint on ranking comes from child acquisition data, namely, from cases of spontaneous, or non-target-like, opacity. Following Wolf (2008), below we will illustrate this claim with the example of spontaneous counterfeeding.

Smith (1973) describes an interesting case of non-target-like chain-shift, whereby velarization of /t, d/ before laterals counterfeeds context-independent occlusivization of stridents to [t, d].

(15)   *Spontaneous chain-shift (Smith 1973)*
   a. /pʌzəl/   →   [pʌdəl] (* → [pʌgəl])
   b. /pʌdəl/   →   [pʌgəl]
   c. /pɪkəl/   →   [pɪkəl]

As it is evident from the data in (15), underlying velars surface faithfully, and underlying /d/ is velarized to [g]. In turn, /d/ derived by occlusivization does not undergo velarization. In light of rule-based approaches, in this case velarization does not apply to derived /d/ because its target is introduced too late in the derivation. Therefore, we are looking at the case of counterfeeding-on-focus. In the tableau below it is demonstrated how OT-CC deals with this case (where *TL is a constraint that penalizes coronal-lateral sequences, i.e. a constraint that prefers velarization):

(16)   *Tableau 3: Spontaneous chain-shift in OT-CC (adapted from Wolf 2008:355)*

	/pʌzəl/	*z	IDENT(contin)	PREC(IDENT(place), IDENT(contin))	*TL	IDENT(place)
a.	< pʌzəl > <Ø>	*!			*	
☞b.	< pʌzəl , pʌdəl > <ID(contin)>		*	*	*	
c.	< pʌzəl , pʌdəl , pʌgəl > <ID(contin), ID(place)>		*	**!		*

The metaconstraint on the ranking of Prec requires that Prec(Ident(place), Ident(contin)) should be ranked below Ident(contin). In order for Prec to favour the opaque candidate, it must also be ranked above its worst violation, i.e. above *TL. The candidate chain with velarization incurs two violations on Prec constraint, because it contains an Ident(contin)-violation that is not preceded by Ident(place)-violation, and is followed by Ident(place)-violation. The above ranking yields the correct output, but it is also conceptually problematic.

McCarthy (2007:119) says that "if Prec constraints are universally present in grammars, then it is reasonable to assume that they are initially ranked at the bottom of the hierarchy, below even the faithfulness constraints". Such ranking would also ensure that the metaconstraint is observed. If we assume that at the initial state Prec constraints occupy the lowest stratum in the hierarchy, then according to the provisions of BCD algorithm (Biased Constraint Demotion, see Prince & Tesar 2004), Prec(Ident(place), Ident(contin)) can only end up dominating *TL in response to the positive evidence. In

our case, the learner clearly has no access to such positive evidence, since neither occlusivization of stridents nor velarization of stops are active in the target language (i.e. English).

Therefore, if we want this analysis to hold, the only way is to assume that the learners have an innate bias for ranking Prec constraints high. The idea of an innate ranking bias is nothing new: innate domination of Markedness over Faithfulness is assumed in order to rule out excessively permissive grammars. Developing this idea, Wolf (2008:355) suggests that in order to accommodate child production data, Prec constraints should be ranked above markedness constraints in the initial state. By transitivity it would also mean that Prec constraints are ranked above faithfulness constraints, which contradicts the metaconstraint on ranking. In fact, as we will see below, it would suffice to say that Prec constraints are ranked above Faithfulness constraints in the initial state. Prec can come to dominate Markedness in the course of acquisition due to the markedness demotion in response to the positive evidence. In any case, ranking Prec constraints above Faithfulness violates the metaconstraint.

According to Wolf(2008:355), there exists an alternative strategy for dealing with developmental chain-shifts, i.e. positional faithfulness proposed by Jesney (2005). According to Jesney(2005:59), the immunity of target /z/ to the general "d to g process" triggered by *TL is the consequence of "preferential feature preservation", enforced by special Identity constraints. That is, input /z/ is said to contain a particularly well-formed feature combination [coronal, +strident]. Privileged relationship that holds between coronality and stridency has the effect that the feature [coronal] is preserved only when it is associated with [+strident] feature in the input.

(17)   *Preferential preservation of [coronal] on input [+strident] segments (from Jesney 2005:63)*

puzzle-puddle-pickle

/s, z, ʃ, tʃ, dʒ/   /t, d, n/   /k, g, ŋ/
[**Cor**, +stri]   [Cor, -stri]   [Dor, -stri]

[t, d]   [k, g, ŋ]
[**Cor**, -stri]   [Dor, -stri]
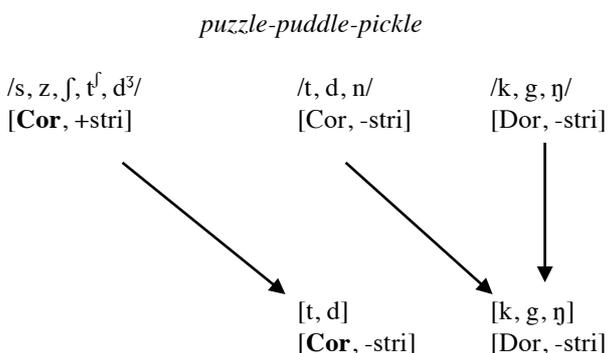
(18)   *Required ranking for the puzzle-puddle-pickle chain shift* (from Jesney (2005:87))

*[+strident], IdentCoronal/[+strident] >> *TL >>IdentCoronal/[-strident]

As noted by Wolf (2008:356), Jesney's approach to child chain shifts (combined with the assumption that Prec constraints are bottom-ranked in the initial state) predicts, *inter alia*, that chain shifts (counterfeeding-on-focus) are the only type of spontaneous opacity that should arise. Wolf continues to say that the discovery of cases of spontaneous counterbleeding would demonstrate the need for Prec constraints being high-ranked in the initial state.

