Building a Phonological Inventory

Veer, B.M. van 't

Citation

Veer, B. M. van 't. (2014, January 7). *Building a Phonological Inventory*. *LOT dissertation series*. Retrieved from https://hdl.handle.net/1887/30242

Author: Veer, B.M. van ‘t
Title: Building a phonological inventory : feature co-occurrence constraints in acquisition
Issue Date: 2015-01-07

CHAPTER 4

Feature Co-occurrence Constraints in Acquisition

4.1 Introduction

In the previous chapter we have explored the Feature Co-occurrence Constraint theory in great detail, and we have seen how the theory enables us to analyse the final state of acquisition of an actual language (Dutch) and what implications follow from the rigorous application of the assumption that the segment inventory is derived solely by a set of feature combinations, and two types of minimalistic feature co-occurrence constraints. While we have seen that features may be innate (section 2.3), in the present chapter we will see that there is little evidence that actual, substance-containing constraints are part of UG (4.6). Rather, we assume that language learners are endowed with constraint templates, which are employed as necessary during acquisition. In this chapter, we will demonstrate the use of Feature Co-occurrence Constraints in language acquisition, specifically the acquisition of the Dutch consonant inventory. We will focus on the system at the level of actual productions of word onsets.

If the actual, feature-referring constraints are not innate but rather constructed from templates and features during acquisition, the question immediately arises at what point in time this happens. The earliest logically possible time is when the second of the two features to which the constraint refers is acquired. The latest logical possibility is anytime after. Of these two options, the first is the only one that can be *a priori* linked to a developmental event; it is hard to predict an event later in acquisition that would trigger the activation of a Feature Co-occurrence Constraint. For this reason, we assume that constraints are actuated no later then when both its features become active in the child's grammar. For every pair of features [F, G], the full set of constraints is activated automatically as soon the second of [F] and [G] is acquired. This is posited to be an automatic process, meaning that *all* constraints involving [F] and [G] are activated, even those whose effects are not seen (these are immediately either demoted or deactivated). Alternatively, we could assume that only those constraints are activated for which the child finds evidence of in the language she is learning. The predictions following from either option with respect to the data are not actually different, but the second option does presuppose more knowledge on the part of the learner: under the first assumption, all possible constraints are evaluated with respect to the data that is in the child's uptake. Under the second option, the learner must evaluate the constraints before they are even activated. Hence, the more logical assumption is to assume that constraint generation is automatic and that constraints for which the child finds no evidence are immediately demoted or deactivated. We will come back to this assumption, especially regarding the question whether constraints are not introduced at a later stage, and see whether it is tenable in section 4.4.

In section 4.2 we will elaborate on the methods that were employed to obtain the data for the current study, and we will discuss some definitional choices that were made (e.g., with respect to developmental stages). Data from one of the children, Jarmo, will serve as an illustration at various points in this section. We shall see that the theory describes his developmental data remarkably well. In section 4.3, we focus on another child, Noortje. Every stage that is specified in her data is treated in detail. At first glance, Noortje's data appear to present some deviations from the theory, but a close examination of these reveals that none of them poses a serious problem. Section 4.4 summarises the main findings, while section 4.5 is devoted to the possible underlying causes for overpredictions, that is, segments that the theory predicts should be present in the inventory at some stage yet remain unattested at that time. Most cases of overpredictions will be found to result from artefacts of data treatment, but some remain. Finally, section 4.7.1 concludes the chapter.

Before we continue, however, let us briefly pause to reflect on the assumptions and predictions of the Feature Co-occurrence Theory.

Most basic assumptions have been noted and motivated in the previous chapter, such as the monovalence of features, the bivalence (and other properties) of Feature Co-occurrence Constraints, the assumption that features may be innate, and the non-specification of coronality and stopness. With respect to the application of the theory to acquisition data, a number of assumptions are added. First of all, we mentioned above that we predict that FCCs are activated at a specific time during development. Furthermore, we will see in section4.6 that we will not assume that the Feature Co-occurrence Constraints are substantively (i.e., specific constraints with individual features) innate. Rather, we will assume that the two types that we use are somehow innate as templates, that learners fill with the features they acquire.

One of the great questions of cognition, one that we will not attempt to solve here, is the question why acquisition proceeds gradually. The child is surrounded by the entire language; what is it that prohibits the entire language from being acquired at once? Or, if we restrict ourselves to the acquisition of the segment inventory, why is there an order of acquisition to begin with? We will discuss a number of theories that have approached the problem of the order of acquisition in the final section of this chapter. Arguments have been proposed based on markedness, frequency and lexical diffusion. The order of acquisition is not our primary concern here, but we must acknowledge that there *is* an order.

Somewhat counterintuitively perhaps, the minimal view of phonology we adopted, which holds that the inventory is epiphenomenal, implies that the inventory is *not* what is acquired. Rather, the child acquires words, which we assume, are analysed in terms of features (contra, for example, Fikkert and Levelt (2008), who propose that featural analysis is not present in the very early stages of development. Whether triggered by frequency, markedness or some other property, the child learns that segments are part of her inventory in a gradual, step-by-step way. She does so, we assume, based solely on positive evidence: only the presence of a segment in the surrounding language can force her to adopt that segment in her inventory. A ban on a given segment (feature combination) cannot follow from lack of exposure to that segment. This is why *all* Feature Co-occurrence Constraints are automatically activated as soon as a feature is acquired; it allows for maximum restrictiveness.

A number of predictions can be made, too, with respect to the application of Feature Co-occurrence Constraint theory acquisition data. For example, in section 3.4, we motivated the non-specifaction of coronality and stopness. This, as was demonstrated, is intimately linked with the use of the constraint types that are part of the current theory, and the ban on positive constraints. The typological prediction is that every inventory has a featurally empty segment; it cannot be banned and hence must receive a phonetic interpretation (*in casu*, /t/). With respect to acquisition, the prediction is that /t/ is the first segment that is acquired (or, more precisely, it is present in the inventory from the first stage on).

A related, more fundamental prediction is a specific instantiation of the Continuity Hypothesis. We do not assume continuity at the surface level, which would entail that the inventory at every stage of development should correspond to an inventory of an existing language (this type of continuity has been shown to exist in the case of the acquisition of syllable types by Levelt, Schiller, and Levelt (1999/2000)). Rather, we predict continuity in the sense that the same type of constraints can be used to model the adult inventory, and the child's inventory alike.