The cases of spontaneous counterbleeding are, in fact, well attested. Barlow & Keare (2008:84), for example, describe a case of non-target-like opacity (exemplified in 19) whereby the process of vowel-lengthening before voiced obstruents is counterbled by word-final voice neutralization. The learner correctly lengthens the vowel preceding the underlyingly voiced obstruent, but devoices the obstruent word-finally, which results in non-surface-apparent outputs.

(19)   *Spontaneous counterbleeding (adapted from Barlow & Keare 2008:84)*
      a. /dɔg/   →   [dɔːk] 'dog' cf.   /dɔgi/ →   [dɔːgi] 'doggie'
      b. /dʌk/   →   [dʌk] 'duck' cf.   /dʌki/ →   [dʌki] 'duckie'

Needless to say that preferential faithfulness cannot handle overapplication. OT-CC analysis of counterbleeding, in turn, requires that the relevant Prec constraint should dominate some faithfulness constraint. Since the learner lacks the positive evidence necessary to establish such a ranking, we have no choice but to assume that the ranking is due to an innate bias in conjunction with BCD.

      The need for Prec constraints to be high-ranked in the initial state has already been contemplated by scholars working on language acquisition. Thus, Dinnsen & Farris-Trimble (2008:115) come to the conclusion that "the fact that opacity effects emerge naturally in the course of early acquisition suggests that in the initial-state, Prec constraints are ranked relatively high among the faithfulness constraints". The fact that Prec comes to dominate markedness constraints (in our case *TL) in the absence of any positive evidence, has been attributed to the "imperfect partial learning" (see Dinnsen & Farris-Trimble (2008) and references therein). In their analysis, Dinnsen & Farris-Trimble (2008) adhere to the original proposal by McCarthy (2007), whereby the ranking of Prec is restricted by the metaconstraint. However, above we have seen that there are reasons to favor discarding the metaconstraint on the ranking of Prec. Once the metaconstraint is discarded, nothing forces Prec to be low-ranked at the initial state. In the next section I will outline my proposal concerning the status of Prec constraints in more detail.

4. Status of Prec constraints

If we assume that the innate bias for Markedness >> Faithfulness ranking is justified, discarding the metaconstraint leaves us with two options with respect to the default ranking of Precedence constraints.

(20)   *Hypothesized Initial State*
      a. Precedence >> Markedness >> Faithfulness
      b. Markedness >> Precedence >> Faithfulness

Let's consider these options in turn. Prec(A,B) can be satisfied in three cases: first, by the faithful candidate that violates neither A nor B; second, by the candidate that violates only A; third, by the candidate that violates both A and B in the order prescribed by the relevant Prec constraint. I will illustrate the consequences of Prec being high-ranked with the example of counterfeeding opacity. If Prec is ranked above markedness at the initial state, the most faithful candidate will be chosen as optimal.

(21)   *Tableau 4: Spontaneous chain-shift in OT-CC (adapted from Wolf 2008:355)*

	/pʌzəl/	PREC(IDENT(place), IDENT(contin))	*z	*TL	IDENT(place)	IDENT(contin)
☞ a.	< pʌzəl > <∅>		*	*		
☹ b.	< pʌzəl , pʌdəl > <ID(contin)>	*		*		*
c.	< pʌzəl , pʌdəl , pʌgəl > <ID(contin), ID(place)>	*!*			*	*

This is clearly not the result that we want, due to two reasons. First, it is inconsistent with the attested early production data. Second, and most important, if the output of child's grammar equals the adult output no learning can take place: BCD will converge on the incorrect and partial ranking Prec >> *z, *TL >> Id(place), Id(contin).

      This ranking hypothesis becomes even more problematic if we assume, following McCarthy (2007), that Prec constraints are innate. It would mean that the uppermost stratum would contain not only

Prec(A,B), but also Prec(B,A). Working together, the two Prec constraints would eliminate all candidates but a faithful parse that always vacuously satisfies Prec.

Option (20 b) is much more plausible. Prec can be viewed as some sort of a quasi-faithfulness constraint, because it is always vacuously satisfied by the faithful candidate. Ranked immediately below markedness, Prec constraints will jointly favor the most faithful of the unmarked candidates. In classic OT, this is exactly the job done by regular faithfulness constraints at the initial state. Therefore, Prec constraints ranked as shown in (20 b) will not interfere with the evaluation process. At some point of the acquisition the learner is supposed to notice that the intended winner (adult form) violates some high-ranked markedness constraint. In the face of such positive evidence she will be prompted to demote the relevant markedness constraint. When Prec constraint comes to dominate markedness, opacity effects might arise (though it's not always that they do, as we will see in Section 5).

Another issue that is worth considering is whether Prec constraints are innate or emergent (see Boersma 1998, Hayes 1999, Fikkert & Levelt 2008 on innate vs. emergent constraints in acquisition). In the original proposal, McCarthy (2007:119) in principle allows for both options, with a remark that the issue is not very important due to the narrow range of logically possible Prec constraints. However, if we allow Prec constraints to be relatively high-ranked at the initial state, we also endow them with the power to crucially affect the evaluation process. Needless to say that Prec(A, B) alone and Prec (A, B) working in conjunction with Prec(B,A) will favor different candidates. Also, if we allow all logically possible Prec constraints to be present in the hierarchy, the evaluation process is at risk of being crucially affected by a large body of completely irrelevant constraints like Prec(C,B), Prec(D,B) etc.

Therefore, the assumption that Prec constraints are constructed by the learner in the process of acquisition seems justified. But it raises another important question: what prompts the learner to construct Prec constraints? In his brief discussion of learnability issues in OT-CC, McCarthy (2007:119) remarks that Prec constraints might be constructed when the learner faces "otherwise refractory data", meaning that the Prec constraints are constructed as a 'last resort' when simple reranking of faithfulness and markedness cannot get the intended winner to be chosen as optimal. The main counterexample to this approach is the very existence of spontaneous opacity effects: the child who is trying to acquire transparent generalizations will never come across contradictory data, and therefore she will never have a reason to construct Prec constraints. So what is the trigger?

I propose that Prec constraints are constructed when the learner detects that the rLUMSeq of at least one of the candidates is not empty. It would be probably naïve to assume that of all logically possible (usually two) options, exactly the constraint that will 'save the day' is constructed. Therefore, we propose that once the learner detects that some candidate has a non-empty rLUMSeq, all logically possible constraints are constructed. The further course of events depends on whether the learner is trying to acquire opaque or transparent generalization. In the former case, the learner will place newly-constructed Prec constraints into the hierarchy one by one. The Prec constraint that will resolve the conflict in the data will be kept, while all the others will be discarded and will not participate in the further ranking. When the learner is trying to acquire transparent generalization, there is no contradiction in the data to start with, so any Prec constraint will 'fit'. Which one is chosen, then, is purely a matter of chance. The choice, however accidental it might be, determines the intermediate stages the child will go through on her way to the adult output. In some cases, as we will see in Section 5, the child who opted for $Prec_1$ will go through the 'opaque' stage, while the outputs of the child who opted for $Prec_2$ will be consistently transparent. Thus the model predicts that spontaneous opacity effects are not obligatory, and that children having chosen different Prec constraint will go through different intermediate stages on their way to the adult output.

5. Spontaneous counterbleeding in modified OT-CC

In this section I will analyze emergent counterbleeding opacity effects demonstrated by the learners acquiring transparent generalizations. I will also address the problem of variation in early production data,

and try to explain why some children demonstrate spontaneous opacity effects, while others have consistently transparent outputs.

5.1 Learner A: overapplication effects

Barlow & Keare (2008) present a case of spontaneous counterbleeding opacity, whereby children neutralize voicing contrast on syllable-final obstruents while still maintaining vowel lengthening before underlyingly voiced obstruents. Relevant child data are provided below.

(22) *Voiced and voiceless obstruents word-initially and intervocalically (Barlow & Keare 2008:84)*
    a. [dɔ:k]   'dog'   [tʰʌp]   'tub'
    b. [bi:ts]   'bridge'   [fis]   'fish'
    c. [tʃi:zi]   'cheese (dim.)'   [fisi]   'fish (dim)'

(23) *Voice contrast neutralized word-finally (Barlow & Keare 2008:84)*
    a. [dɔ:k]   'dog'   [dʌk]   'duck'
    b. [dʌ:f]   'glove'   [wif]   'leaf'

(24) *Morphophonemic alternations for target morpheme-final voiced obstruents (Barlow & Keare 2008:85)*
    a. [dɔ:k]   'dog'   [dɔ:gi]   'dog (dim.)'
    b. [dʌ:f]   'glove'   [gʌ:vi]   'glove (dim.)'

Before I analyze child production at this stage, I have to make explicit certain assumptions about the state of the child's grammar at this point. Errors in production indicate that the child is still struggling with language-specific phonotactics. Productive knowledge of voicing patterns also indicates that lexical representations of voiced and voiceless obstruents are adult-like (Barlow & Keare 2008:85). I also assume that at this point of acquisition the child has already departed from the identity map with respect to vowel length, having observed the fact that the length of a vowel is always predictable from the quality of the following consonant. In other words, the underlying representations the child has internalized are completely adult-like at this point in acquisition (i.e. /dɔg/ for [dɔ:g] etc.)

    Provided that the above assumptions are correct, the error-pattern we are looking at is a case of counterbleeding opacity, whereby vowel-lengthening applies out of context, i.e. before a voiceless consonant. This counterbleeding opacity is spontaneous, or non-target-like, because neither lengthening of vowels before voiceless consonants nor final devoicing happen in the variety of English that the child is trying to acquire.

(25) *Spontaneous counterbleeding (adapted from Barlow & Keare 2008)*
    a. /dɔg/   →   [dɔ:k]   'dog'
    b. /dɔgi/   →   [dɔ:gi]   'dog(dim.)'
    c. /dʌk/   →   [dʌk]   'duck'
    d. /dʌki/   →   [dʌki]   'duck (dim.)'

In (25) we can see that the learner lengthens a vowel followed by the underlyingly voiced obstruent, and at the same time devoices the triggering obstruent. Below I will analyse this case of overapplication in the light of OT-CC. Following Barlow & Keare (2008), I employ the following constraints in the analysis:

(26)   *Devoicing and shortening constraints Barlow & Keare (2008:87)*
   a. ***VC̬** – short vowels before voiced obstruents are prohibited.
   b. ***C̬#** – voiced coda obstruents are prohibited.
   c. **ID[length]** – input and output segments should have identical specifications for [length].
   d. **ID[voice]** – input and output segments should have identical specifications for [voice].

Barlow & Keare 2008 present the analysis of this case of spontaneous counterbleeding in terms of OT-CC. Consistent with the original proposal by McCarthy (2007:119), they assume that precedence constraints are innate, and that at the initial state they are ranked at the bottom of the hierarchy. Furthermore, their analysis respects the metaconstraint on the ranking of precedence constraints. The analysis is given below, where Prec stands for Prec(ID[length], ID[voice]):

(27)   *Tableau 5: Spontaneous counterbleeding in OT-CC (from Barlow & Keare 2008:88)*

	/dɔg/ 'dog'	*VC̬	*C̬#	IDENT(voice)	PREC	IDENT(length)
a.	<dɔg> faithful candidate	*!	*			
b.	<dɔg, dɔːg> IDENT(length)		*!			*
☞c.	<dɔg, dɔːg, dɔːk> IDENT(length), IDENT(voice)			*		*
d.	<dɔg, dɔk> IDENT(voice)			*	*!	

In Tableau 5 above, the faithful candidate violates both high-ranked markedness constraints, and is therefore eliminated. Candidate (b), the intended winner, also violates a high-ranked markedness constraint against voiced codas. Transparent candidate (d) incurs a fatal violation of the Prec constraint, and thus opaque candidate (c) is correctly predicted to win.