Finally, although it was noted that the current proposal makes no specific predictions with respect to the order of acquisition of features, it does limit the amount of possible acquisition paths by making predictions about possible and impossible feature combinations.

Name	Sex	First session	Last session	Nr. of utterances
Catootje	F	1;10.12	2;7.4	2210
Eva	F	1;4.12	1;11.8	895
Jarmo	M	1;4.18	2;4.1	1544
Noortje	F	1;7.14	2;11.0	1867
Robin	M	1;4;14	2;4.28	2283
Tirza	F	1;6.14	2;6.12	1681
Tom	M	1;0.10	2;2.2	1761

Table 4.1: General information about the selected children

4.2 Methods

The data used to test the theory of Feature Co-occurrence Constraints in development are taken from the CLPF corpus of Dutch child language (Fikkert, 1994; Levelt, 1994). The corpus contains longitudinal recordings of twelve children, all acquiring Dutch. Not all of the children in the CLPF database are represented in the current study, because in order to be able to say something about the consonantal development of the child, we need to have a time-window that does not begin too late, nor end too soon. Seven children are included in the final selection, because they provide enough data and because the recordings took place at a time interval in which we can observe segmental development. The seven children are Catootje, Eva, Jarmo, Noortje, Robin, Tirza and Tom. Table 4.2 gives some general information about the children (based on Table (I) in Levelt (1994)).

The data for the current study was collected from the CLPF database (Fikkert, 1994; Levelt, 1994), as available from PhonBank.[1] The database was loaded and searched in the Phon software application (Rose et al., 2006; Rose & MacWhinney, to appear), using a script designed primarily by Greg Hedlund.[2] In the Phon design structure, the database consists of Projects, Coropora, Sessions and Records. The Project is the entire CLPF database, in which every child represents a Corpus. The Corpora are further subdivided into Sessions, and Sessions into Records. Records contain utterances, consisting of word groups. These are represented on a number of 'Tiers', most important of which for our present purposes is the Actual Tier. This contains the IPA transcription of the utterance; in addition to segmental information, stress and syllabic affiliation and position also are coded.

One of the strongest assets of the Phon interface for phonological databases is its built in search capabilities. The data from the seven children were searched

---

[1] See `http://childes.psy.cmu.edu/phon/`

[2] A free and open source copy of Phon can be obtained from `http://phon.ling.mun.ca/phontrac/wiki/Downloads`

for each of the members of the inventory of Dutch, as given in table 4.2 (see also section 3.2 on the segments and features of Dutch).

	Bilabial	Labiodental	Alveolar	Palatal	Velar
Plosives	p, b		t, d	c	k, (g)
Fricatives		f, v	s, z	ʃ	x
Nasals	m		n		ŋ
Liquids			l, r		
Glides		ʋ		j	

Table 4.2: The consonants of Dutch (after Booij, 1995)

The results of these searches were tabulated in a spread sheet, after which they were processed to exclude false positives, by the following criteria:

1. First, all single occurrences of any segment in any session were deleted
2. Next, any session that contained precisely two occurrences of a given segment, but were followed by at least two sessions with no occurrence of that segment, were discarded with respect to the segment
3. Then, early occurrences of segments were checked to see over how many lexical items the occurrences were distributed. The number of occurrences was replaced by the number of lexical items, to exclude high hits for segments that only occur in single words.[3]
4. Finally, steps 1 and 2 were repeated.

It may seem somewhat indirect to start the search for features by searching for segments, but there is one compelling reason to do this; had the search been performed for individual features, it would have been very difficult – if not impossible – to find restrictions on feature co-occurrences. To illustrate, looking for [labial] returns all labials, without discrimination for which features they co-occur with (or do not). It teaches nothing about co-occurrences; the returned results will have to broken down further. Conversely, simply looking for all labial consonants allows one to know not only when the first labial occurs, but also how the feature combines in various constellations. This advantage outweighs the obvious fact that looking for all labial segments limits the possibilities to those labial segments that are included as search parameters; i.e., the list of segments given in table (4.2) above. It is highly unlikely that the children systematically produce segments that are not in the Dutch inventory, first of all because they are not in the input, and secondly, because the inventory contains relatively unmarked segments. Systematic substitutions are thus likely to be structure preserving.

---

[3] If a child were to say *bal* [bɑl] 'ball' 6 times in a session but no other word starting with [b], only 1 [b] was counted, for example.

A number of exceptions were made to this procedure. Most notably, the search for /r/ was broken down into searches for the feature [rhotic][4] and the segment [ʁ] (which, in the Phon system, is not coded for rhoticity – even though it is a frequent phonetic variant of rhotics). This was done to capture all occurrences of /r/ which, in Dutch, has a wide range of surface shapes; from trill to fricative, from apical to uvular, voiced and voiceless (Sebregts, ms.). The reason why in this case, the search for a feature rather then for individual segments is warranted, is fairly straightforward: underlyingly, [rhotic] only describes one segment. Another exception was made for the dorsal fricative, because of variations in the transcription. This segment is sometimes transcribed as uvular, sometimes as velar; in the latter case, sometimes as voiced, sometimes as voiceless. There are no lexical contrasts between these options. Hence, searches were performed for all of {x, χ, ɣ}, which were then pooled into one segment labeled /x/.

After the procedure outlined above, the spread sheets were rearranged to derive Guttman scales (Torgerson, 1963). Guttman scales are a tool to reveal an order in data; originally, they were designed to test for hierarchical order in response patterns for questionnaires, but they have been shown to be suitable to reveal temporal orders, too, for example, in language acquisition research (Barton, 1976; Fikkert & Levelt, 2008; Levelt et al., 1999/2000). The Guttman scales revealed an order of acquisition for *segments*, but in this thesis segments have no other ontological status then as the epiphenomenal surface reflection of feature co-occurrences. An example of such a scale can be seen below in (39):

(39) Guttman Scale for Jarmo's onsets

---

[4] The features in PhonBank are defined as pre-theoretically as possible, and are somewhat different from the definitions used for our analysis. The feature [rhotic] is not in our feature set, but it is in PhonBank.

Jarmo	k	t	d	p	b	m	ʋ	l	n	f	s	x	r	ʃ
1989-09-22														
1989-10-06														
1989-10-31	2													
1989-11-17	2													
1989-12-01	2	6	7											
1989-12-19	5	8	8	8	2									
1990-01-02	2	9	4	11	4									
1990-01-16	5	9	2	9	5									
1990-01-30	3	9	2	12	6									
1990-02-13	4	7	2	9	6	12								
1990-02-27		2	3	4	8	4								
1990-03-13	10	11	2	8	6	3	2							
1990-03-27	11	9	3	9	9	2								
1990-04-10	7	6		5	2	2	2							
1990-04-24	18	30	4	4	5	3	2	5						
1990-05-08	16	38	8	17	6	7	3	10	5	2				
1990-06-01	24	16	6	10		3	4	5	3	3				
1990-06-12	17	23	6	9	2	4		14		3				
1990-06-26	21	40		15	2	8		12	8		3	3		
1990-07-10	19	32	11	20	8	2	5	9	10		3	4		
1990-07-31	5	15	4	11	6	2		11	4		2	3		
1990-08-13	24	25	13	12	24	14	9	16	16	2	11	3	2	2
1990-09-05	24	10	20	11	12	2	10	22	10	4	11		2	

The Guttman scales list the inventory of each child at each stage. From these segmental inventories, inventories of active distinctive features were derived according to the schema in table 4.3.

	p	b	t	d	c	k	f	v	s	z	ʃ	x	m	n	ŋ	l	r	ʋ	j
[continuant]							+	+	+	+	+	+					+	+	+
[nasal]													+	+	+				
[approximant]																+	+	+	+
[liquid]																+	+		
[voice]		+		+				+		+									
[distributed]					+						+								
[labial]	+	+					+	+					+					+	
[dorsal]						+						+			+				

Table 4.3: Feature specifications for Dutch.

Before turning to the procedure with which the emergence and demotion of the individual Feature Co-occurrence Constraints is determined, let us take a brief moment to discuss the way in which the notion ‘stage’ is used in the current thesis.

4.2.1 Stages

Although language acquisition is a gradient process which proceeds with rushes and delays, it is common practice in language acquisition research to segment the time during which a given phenomenon develops into segments, usually called 'stages'. Two general reasons underly this practice: first, it is extremely difficult if not impossible to collect continuous data, as it implies continuously collecting data. Although diary studies (such as Preyer (1895); Röttger (1931); Leopold (1939 – 1949)) can be seen as approximations of continuous data, these are often limited to data from a single child. Therefore, data is often collected at regular intervals; examples of results of this are the CLPF corpus used in the current study (Fikkert, 1994; Levelt, 1994), or the Goad/Rose corpus of Canadian French (Rose, 2000). The second reason is that it is unclear whether continuous data yields dramatically more insights then interval data, especially if the intervals are short enough.

Even if data are collected at regular intervals, these data points do not automatically coincide with developmental stages. It might be the case that a child's progress at phonological property A stagnates for some months to then rapidly develop to the adult form. In such a case it makes no sense to segment the first period into several stages; at least not with respect to property A. This example also shows that there is no *a priori* need for stages of different aspects of grammatical acquisition to coincide.

There are several ways to define stages. One such way is developed in Ingram (1989). Ingram proposes to segment development into three types of stages: continuous stages, in which change occurs at a steady rate, acceleration stages, in which development proceeds rapidly, and transition stages, at which no change occurs: it represents a transition between two maturational phases. Naturally, the Final State is a special case of a transition stage, even if the terms 'Final' and 'transition' are somewhat at odds with each other. Hence, Ingram mentions a fourth stage, the plateau stage, at which competence has reached the level of the Final Sate and no change is necessary. A typical s-shaped developmental curve, for example, consists of a continuous stage, followed by an acceleration stage, which is succeeded by yet another continuous stage. Although Ingram's stages have proved to be of great use and influence in language acquisition research, they require some type of "goodness" to be measurable. For example, van 't Veer (2012) describes the actual realisations of target liquids by Catootje (one of the subjects in the CLPF corpus, see below), and finds that the realisation of her liquids are scattered over various possibilities (including $\emptyset$) and only slowly improve initially, after which the child goes through a stage of rapid improvement (of the rate of target-to-non-target substitutions). Finally, when production is nearly always correct, the rate of improvement drops to a lower pace – the final continuous stage.

In the current study, stages are defined differently. For one thing, the rate of acquisition is not our primary concern, and furthermore, stages defined on the basis of the rate of development are not fine-grained enough. To determine a rate in the first, place, a stage needs to contain at least two different changes in the child's observable size (for example, additions of segments to the inventory). We are interested in the changes *per se*, and hence it was decided to define a stage as the period of one recording session up to but not including the next recording session in which the inventory has grown. This allows us to discard data points at which no change occurs, yet gives us a fine-grained enough look at development.[5]

4.2.2 Finding FCCs

To find the relevant Feature Co-occurrence Constraints, the featural inventories as described above are fed into the algorithm described in section 3.2.3 above. The steps of the algorithm are briefly reproduced below. The definitions of the FCCs are reproduced in (40).

(40) a. *FG
assign a violation mark for every segment Σ iff [F] is dominated by Σ and [G] is dominated by Σ (*c-constraint*)

b. F→G
assign a violation mark for every segment Σ iff [F] is dominated by Σ and [G] is not in Σ(*i-constraint*)

The first step in the algorithm is to list the features that are active for the current stage. Based on the segments in the inventory and the data in table (4.3), every combination of two features in the matrix is assigned "–" where both features are the same, "0" where the combination is not attested, and "1" where it is. As every cell containing "0" refers to a combination of two features that is logically possible in the phonological grammar at that stage, and yet not legal, that combination is listed as a c-constraint.

Next, a list containing every possible feature combination at the current stage is compared to every cell containing "1", where the feature heading the row is [F], and the feature heading the column is [G]. If [G] is present in every segment that contains [F], [F]→[G] is satisfied. If a segment is found that contains [F] but not [G], [F]→[G] is listed as being violated. Finally, a checking procedure is included to list any over- and undergenerated segments; overgenerated segments are attested segments that are not ruled out by the FCCs that the algorithm found, whereas undergenerated segments are feature combinations that are ruled out, but which are attested.