  Though the ranking illustrated above correctly accounts for the data, it raises a familiar question first sounded by Wolf (2008:355). What evidence do learners have for ranking Prec constraint that high? The answer is: none. On the basis of the positive evidence from the target-language, the learner must conclude that in order to get the optimal candidate (here: [dɔːg]) to win, all is needed is the demotion of *C̬# below ID[voice]. The ranking shown in Tableau 5 could have resulted only if the learner maliciously intended to derive the 'illegal' opaque output. Acquisition algorithms available to date do not provide for such a possibility, and therefore the above analysis cannot be taken to reflect a plausible learning situation.

  Below I present the acquisition model based on BCD (see Prince & Tesar 2004) and Minimal GLA (Gradual Learning Algorithm, see Boersma 2008). In accordance with these works, I assume that in the initial state markedness constraints dominate faithfulness constraints. I also assume based on the considerations outlined in the previous section that a precedence constraint, when constructed and placed into the hierarchy, occupies a stratum immediately below markedness and above faithfulness. In addition to the markedness constraints suggested by Barlow & Keare (2008), I also postulate a context-free markedness constraint against long vowels, *LongV, ranked below the context-sensitive *VC̬. Therefore, the initial state ranking is as shown below:

(28)   *Ranking at the Initial State*

  *VC̬, *C̬# >> *LongV >> *Prec Constraint* >> ID[length], ID[voice]

In error-driven constraint demotion algorithms (see Tesar 1995, Prince & Tesar 2004, Boersma 2008), a suboptimal form or 'loser' is an output of the learner's current grammar. In OT-CC, however, suboptimal

competing candidates are generated by the learner on the basis of the input and current constraint ranking. This solves the problem of the potentially infinite search space of competing output forms (Tesar 1995) and, combined with Minimal GLA (Boersma 2008), ensures that the informative loser is always available for the language learner.

First the learner has to construct potential outputs for the input /dɔg/ on the basis of her current grammar.

(29)   *Candidate chains for the input /dɔg/*
        a.   <dɔg>
        b.   <dɔg, dɔk>, because *C# >> ID[voice]
        c.   < dɔg, dɔ:g>, because *VC̬ >> ID[length]
        d.   <dɔg, dɔ:g, dɔ:k>, *VC̬ >> ID[length], and *C# >> ID[voice]

(30)   *Candidates for the input /dɔg/ ordered as (in, out, L-set, rLUMseq)*
        a.   (/dɔg/, dɔg, Ø,Ø)
        b.   (/dɔg/, dɔk, ID[voice], Ø)
        c.   (/dɔg/, dɔ:g, ID[length], Ø)
        d.   (/dɔg/, dɔ:k, {ID[length], ID[voice]}, < ID[length], ID[voice]>)

Once the candidate chains have been constructed, the learner detects that the candidate set contains a candidate whose rLUMseq is not empty. This observation prompts the learner to construct Prec constraints: Prec(ID[length], ID[voice]) and Prec(ID[voice], ID[length]). As I mentioned in the previous section, which constraint gets chosen is purely incidental in this case. However, for the sake of illustration we will consider both scenarios in turn.

Suppose the learner has chosen Prec(ID[length], ID[voice])[2] and inserted it into its designated slot in the hierarchy. Then, in accordance with Minimal GLA (Boersma 2008), the learner has to choose some random total ranking of constraints consistent with her current stratified grammar and compute the output of her grammar. She that it differs from the optimal output. This is illustrated in the Tableau 6 below.

(31)   *Tableau 6: Initial stage: transparent outputs*

	/dɔg/	*VC̬	*C#	*LongV	PREC(IDENT(length), IDENT(voice))	IDENT(length)	IDENT(voice)
a.	dɔg	*	*!				
☛※ b.	dɔk				*		*
☹ c.	dɔ:g		*!	*		*	
d.	dɔ:k			*!		*	*

In Tableau 6 above, faithful candidate (a) is ruled out by the high-ranked markedness constraint banning short vowels followed by voiced obstruents. The intended winner in (c) incurs a fatal violation of the markedness constraint against voiced codas, while opaque form (d) crucially violates *LongV. Thus, the transparent unmarked candidate in (b) is judged optimal.

Having computed the output of her current grammar, the child detects that it is different from the adult output: the child's grammar produces the output [dɔk], while predicting 'target' output [dɔ:g] to be suboptimal. The learner then constructs winner-loser pairs for the input /dɔg/, and performs the demotions necessary to make the form [dɔ:g] the winner. Constraint demotion is carried out in accordance with the BCD algorithm (see Prince & Tesar 2004).

---

[2] Note that in case of (spontaneous) counterbleeding it is the insertion of the Prec constraint that corresponds to the actual ordering of LUMs in the candidate chain that leads to opaque outputs. In case of (spontaneous) counterfeeding (see (16)) it is the insertion of the Prec constraint that refers to the reverse order of LUMs that leads to opaque outputs.

(32)   *Tableau 7: Winner-loser pairs for the input /dɔg/*

/dɔg/	*V̌C̬	*C#	*LongV	PREC (IDENT(length), IDENT(voice))	IDENT(length)	IDENT(voice)
a. dɔg	*	*				
c. dɔ:g		✪	✪		✪	
b. dɔk				*		*
c. dɔ:g		✪	✪		✪	
d. dɔ:k			*		*	*
c. dɔ:g		✪	✪		✪	

(33)   *Tableau 8: Winner-loser pairs for the input /dɔg/*

/dɔg/	*V̌C̬	*C#	*LongV	PREC(IDENT(length), IDENT(voice))	IDENT(length)	IDENT(voice)
a. dɔg <<c. dɔ:g	W		L		L	
b. dɔk <<c. dɔ:g		L	L	W	L	W
d. dɔ:k <<c. dɔ:g		L				W

Since no ranking of markedness constraints can reconcile the learner with the data, the learner has to demote some markedness constraint. In accordance with Minimal GLA (Boersma 2008), the demotion is gradual, meaning that one constraint can be demoted by one stratum at a time. The first demotion of *LongV will result in the following hierarchy:

(34)   *Tableau 9: Intermediate stage: opaque outputs*

/dɔg/	*V̌C̬	*C#	PREC(IDENT(length), IDENT(voice))	*LongV	IDENT(length)	IDENT(voice)
a. dɔg	*!	*				
b. dɔk			*!			*
☹c. dɔ:g		*!		*	*	
☞d. dɔ:k				*	*	*

In Tableau 9 above, candidate (a) is ruled out by the high-ranked markedness constraint banning short vowels before voiced obstruents. Candidate (b) is eliminated due to the violation of the Prec constraint. The intended winner, candidate (c) fatally violates the still-high-ranked *C#. Therefore, opaque candidate (d) is selected as an optimal output.