The result of this procedure is that we have, for each stage of every child, a list of features that are acquired, a list of violated constraints, and lists of over- and undergenerated segments. Ideally, the latter two lists are empty; the attested inventory is then congruent with the state of the grammar. It should be noted that each stage is treated individually by the algorithm, even if this is not the case for the child (who, while having no knowledge of the future, builds on knowledge already acquired). Hence, we can make a number of predictions: if the algorithm that is proposed here is a good model of acquisition, it will

---

[5] Note, too, that stages are always defined with respect to a specific trait or behaviour. In our case, it is the development of segmental material in the word-onset position.

converge, from one stage to the next, on those parts of the inventory that do not change. In other words, gaps that remain in the inventory remain there because of the same constraints.

It is worthwhile to pause for a moment and consider the limits of the procedure; the type of scenario that is predicted to be impossible. For one thing, the theory cannot rule out segments for which it is necessary to posit co-occurrence restrictions for more than two features. In this case, the theory predicts more segments than are attested. Related to this issue is that no conditional restrictions can be posed. It is impossible to ban, for example, [continuant] combining with [labial], but only when [voice] is involved:*[cont, lab, voice]. Such a constraint would be necessary in the case where /f/, /b/ and /z/ are legal but not /v/. In other words, the following (sub-)inventory is predicted not to occur:

(41)	b	f	*v
			z

The segment /v/ has precisely the three features given as an example above: it is [cont, lab, voice]. In terms of c-constraints, there are three possibilities to rule it out:

(42) *[cont, lab]
*[lab, voice]
*[cont, voice]

The first constraint in 42 is violated by /f/, the second by /b/ and the third by /z/. With all of the relevant c-constraints independently violated by other segments, none can be used to exclude /v/.[6] Given the active features in our example inventory, the following i-constraints are active:

(43) [cont]→[lab]
[lab]→[cont]
[lab]→[voice]
[voice]→[lab]
[cont]→[voice]
[voice]→[cont]

Here, /z/ violates the first and fourth constraint in 43, as it is not a labial. The second constraint is violated by /b/, as is the sixth, a non-continuant. The third and fifth constraints are violated by /f/, because it is not voiced. Again, we see that each available constraint is violated by a segment other than /v/, and hence, it cannot be ruled out. In section 4.5 below, we will come back in more detail to this type of impossible inventory, which is related to set-subset relations of feature combinations (segments). For now, however, it should be noted that we predict that no 'conditional gaps' occur in the data.

---

[6] Note that constraint conjunction as proposed in OT would allow a constraint *[cont, lab]&*[lab, voice]. Constraint conjunction is an incredibly powerful tool, and we shall not employ it here.

In the introduction, we adopted an assumption that constraints over [F,G] are activated as soon as the second member of the pair F and G is acquired, an assumption that we will see below is not contradicted. In the face of positive evidence, that is to say, after having learned that the segment is part of the surrounding language, the child can revoke constraints immediately or later during development. However, although constraints impose co-occurrence restrictions, it is possible to describe a situation where segment A is allowed at stage S, but no longer so at S+1. The way in which this can be done is highly limited however. Let us first consider the type of shrinking inventory that *cannot* be described, using the following scenario:

(44) Impossible scenario: the shrinking inventory (1)

Stage	*Features*	*Inventory*
S-1	[F]	/f/
S	[F], [G]	/f/, /f͜g/ /g/
S+1	[F], [G]	/f/, /g/

In this scenario, stage S heralds the acquisition of the feature [G]. This feature is uninhibited in its combinations: it is expressed in the segment /g/ as well as in /f͜g/. At stage S+1, it is no longer permitted; to account for this, the constraint *[F, G] must be activated. This is not possible, however, as both [F] and [G] were acquired before stage S+1. The second of the two, [G], was acquired at stage S, which means that the constraint *[F, G] was activated at that same stage, but demoted/de-activated immediately, due to the presence of /f͜g/ in the inventory.[7]

The same argument holds with respect to i-constraints. Consider the marginally different scenario below:

(45) Impossible scenario: the shrinking inventory (2)

Stage	*Features*	*Inventory*
S-1	[F]	/f/
S	[F], [G]	/f/, /f͜g/ /g/
S+1	[F], [G]	/f/, /f͜g/

Here, it is the segment /g/ that is lost after it is acquired. To account for this, a constraint [G]→[F] must be activated, but we run into the same timing problem as we did before.

The only possible way to describe a inventory that disallows a segment after allowing it at an earlier stage is when three features are involved:

---

[7] It should be noted that under most OT theories of acquisition constraints are never promoted, only demoted. Moving *[F, G] up in the hierarchy is no option.

(46) Possible scenario: the shrinking inventory

Stage	*Features*	*Inventory*	*Active constraints*
S-1	[F]	/f/	
S	[F], [G]	/f/, /f͜g/	[G]→[F]
S+1	[F], [G], [I]	/f/, /g͜i/	[G]→[I]

Here, /g/ is ruled out by [G]→[F] at stage S. At S+1, a new feature is acquired: [I]. It is only present in one segment, /g͜i/. This means that the constraint [G]→[F] must be revoked: it is violated by the new segment. However, the non-existence of /i/ means that a new constraint, [G]→[I], is introduced. This constraint bans the previously legal segment /f͜g/.

In practice, this situation is not encountered because of the way the Guttmann scales are interpreted: gaps are interpreted as accidental in the sense that they are considered artefacts of the sampling methods. Once a segment passes the inclusion criteria described above, it is considered to be acquired permanently. It would be interesting to see whether the gaps found in the Guttmann scales conform to the predictions outlined here, but that is beyond the scope of the current thesis.

Before turning to a detailed exploration of the acquisition of an individual child, we must consider two additional aspects of the proposed theory: constraint inactivity and constraint redundancy.

Let us consider the inventory of Jarmo at early stages.[8] As can be seen in (47a), Jarmo quickly acquires the full range of places of articulation (by stage 3, every major PoA is represented). At stage 4, /m/ is acquired, and with it, the feature [nasal] (see 47b). This is the only nasal for some time, however; only at stage 7 is it accompanied by /n/. The acquisition of the feature [nasal] triggers the activation of two constraints: *[nasal, dorsal], and [nasal]→[labial] (table 4.4 lists all constraints that are active in Jarmo's development). The effect of the first is to ban /ŋ/, whereas the second bans /n/. Since we are looking at (word) onsets only, it is to be expected that the former constraint remains active; this is indeed the case (see table 4.4). At stage 7, however, something must change; /n/ is now an admissible segment in Jarmo's phonological grammar. For this reason, [nasal]→[labial] must be revoked, or, in OT-terms, demoted to a place where it no longer has any influence. The important thing to remember here is that, while OT-type grammars offer a readymade solution for rendering constraints less influential (demotion), a mechanism for constraint de-activation (be it partial or complete) is a necessity for every theory that aims to combine constraints and acquisition data: children's grammars are simply more restrictive at some stages than they might be at a later stage.

(47) a. Jarmo's inventory of segments

---

[8] The inventories, acquired features and active constraints for each child are listed in Appendix C

Stage	*Inventory*													
1	k													
2	k	t	d											
3	k	t	d	p	b									
4	k	t	d	p	b	m								
5	k	t	d	p	b	m	ʋ							
6	k	t	d	p	b	m	ʋ	l						
7	k	t	d	p	b	m	ʋ	l	n	f				
8	k	t	d	p	b	m	ʋ	l	n	f	s	x		
9	k	t	d	p	b	m	ʋ	l	n	f	s	x	r	ʃ

b. Jarmo's inventory of features

Stage	*Features*
1	[dors]
2	[dors], [voice]
3	[dors], [voice], [lab]
4	[dors], [voice], [lab], [nas]
5	[dors], [voice], [lab], [nas], [cont], [apprx]
6	[dors], [voice], [lab], [nas], [cont], [apprx], [liq]
7	[dors], [voice], [lab], [nas], [cont], [apprx], [liq]
8	[dors], [voice], [lab], [nas], [cont], [apprx], [liq]
9	[dors], [voice], [lab], [nas], [cont], [apprx], [liq], [dist]

The example of Jarmo's nasals is useful to illustrate another aspect of the theory as it is proposed here. At stage 4, when Jarmo acquired the feature [nasal] used only in the segment /m/, two constraints are activated, as we have seen: [nasal]→[labial] to ban /n/ and *[nasal, dorsal] to prevent /ŋ/ from surfacing as a legal segment. The reader may have noticed, however, that the former constraint in fact does both jobs: by requiring [nasal] to only co-occur with [labial], it effectively bans both /n/ *and* /ŋ/. Yet, an additional constraint is activated. At first glance this may seem like an unnecessary complication, but it is important to consider that the child has no knowledge of the future. If only [nasal]→[labial] were to be activated at stage 4, there learner would be forced to reanalyse the inventory and posit *[nasal, dorsal] after all. We assumed that all FCCs would be acquired no later than the stage at which its second feature is acquired (see also section 4.4); if constraints were not redundantly activated, this position would not be tenable.

With this in hand, we can now investigate the acquisition of the segment inventory, focussing on the word-onset. Although it is interesting to consider other positions, as well, the word-onset is usually the first consonantal slot in which segmental knowledge is acquired. To illustrate, the development of the coda starts at a later point, when the child already has some knowledge of sub-segmental phonology; hence, the development of other positions is 'contaminated' by earlier knowledge. In addition, not many children in our database provide a reasonable amount of data from other syllable positions, acquiring only a very limited inventory during the time window of the recordings. For

Stage	1	2	3	4	5	6	7	8	9
Constraints					[apprx]→[cont]	[apprx]→[cont]	[apprx]→[cont]	[apprx]→[cont]	[apprx]→[cont]
					[apprx]→[lab]	[apprx]→[lab]	[apprx]→[lab]	[apprx]→[lab]	[apprx]→[lab]
					[cont]→[apprx]	[cont]→[apprx]			
					[cont]→[lab]	[cont]→[lab]	[cont]→[lab]		
									[dist]→[cont]
									*[apprx, dist]
				[nas]→[lab]	[nas]→[lab]	[nas]→[lab]			
						*[apprx, liq]	*[apprx, liq]	*[apprx, liq]	*[apprx, liq]
						*[cont, liq]	*[cont, liq]	*[cont, liq]	
					*[dors, apprx]	*[dors, apprx]	*[dors, apprx]	*[dors, apprx]	*[dors, apprx]
					*[dors, cont]	*[dors, cont]	*[dors, cont]		
									*[dors, dist]
			*[dors, lab]	*[dors, lab]	*[dors, lab]	*[dors, lab]	*[dors, lab]	*[dors, lab]	*[dors, lab]
						*[dors, liq]	*[dors, liq]	*[dors, liq]	*[dors, liq]
				*[dors, nas]	*[dors, nas]	*[dors, nas]	*[dors, nas]	*[dors, nas]	*[dors, nas]
		*[dors, voice]	*[dors, voice]	*[dors, voice]	*[dors, voice]	*[dors, voice]	*[dors, voice]	*[dors, voice]	*[dors, voice]
									*[lab, dist]
						*[lab, liq]	*[lab, liq]	*[lab, liq]	*[lab, liq]
									*[liq, dist]
					*[nas, apprx]	*[nas, apprx]	*[nas, apprx]	*[nas, apprx]	*[nas, apprx]
					*[nas, cont]	*[nas, cont]	*[nas, cont]	*[nas, cont]	*[nas, cont]
									*[nas, dist]
						*[nas, liq]	*[nas, liq]	*[nas, liq]	*[nas, liq]
					*[voice, apprx]	*[voice, apprx]	*[voice, apprx]	*[voice, apprx]	*[voice, apprx]
					*[voice, cont]	*[voice, cont]	*[voice, cont]	*[voice, cont]	*[voice, cont]
									*[voice, dist]
						*[voice, liq]	*[voice, liq]	*[voice, liq]	*[voice, liq]
				*[voice, nas]	*[voice, nas]	*[voice, nas]	*[voice, nas]	*[voice, nas]	*[voice, nas]

Table 4.4: Feature Co-occurrence Constraints in Jarmo's Actual word onset productions

these reasons, we will limit the discussion to the word onset position.

4.3 Constraints in the child's grammar

In the previous section, we have illustrated some of the aspects of the current theory with data from one child, Jarmo. His data represents an almost ideal case, in which all predictions are borne out (save for the overprediction of /t/ in the first stage - but see below). In this section, we will examine every stage of the development of another child, Noortje. She is selected because her data present the most cases of apparent anomalies. As we will see however, the issues that seem to arise will all be resolved under closer inspection.