     On the basis of winner-loser pairs, the learner again makes necessary adjustments to her grammar (gradually), until she arrives at the ranking where ID[voice] dominates *C#. At this point the learner has converged on the target grammar[3].

(35)   *Tableau 10: Final stage: transparent outputs*

/dɔg/	*V̌C̬	PREC (IDENT(length), IDENT(voice))	*LongV	IDENT(voice)	*C#	IDENT(length)
a. dɔg	*!				*	
☞b. dɔ:g			*		*	*

In Tableau 10 above, candidate (a) is ruled out by a high-ranked markedness constraint *V̌C̬. Therefore, candidate (b) is correctly selected as an optimal output, i.e. the child has learned the adult ranking. It is worth highlighting that the final grammar in Tableau 10 is restrictive and also fully consistent with the

---

[3] Note that since *C# is dominated by ID[voice], forms *[dɔk] and *[dɔ:k] are no longer in the candidate set.

Richness of the Base principle, i.e. it maps any input onto some language-legal output. Thus, the hypothetical input /dɔ:g/ is correctly mapped to [dɔ:g], while hypothetical /dɔ:k/ will be mapped to [dɔk].

It is also noteworthy that the input form /dɔgi/ is correctly mapped to [dɔ:gi] throughout the learning process, which is consistent with the attested data.

(36)   *Candidate chains for the input /dɔgi/*
         a.   < dɔgi >
         b.   < dɔgi, dɔ:gi>, because *VČ̬ >> ID[length]

(37)   *Candidates for the input /dɔgi/ ordered as (in, out, L-set, rLUMseq)*
         a.   (/dɔgi/, dɔgi, Ø,Ø)
         b.   (/dɔgi/, dɔ:gi, ID[length], Ø)

(38)   *Tableau 11: Final stage*

	/dɔgi/	*VČ̬	Prec(Ident(length), Ident(voice))	*LongV	*C#	Ident(length)	Ident(voice)
a.	dɔgi	*!					
☞b.	dɔ:gi			*		*	

In this subsection we have seen that OT-CC correctly accounts for the emergence and subsequent loss of overapplication effects in developing grammars. In the next subsection I will show that the modified OT-CC can account for cross-subject variation, which is widely documented in the studies of early production.

5.2   *Learner B: alternative acquisition path*

It has been observed that not all children go through the 'opaque' stage on their way to adult grammar: some learners (see Dinnsen et al. 2000) consistently produce transparent outputs. According to Dinnsen et al. (2000:337), "such variability is a hallmark of developing systems and is in no way atypical". It goes without saying that a good acquisition model should be able to account for this state of affairs.

In the previous subsection we have demonstrated that OT-CC is able to account for the emergence and disappearance of overapplication opacity effects in developing grammars. In this subsection I will show that though OT-CC predicts (in full consistency with the attested data) that opacity effects might arise in the course of acquisition of transparent generalizations, it does not claim that such effects must and will arise.

Recall from Section 4 that Prec constraints are constructed by the learner when she detects that the rLUMSeq of some candidate is not empty. I have assumed that the learner constructs all logically possible Prec constraints (usually 'all' means 'two'). The further course of events depends on whether the learner is trying to acquire an opaque or a transparent generalization. In the former case, the learner places newly-constructed Prec constraints into the hierarchy one by one. The Prec constraint that resolves the conflict in the data is kept, while all the others are discarded and do not participate in further ranking. When the learner is trying to acquire a transparent generalization, there is no contradiction in the data to start with, so any Prec constraint will 'fit'. Which one is chosen is a matter of chance. The choice, however accidental it might be, determines the intermediate stages the child goes through on her way to the adult output. In some cases, the child who opted for $Prec_1$ goes through the 'opaque' stage, while the outputs of the child who opted for $Prec_2$ are consistently transparent.

In Subsection 5.1 I have shown that a child who has demonstrated overapplication opacity effects while trying to acquire transparent generalization of English must have chosen Prec(ID[length], ID[voice]). In this section I will consider the alternative scenario and see what intermediate stages the learner would have gone through had she opted for Prec(ID[voice], ID[length]) instead.

The postulation of the 'alternative' Prec constraint, of course, does not influence the outcome of the learning process: the learner is still able to eventually converge on the correct adult output. However, the learning path of the hypothetical Learner B is predicted to be different from that of Learner A illustrated above. Note also that Learner B has exactly the same candidate set as Learner A, because Prec constraints do not participate in the construction of candidate chains.

Having placed Prec(ID[voice], ID[length]) into the hierarchy, Learner B computes the output of her current grammar. This is illustrated in Tableau 12 below.

(39) *Tableau 12: Initial stage: transparent outputs*

	/dɔg/	*VC̬	*C#	*LongV	PREC (IDENT(voice), IDENT(length))	IDENT(length)	IDENT(voice)
a.	dɔg	*!	*				
☞b.	dɔk						*
☹ c.	dɔ:g		*!	*	*	*	
d.	dɔ:k			*!	**	*	*

In Tableau 12, a fully faithful candidate (a) violates both high-ranked markedness constraints. The intended winner, candidate (c), violates the markedness constraint against voiced codas. The opaque candidate (d) incurs a fatal violation of the *LongV constraint. Therefore, transparent candidate (b) is selected as optimal. Note that transparent candidate (b) harmonically bounds opaque candidate (d) – this ensures that Learner B will consistently have transparent outputs in the course of acquisition.

Having detected that the output of her grammar differs from the adult output, Learner B constructs winner-loser pairs for the input /dɔg/.