4.3.1 Noortje

Stage 1

Inventory /m/
Features [nasal], [labial]
Constraints

Active constraints	*Inactive constraints*
[nas]→[lab]	*[nas,lab]
[lab]→[nas]	

In Noortje's first stage, her onset inventory consists only of [m], for which two features are necessary: [nasal] and [labial]. Two constraints are activated: [nas]→[lab] and the reverse, [lab]→[nas]. The former rules out a segment consisting of just [nasal], which would be interpreted as /n/. The second constraint rules out a segment [labial], which corresponds to /p/. Finally, we see that it is not possible to rule out the empty segment (interpreted as /t/). We will explore the issue of overprediction in full detail in section 4.5.

Stage 2

Inventory /m, p, t, k/
Features [nasal], [labial], [dorsal]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[dors, lab]	[nas]→[lab]
[lab]→[nas]	[nas]→[dors]
	[dors]→[nas]
	[lab]→[dors]
	[dors]→[lab]

At the second stage, Noortje's inventory grows substantially, adding the voiceless stops to the set. As we mentioned above, /t/ was already predicted to be in the inventory. The feature set in stage 1 is already sufficient to allow for /p/, the only thing that needs to change is the deactivation (or demotion) of the constraint requiring labials to be nasal: [lab]→[nas] is revoked. Finally, [dorsal] is added to allow for /k/, but it is accompanied by two constraints to restrict its combinatorics: *[lab, dors] prevents doubly articulated segments and *[nas, dors] prevents /ŋ/. Since we are considering the onset inventory only, both constraints will remain active in the grammar.

Stage 3

Inventory /m, p, t, k, n/
Features [nasal], [labial], [dorsal]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[dors, lab]	[nas]→[lab]
	[lab]→[nas]
	[nas]→[dors]
	[dors]→[nas]
	[lab]→[dors]
	[dors]→[lab]

Stage 3 heralds the addition of /n/. The feature needed to produce this segment, [nasal], is already in the inventory. By revoking [nas]→[lab], /n/ is now a legal segment.

Stage 4

Inventory /m, p, t, k, n, b/
Features [nasal], [labial], [dorsal], [voice]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[nas, voice]	*[lab, voice]
*[lab, dors]	[nas]→[lab]
*[dors, voice]	[lab]→[nas]
[voice]→[lab]	[nas]→[dors]
	[dors]→[nas]
	[nas]→[voice]
	[voice]→[nas]
	[lab]→[dors]
	[dors]→[lab]
	[lab]→[voice]
	[dors]→[voice]
	[voice]→[dors]

Voicing is introduced at stage 4, but it is restricted to occurring with [labial]. Hence, three new constraints are activated: [voice]→[lab], *[dors, voice], and *[nas, voice]. The i-constraint forces [voice] to be with [labial], banning both /d/ and /ɣ/. One might think, then, that the constrain *[dors, voice] is unnecessary, and in fact, it is redundant. However, the requirement that only labials be voiced must be revoked at some point to allow for /d/, and it is important to note that the learner has little or no way of knowing what will happen in the future. Hence, the most prudent strategy is to activate constraints regardless of whether they are redundant, or not. It should be noted that this strategy is also in accordance with the subset principle, which can be paraphrased to hold that the most restrictive grammar should be preferred by the learner.

Stage 5

Inventory /m, p, t, k, n, b, d/
Features [nasal], [labial], [dorsal], [voice]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[nas, voice]	*[lab, voice]
*[lab, dors]	[nas]→[lab]
*[dors, voice]	[lab]→[nas]
	[nas]→[dors]
	[dors]→[nas]
	[nas]→[voice]
	[voice]→[nas]
	[lab]→[dors]
	[dors]→[lab]
	[lab]→[voice]
	[voice]→[lab]
	[dors]→[voice]
	[voice]→[dors]

At stage 5, this is indeed what happens: [voice]→[lab] is revoked or demoted, because /d/ is added to the inventory. /ɣ/ is still ruled out by *[dors, voice], and continues to be so.

Stage 6

Inventory /m, p, t, k, n, b, d, s/
Features [nasal], [labial], [dorsal], [voice], [continuant]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[nas, voice]	*[lab, voice]
*[nas, cont]	[nas]→[lab]
*[lab, dors]	[lab]→[nas]
*[lab, cont]	[nas]→[dors]
*[dors, voice]	[dors]→[nas]
*[dors, cont]	[nas]→[voice]
*[voice, cont]	[voice]→[nas]
	[nas]→[cont]
	[cont]→[nas]
	[lab]→[dors]
	[dors]→[lab]
	[lab]→[voice]
	[voice]→[lab]
	[lab]→[cont]
	[cont]→[lab]
	[dors]→[voice]
	[voice]→[dors]
	[dors]→[cont]
	[cont]→[dors]
	[voice]→[cont]
	[cont]→[voice]

No constraints are revoked at stage 6, but [continuant] is acquired because /s/ is added to the inventory. Since no other continuants are in the inventory (yet), the constraints *[dors, cont], *[lab, cont], *[voice, cont] and *[nas, cont] remain activated. Again, one of these constraints, *[nas, cont], will remain active in the mature grammar.

Stage 7

Inventory /m, p, t, k, n, b, d, s, ʋ/
Features [nasal], [labial], [dorsal], [voice], [continuant], [approximant]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[nas, voice]	*[lab, voice]
*[nas, cont]	*[lab, cont]
*[nas, apprx]	*[lab, apprx]
*[lab, dors]	*[cont, apprx]
*[dors, voice]	[nas]→[lab]
*[dors, cont]	[lab]→[nas]
*[dors, apprx]	[nas]→[dors]
*[voice, cont]	[dors]→[nas]
*[voice, apprx]	[nas]→[voice]
[apprx]→[lab]	[voice]→[nas]
[apprx]→[cont]	[nas]→[cont]
	[cont]→[nas]
	[nas]→[apprx]
	[apprx]→[nas]
	[lab]→[dors]
	[dors]→[lab]
	[lab]→[voice]
	[voice]→[lab]
	[lab]→[cont]
	[cont]→[lab]
	[lab]→[apprx]
	[dors]→[voice]
	[voice]→[dors]
	[dors]→[cont]
	[cont]→[dors]
	[dors]→[apprx]
	[apprx]→[dors]
	[voice]→[cont]
	[cont]→[voice]
	[voice]→[apprx]
	[apprx]→[voice]
	[cont]→[apprx]

The first glide appears at stage 7, when the feature [approximant] is acquired. Severe co-occurrence restrictions hold for this feature, with c-constraints being activated for [approximant] with nearly every other feature that has been acquired up to now: *[dors, apprx], *[nas, apprx], *[voice, apprx]. This does not suffice to rule out /j/ [cont, apprx], however, so the i-constraint [apprx]→[lab] is activated as well. Approximants must be continuant, and so we see the activation of [apprx]→[cont].