(40) *Tableau 13: winner-loser pairs for the input /dɔg/*

	/dɔg/	*VC̬	*C#	*LongV	PREC(IDENT(voice), IDENT(length))	IDENT(length)	IDENT(voice)
a.	dɔg	*	*				
c.	dɔ:g		✪	✪	✪	✪	
b.	dɔk						*
c.	dɔ:g		✪	✪	✪	✪	
d.	dɔ:k			*	*	*	*	*
c.	dɔ:g		✪	✪	✪		✪	

(41) *Tableau 14: winner-loser pairs for the input /dɔg/*

/dɔg/	*VC̬	*C#	*LongV	PREC(IDENT(voice), IDENT(length))	IDENT(length)	IDENT(voice)
a. dɔg <<c. dɔ:g	W		L	L	L	
b. dɔk <<c. dɔ:g		L	L	L	L	W
d. dɔ:k <<c. dɔ:g		L		W		W

Now the learner demotes the constraints violated by the intended winner in accordance with the BCD algorithm (Prince & Tesar 2004). In accordance with Minimal GLA (see Boersma 2008), the demotion should be minimal and gradual: one constraint can be demoted by one stratum at a time. At some point, Learner B will arrive at the target hierarchy as shown in Tableau 15 below[4].

---

[4] Note that since ID[voice] dominates **C#,** forms *[dɔ:k] and *[dɔk] are no longer in the candidate set.

(42)   *Tableau 15: Final state: transparent outputs*

/dɔg/	*VC̬	ID[voice]	*C#	*LongV	PREC (IDENT(voice), IDENT(length))	IDENT(length)
a. dɔg	*!		*			
☞b. dɔ:g			*	*	*	*

In Tableau 15, faithful candidate (a) violates a high-ranked constraint prohibiting short vowels before voiced obstruents. Therefore, candidate (b) is correctly chosen to be optimal: Learner B has successfully converged on the target grammar.

5.3   More cross-subject variation data

Dinnsen *et al.* (2000) provide data from two children suffering from phonological delay, which might be taken to exemplify both opaque and transparent acquisition paths. Thus, Child A showed evidence for the counterbleeding interaction of coda deletion and vowel lengthening before voiced consonants, whereby the vowel was lengthened even if the triggering consonant was deleted. Child C, in turn, demonstrated transparent interaction of vowel lengthening and coda deletion, whereby vowels preceding the deleted voiced consonant remained short. Consider the production data below:

(43)   *Child A (age 7;2) (from Dinnsen et al. 2000:325)*
         a.   [kæ:]   'cab'   [ka]   'cop'
         b.   [kɪ:]   'kid'   [pæ]   'pat'
         c.   [dɔ:]   'dog'   [dʌ]   'duck'

(44)   *Child C (age 3;10) (from Dinnsen et al. 2000:338)[5]*
         a.   [dɔ]   'dog'   [tʌ]   'truck'
         b.   [dæ]   'dad'   [i]   'eat'

If we assume that coda deletion is due to the markedness constraint NoCoda ranked above MAX, and vowel lengthening is due to *VC̬ ranked above Ident(length), the following candidate chains for the input /dɔg/ can be constructed:

(45)   *Candidate chains for the input /dɔg/*
         a.   <dɔg>
         b.   <dɔg, dɔ>, because NoCoda >> MAX
         c.   < dɔg, dɔ:g>, because *VC̬ >> ID[length]
         d.   <dɔg, dɔ:g, dɔ:>, *VC̬ >> ID[length], and NoCoda >> MAX

(46)   *Candidates for the input /dɔg/ ordered as (in, out, L-set, rLUMseq)*
         a.   (/dɔg/, dɔg, Ø,Ø)
         b.   (/dɔg/, dɔ, MAX, Ø)
         c.   (/dɔg/, dɔ:g, ID[length], Ø)
         d.   (/dɔg/, dɔ:, {ID[length], ID[voice]}, < ID[length], MAX>)

---

[5] In fact, the situation is a bit more complex since Child C also deletes obstruents word-medially. Dinnsen et al. 2000 attribute it to the ranking NoCoda >>*VC >>MAX, where *VC is a markedness constraint against post-vocalic obstruents. In case of Child A, the ranking is NoCoda >> MAX >> *VC, to the effect that only coda obstruents are deleted. Since both children treat coda obstruents in exactly the same way, for the purposes of our analysis we will disregard the effects of *VC and assume the ranking NoCoda >> Max for both children.

In case of Child A, candidate (d) is selected as an optimal output at some point of acquisition. In case of Child C, it is candidate (b). Below we will show that our assumptions about the status of Prec constraints can account for the production data of both children.

Having constructed candidate chains and corresponding candidates, as shown in (45) and (46) above, both children will detect that rLUMSeq of one of the candidates is not empty. Therefore, the children will construct both logically possible Precedence constraints, in this case these are Prec(Ident(length), MAX)) and Prec(MAX, Ident(length)). At this point the acquisition path forks: Child A chooses Prec(Ident(length), MAX), while Child C chooses Prec(MAX, Ident(length)). As a result of the choice, Child A will go through the opaque stage, while Child C will have consistently transparent outputs.

(47)   *Tableau 16*: *Child A: opaque intermediate stage*

	/dɔg/	*VC̬	NoCoda	Prec (Ident(length), MAX)	*LongV	Ident(length)	MAX
a.	dɔg	*!	*				
b.	dɔ			*!			*
☹ c.	dɔːg		*!		*	*	
�" d.	dɔː				*	*	*

In Tableau 16 above, candidate (a) is ruled out due to the violation of the high-ranked markedness constraint against short vowels followed by voiced obstruents. Unmarked candidate (b) fatally violates the Prec constraint, since MAX-violating LUM in its L-set is not preceded by ID[length]-violating LUM. The intended winner, candidate (c) fatally violates the high-ranked NoCoda. Therefore, overapplication candidate (d) is selected as optimal, in consistency with the attested data.

(48)   *Tableau 17*: *Child C: transparent intermediate stage*

	/dɔg/	*VC̬	NoCoda	Prec(MAX, Ident(length))	*LongV	Ident(length)	MAX
a.	dɔg	*!	*				
�" b.	dɔ						*
☹ c.	dɔːg		*!	*	*	*	
d.	dɔː			**!	*	*	*

The situation with Child C shown in Tableau 17 is somewhat different. Just like in case with Child A, candidates (a) and (c) are ruled out by the high-ranked markedness constraints. Opaque candidate (d) fatally violates the Prec constraint: ID[length]-violating LUM in its L-set is followed by MAX-violating LUM, which is the opposite of what the Prec constraint requires. Therefore, transparent candidate (b) is selected as optimal. Note also that opaque candidate (d) is harmonically bounded by (b), which means that Child C is predicted not to have opaque outputs at any point of acquisition.

In a nutshell, from our above analysis it follows that differences in production of Child A and Child C are due to the fact that these children follow *different acquisition paths*. However, this is not the only way to look at the data in question.

Dinnsen *et al*. (2000) analyze the production data of Child A and Child C in the light of Sympathy Theory (see McCarthy 1999). Under their approach, differences in production between Child A and Child C are due to the fact that at the moment when the data were attested the children were at the *different stages of development*. This is schematically represented below:

(49)   *Stages of development (from Dinnsen et al. 2000:343)*

> Stage 1: Markedness >> Sympathy >> IO Faith
>   (e.g., Child C, transparent outputs)
> Stage 2: Sympathy >> Markedness >> IO Faith
>   (e.g., Child A, opaque outputs)
> Stage 3: Sympathy >> IO Faith >> Markedness
>   (e.g., archetypical fully developed language, transparent outputs)

According to Dinnsen *et al*. (2000), a sympathy constraint can affect evaluation only if it dominates some markedness constraint. Therefore, during the first stage of acquisition, the effect of the sympathy constraint is not visible, and the learner produces transparent outputs (according to Dinnsen *et al*. 2000, this stage is exemplified by Child C). Then, on the basis of positive evidence, markedness constraints are gradually demoted, so that at some point they are dominated by the sympathy constraint. This is when the opaque outputs are produced (this is allegedly the stage Child A is at). Finally, markedness constraints get demoted below IO faithfulness constraints, rendering sympathy constraints inert.

    The insight expressed by Dinnsen *et al*. (2000) is largely consistent with the OT-CC-based acquisition model outlined in this section. Thus, above we have seen that Prec, just like sympathy, can influence the outcome of evaluation if it is ranked above some markedness constraint. Just like sympathy, Prec can only come to dominate markedness if markedness is demoted on the basis of positive evidence. Just like sympathy, Prec constraints become inert if relevant markedness constraints get demoted below faithfulness constraints. Recall from Subsection 5.1 that during the initial stage of acquisition our Learner A had transparent outputs, and it was only during the later stages that opacity effects emerged. Therefore, just like Dinnsen *et al*. (2000), we could have claimed that differences in production between Child A and Child C are merely due to the fact that the children are in different stages of development.

    However, certain considerations prompt us to be cautious with drawing conclusions. Firstly, the account presented in Dinnsen *et al*. (2000) predicts that opacity effects will obligatorily arise at a certain point in acquisition. Moreover, the 'opaque' stage must follow and not precede the 'transparent' stage. Intuitively, this claim seems to be too strong; however, scrupulous longitudinal studies are necessary to see if it is warranted. Secondly, though Child C (aged 3;10) is indeed younger than Child A (aged 7;2), the fact that both children suffer from phonological delay might prevent us from taking this particular case to exemplify continuity in phonological development. To this end, it would be especially interesting to carry out a dynamic investigation of several typically developing children and see if any correlation can be drawn between their development and emergence of opacity effects (if any). Therefore, in the light of the aforementioned, our hypothesis is that differences in production between Child A and Child C are due to the fact that the children have postulated different Prec constraints in the course of the acquisition. The hypothesized learning progression for Child A and Child C is summarized below:

(50)   *Hypothesized learning progression for Child A and Child C*

Child A

> **Early stage: transparent outputs**
>   *VC̬ , NoCoda >> *LongV >> Prec(ID[length],MAX) >> ID[length], MAX
>
> **Intermediate stage: opaque outputs**
>   *VC̬ , NoCoda >> Prec(ID[length],MAX) >> *LongV >> ID[length], MAX
>
> **Final stage: target-appropriate outputs**
>   *VC̬ >> Prec(ID[length],MAX) >> *LongV >> ID[length], MAX >> NoCoda

<u>Child C</u>

Early stage: transparent outputs
$$\text{*V\textasciibreve C}, \text{NoCoda} \gg \text{*LongV} \gg \text{Prec(MAX, ID[length])} \gg \text{ID[length], MAX}$$

Intermediate stage: transparent outputs
$$\text{*V\textasciibreve C}, \text{NoCoda} \gg \text{Prec(MAX, ID[length])} \gg \text{*LongV} \gg \text{ID[length], MAX}$$

Final stage: target-appropriate outputs
$$\text{*V\textasciibreve C} \gg \text{Prec(MAX, ID[length])} \gg \text{*LongV} \gg \text{ID[length], MAX} \gg \text{NoCoda}$$

In this section I have demonstrated that OT-CC can successfully account for the emergence and subsequent disappearance of overapplication effects in developing grammars. Importantly, we have seen that though OT-CC predicts (in full consistency with the attested data) that opacity effects might arise in the course of acquisition of transparent generalizations, it does not claim that such effects must and will arise: whether or not the learner will have opaque outputs depends entirely on which Prec constraint is chosen. Thus, OT-CC predicts a certain degree of variation in early production data. In the next section we will see if the assumptions we have made and conclusions we have reached so far can help us to analyse emergent underapplication effects in developing grammars.