One constraint must be revoked in order to allow for /ʋ/, which has the features [lab, cont, apprx], to be in the inventory: *[lab, cont]. This results in another overprediction, namely that of /f/, which is [lab, cont].

Stage 8

Inventory /m, p, t, k, n, b, d, s, ʋ, f/
Features [nasal], [labial], [dorsal], [voice], [continuant], [approximant]
Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[nas, voice]	*[lab, voice]
*[nas, cont]	*[lab, cont]
*[nas, apprx]	*[lab, apprx]
*[lab, dors]	*[cont, apprx]
*[dors, voice]	[nas]→[lab]
*[dors, cont]	[lab]→[nas]
*[dors, apprx]	[nas]→[dors]
*[voice, cont]	[dors]→[nas]
*[voice, apprx]	[nas]→[voice]
[apprx]→[lab]	[voice]→[nas]
[apprx]→[cont]	[nas]→[cont]
	[cont]→[nas]
	[nas]→[apprx]
	[apprx]→[nas]
	[lab]→[dors]
	[dors]→[lab]
	[lab]→[voice]
	[voice]→[lab]
	[lab]→[cont]
	[cont]→[lab]
	[lab]→[apprx]
	[dors]→[voice]
	[voice]→[dors]
	[dors]→[cont]
	[cont]→[dors]
	[dors]→[apprx]
	[apprx]→[dors]
	[voice]→[cont]
	[cont]→[voice]
	[voice]→[apprx]
	[apprx]→[voice]
	[cont]→[apprx]

In stage 8 the overprediction introduced in stage 7 is resolved: the segment /f/ is now in the attested inventory. No new features are added, and therefore no constraints are introduced, nor is any constraint revoked.

Stage 9

Inventory /m, p, t, k, n, b, d, s, ʋ, f, l, j, x/
Features [nasal], [labial], [dorsal], [voice], [continuant], [approximant] , [liquid]
Constraints (see next page)
At the final stage within the time span of the recordings, Noortje has not acquired every segment yet. Notably missing are voiced fricatives and /r/. The other liquid, /l/, is acquired at this point. This entails the acquisition of the feature [liquid], which may not co-occur with any of the place features (*[dors, liq], *[lab, liq]), nor with any of the other manner features (*[apprx, liq], *[cont, liq], where the latter rules out /r/), nor with voice (*[voice, liq]). *[dors, cont] is lifted to allow for /x/, and the revocation of [apprx]→[lab] accounts for the acquisition of /j/.

Constraints

Active constraints	*Inactive constraints*
*[nas, dors]	*[nas, lab]
*[nas, voice]	*[lab, voice]
*[nas, cont]	*[lab, cont]
*[nas, apprx]	*[lab, apprx]
*[nas, liq]	*[dors, cont]
*[lab, dors]	*[cont, apprx]
*[lab, liq]	[nas]→[lab]
*[dors, voice]	[lab]→[nas]
*[dors, apprx]	[nas]→[dors]
*[dors, liq]	[dors]→[nas]
*[voice, cont]	[nas]→[voice]
*[voice, apprx]	[voice]→[nas]
*[voice, liq]	[nas]→[cont]
*[cont, liq]	[cont]→[nas]
*[apprx, liq]	[nas]→[apprx]
[apprx]→[cont]	[apprx]→[nas]
	[nas]→[liq]
	[liq]→[nas]
	[lab]→[dors]
	[dors]→[lab]
	[lab]→[voice]
	[voice]→[lab]
	[lab]→[cont]
	[cont]→[lab]
	[lab]→[apprx]
	[apprx]→[lab]
	[lab]→[liq]
	[liq]→[lab]
	[dors]→[voice]
	[voice]→[dors]
	[dors]→[cont]
	[cont]→[dors]
	[dors]→[apprx]
	[apprx]→[dors]
	[dors]→[liq]
	[liq]→[dors]
	[voice]→[cont]
	[cont]→[voice]
	[voice]→[apprx]
	[apprx]→[voice]
	[voice]→[liq]
	[liq]→[voice]
	[cont]→[apprx]
	[cont]→[liq]
	[liq]→[cont]
	[apprx]→[liq]
	[liq]→[apprx

Having now fully explored the developmental path of Noortje's onset inventory in terms of features and Feature Co-occurrence Constraints, let us investigate how she and Jarmo, whose data we examined in section 4.2, compare to the other children in the database.

4.3.2 Constraints: Noortje, Jarmo and the other children

The total number of constraints employed by Noortje is far lower than the number that is logically possible (for nine features, the number of possible constraints is 1.5n(n-1) = 108); the theory is thus highly economic in terms of the number of constraints that remain activated. Table 4.5 lists, for the actual productions, all constraints that are active in the current study, and the children in whose grammars the constraints are activated.

It can be seen that there is a high degree of overlap between the constraints that the children are using, especially in the set of c-constraints. Comparing the children's constraints to the ones active in Dutch, we see that there are some that are active in the children's grammars, but not in the Adult grammar; these are transient constraints that will be revoked or demoted at some point after our window on development ends. More interestingly, there are also constraints that are active in Dutch, but not in the children's grammars. First of all, it is important to note that there are no constraints that are exclusive to the mature grammar. Secondly, the vast majority of constraints in this category involve the feature [distributed]. These constraints are only active in the grammars of Jarmo and Tirza. This is because Jarmo and Tirza are also the only two children who have acquired the feature by the end of the developmental window.[9]

Given that we have nine features, and that for every combination of two features, three constraints are possible (one c-constraint, and two i-constraints), the number of possible constraints is quite large (108). In chapter 3, we saw that for adult Dutch, only 19 FCCs are required to be active: 18 c-constraints and one i-constraint.These are represented in the final column in table 4.5 repeated in (48) below:

(48) a. C-CONSTRAINTS
*[lab, dors]
*[lab, dist]
*[lab, liq]
*[voice, dors]
*[voice, dist]
*[voice, nas]
*[voice, liq]
*[voice, apprx]
*[dors, dist]
*[dors, liq]
*[dors, apprx]

---

[9] The one other constraint that is in the mature grammar but not in all of the children's, is *[liq, apprx], which only Catootje, Tirza and Tom have. It should be noted that this constraint is formally and functionally equivalent to *[apprx, liq], which is in all of the children's systems. The fact that some children have a redundant formalisation of this co-occurrence restriction, whereas others have only a single constraint, reflects a difference in the order of acquisition of both features. Catootje, Tirza and Tom, who have both versions, acquire [liquid] before [approximant], whereas the other children have the opposite acquisition order.