6. Conclusions

The focus of this paper is a recent OT-based theory of opacity called *Optimality Theory with Candidate Chains* (OT-CC, see McCarthy 2007), which represents the synthesis of OT with derivations. I question the status of Precedence constraints and introduce the following changes to the original formulation of OT-CC proposed by McCarthy (2007):

- There is no *B >> Prec(A,B)* metaconstraint on the ranking of Prec constraints.
- Prec constraints are not innate.
- All logically possible Prec constraints are constructed when the learner detects a non-empty rLUMSeq.
- Newly constructed Prec constraints are placed into the hierarchy one by one.
- Prec constraints are subject to the ranking bias, i.e. they are inserted into the hierarchy immediately below markedness constraints and above faithfulness constraints.
- If the learner is trying to acquire an opaque generalization, the Prec constraint that resolves the conflict in the data is kept, all others are discarded.
- If the learner is trying to acquire a transparent generalization, the choice of Prec constraint is incidental and sometimes results in the emergence of developmental opacity effects.

I demonstrate that the adjustments I have introduced make it possible for OT-CC to account for the emergence and subsequent loss of spontaneous counterbleeding effects in developing grammars. Importantly, I show that though OT-CC predicts (in full consistency with the attested data) that opacity effects might arise in the course of acquisition of transparent generalizations, it does not claim that such effects must and will arise. Thus, my analysis predicts a certain degree of variation in early production data. I demonstrate that constraint rankings corresponding to the attested error patterns arise naturally in the course of acquisition, and the order in which they arise is fully consistent with the attested data. Besides, I show that there is a high degree of continuity in transition from one developmental stage to the next. I also demonstrate that acquisition of target-like opacity effects can be modeled within OT-CC (McCarthy 2007) with the minimum of machinery and theory-internal stipulation. In my analyses I

closely follow current stands on such acquisition issues as the nature of the Initial State, acquisition of phonotactics, underlying representations and morphophonology. Acquisition models I have presented are fully compatible with BCD (Prince & Tesar 2004) and Minimal GLA (Boersma 2008).

The approach advocated in this paper demonstrably solves many problems inherent in the previous approaches to opacity, while at the same time preserving their merits. Therefore, it makes a good starting point for the further theoretical and empirical research on the problem of acquisition of spontaneous and target-like opacity effects.

References

Barlow, Jessica A. and Amanda Keare. 2008. Acquisition of final voicing: An acoustic and theoretical account. Indiana University Working Papers in Linguistics: Volume 6.

Bermúdez-Otero, Ricardo 2003. The acquisition of phonological opacity. In Spenader et al. (eds) *Variation within Optimality Theory*, 25-36. Stockholm: Department of Linguistics, Stockholm University [available on ROA--593--0403].

Boersma, Paul. 1998. Functional Phonology. *Formalizing the Interactions between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Boersma, Paul. 2008. Some correct error-driven versions of the Constraint Demotion algorithm. ROA-980.

Dinnsen, Daniel A. 2008. A typology of opacity effects in acquisition. In Daniel A. Dinnsen and Judith A. Gierut (eds.) *Optimality Theory, Phonological Acquisition and Disorders*. London, Oakville: Equinox

Dinnsen, Daniel A., Laura W. McGarrity, Kathleen M. O'Connor and Kimberly A. B. Swanson. 2000. On the role of sympathy in acquisition. Language Acquisition 8:321-361.

Dinnsen, Daniel A., Kathleen M. O'Connor and Judith A. Gierut. 2001. The puzzle-puddle-pickle problem and the Duke-of-York gambit in acquisition. J. Linguistics 37: 503-525.

Dinnsen, Daniel A. and Ashley W. Farris-Trimble.2008. An opacity-tolerant conspiracy in phonological acquisition. Indiana University Working Papers in Linguistics: Volume 6.

Farris-Trimble, Ashley W. 2008. Cumulative faithfulness effects in phonology. ROA-991.

Fikkert, Paula and Claartje Levelt. 2008. "How does place fall into place? The lexicon and emergent constraints in children's developing phonological grammar". In Peter Avery, Elan Dresher & Keren Rice (eds.) *Contrast in Phonology: Perception and Acquisition*, Berlin: Mouton de Gruyter.

Hayes, Bruce. 1999. Phonetically-driven phonology: the role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer, and K. Wheatly (eds.) *Functionalism and Formalism in Linguistics, Vol. 1: General Papers*. Amsterdam: John Benjamins. 243-285.

Hyman, Larry. 1993. Problems for rule ordering in phonology: Two Bantu test cases. In John Goldsmith (ed.), *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago: University of Chicago Press, pp. 195-222.

Jesney, Karen. 2005. Chain shift in phonological acquisition. MA thesis. Available at [http://people.umass.edu/kjesney/papers.html]

Kaplan, Aaron. 2011. Harmonic Improvement without Candidate Chains in Chamorro. *Linguistic Inquiry* 42 (4): 631−650.

Kavitskaya, Darya and Peter Staroverov. 2010. When an interaction is both opaque and transparent: the paradox of fed counterfeeding. *Phonology* 27 (2): 255−288.

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

McCarthy, John J. 1999. Sympathy and phonological opacity. ROA-252

McCarthy, John J. 2007. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London, Oakville: Equinox

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: constraint interaction in Generative Grammar. ROA-537

Prince, Alan and BruceTesar. 2004. Learning phonotactic distributions. In Rene Kager, Joe Pater, Wim Zonneveld (eds.) *Constraints in Phonology Acquisition*. Cambridge: Cambridge University press.

Smith, Neilson V. 1973. *The Acquisition of Phonology: a Case Study*. Cambridge: Cambridge University Press.

Smolensky, Paul, Lisa Davidson and Peter Jusczyk. 2004. The initial and final states: theoretical implications and experimental explorations of the Richness of the Base. In Rene Kager, Joe Pater, Wim Zonneveld (eds.) *Constraints in Phonological Acquisition*. Cambridge University Press

Tesar, Bruce. 1995. Computational Optimality Theory. ROA-90.

Tesar, Bruce and Paul Smolensky. 1993. The learnability of Optimality Theory: an algorithm and some basic complexity results. ROA-2.

Wolf, Matthew. 2008. Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model. ROA-996