	Catootje	*Eva*	*Jarmo*	*Noortje*	*Robin*	*Tirza*	*Tom*	*Adult Dutch*
[apprx]→[cont]	✓	✓	✓	✓	✓	✓	✓	✓
[apprx]→[lab]			✓	✓	✓		✓	
[cont]→[apprx]	✓	✓	✓					
[cont]→[lab]			✓				✓	
[dist]→[cont]			✓			✓		
[lab]→[voice]					✓			
[nas]→[lab]			✓	✓				
[voice]→[lab]	✓					✓	✓	
*[apprx, dist]			✓			✓		✓
*[apprx, dors]	✓	✓	✓	✓	✓	✓	✓	✓
*[apprx, liq]	✓	✓	✓	✓	✓	✓	✓	✓
*[cont, dors]	✓	✓	✓	✓	✓	✓		
*[cont, liq]	✓	✓	✓	✓	✓	✓	✓	
*[dors, dist]			✓			✓		✓
*[dors, lab]	✓	✓	✓	✓	✓	✓	✓	✓
*[dors, liq]	✓	✓	✓	✓	✓	✓	✓	✓
*[dors, nas]	✓	✓	✓	✓	✓	✓	✓	
*[dors, voice]	✓	✓	✓	✓	✓	✓	✓	✓
*[lab, apprx]	✓					✓		
*[lab, cont]	✓			✓	✓	✓		
*[lab, dist]			✓			✓		✓
*[lab, liq]	✓	✓	✓	✓	✓	✓	✓	✓
*[lab, nas]		✓			✓			
*[liq, apprx]	✓					✓	✓	✓
*[liq, dist]			✓			✓		✓
*[nas, apprx]	✓	✓	✓	✓	✓	✓	✓	✓
*[nas, cont]	✓	✓	✓	✓	✓	✓	✓	✓
*[nas, dist]			✓			✓		✓
*[nas, liq]	✓	✓	✓	✓	✓	✓	✓	✓
*[nas, voice]	✓	✓	✓	✓	✓	✓	✓	✓
*[voice, apprx]	✓	✓	✓	✓	✓	✓	✓	✓
*[voice, cont]	✓	✓	✓	✓	✓	✓	✓	
*[voice, dist]			✓			✓		✓
*[voice, liq]	✓	✓	✓	✓	✓	✓	✓	✓

Table 4.5: Constraints activated by the children in the current study

*[cont, nas]
*[dist, nas]
*[dist, liq]
*[dist, apprx]
*[nas, liq]
*[nas, apprx]
*[liq, apprx]

b. I-CONSTRAINTS
[apprx]→[cont]

Table 4.6 condensed the data in table 4.5 by listing the number of different constraints activated per child. In some cases, the number for the children is smaller than the number of constraints for the adult grammar, but the opposite is the more frequent situation. The list in (48) concerns the inventory at all positions, whereas for the children, we are only considering word onsets exclusively. /ŋ/, for example, is in the inventory of Dutch, but it is not allowed in onsets. To account for this difference, the adult grammar has one constraint *less*: we find *[nas, dors] in the grammars of every child (and at every level), but it is not part of the list in (48).

More importantly, a look at the inventories in appendix C reveals that by the end of the recording sessions, not all children have acquired the full inventory (see also our discussion of Noortje's inventory, above). The two largest numbers in table 4.6 are found for the inventories of Jarmo and Tirza. These are also the only two inventories where the feature [distributed] is acquired (see above). This feature causes seven of the constraints in (48), its absence in most of the children's inventories accounts for much of the observed difference in the amount of constraints.

	Constraints
Catootje	22
Eva	19
Jarmo	28
Noortje	21
Robin	21
Tirza	27
Tom	20

Table 4.6: Number of different constraints used per child

In table 4.7, we see that there are 34 constrains in the current study. Appendix B presents the constraints per child, in the same fashion as table 4.5 does. Of interest are those constraints that do not remain active, that is, those that are unique to the children's grammars. One possible way in which the method developed here can be used is to study temporary constraints in child language, in a way similar to Fikkert and Levelt (2008). Example (49) lists the constraints that are unique to the inventories of the children studied here.

(49) a. C-CONSTRAINTS
*[cont, dors]
*[cont, liq]
*[dors, nas]
*[lab, apprx]
*[lab, cont]
*[lab, nas]
*[voice, cont]

b. I-CONSTRAINTS
[apprx]→[lab]
[cont]→[apprx]
[cont]→[lab]
[dist]→[cont]
[lab]→[voice]
[nas]→[lab]
[voice]→[lab]

The first thing that we see is that while for the mature inventory only a single i-constraint is needed, the children use i-constraints far more frequently. This is likely to be due to the nature of FCCs. For every feature, c-constraint specify a single co-occurrence restriction, but i-constraints ban every possible restriction *except* for the specified co-occurrence. Hence, i-constraints are much more powerful, and perhaps too powerful for a full-fledged inventory.

It is worth repeating, furthermore, that the number of constraints is much smaller than it would be, if every possible constraint were activated and subsequently ranked. It is unclear whether there is, at this point, any evidence that suggests a correlation between the number of constraints in a grammar and its efficiency, fitness, learnability or any other measure of goodness (note that in the original conception of OT, there is no pressure to keep the number of constraints to a minimum; the goal of factorial typologies is to reduce the number of ad-hoc constraints, not the number of constraints *per se*.

4.3.3 Summary

In this section, we have seen a demonstration of Feature Co-occurrence Constraints in acquisition, by carefully observing the development of Noortje's inventory of onset productions. In addition, we have seen that over all, only a small number of possible FCCs become activated. I take this to be a good

	Active
[apprx]→[cont]	✓
[apprx]→[lab]	✓
[cont]→[apprx]	✓
[cont]→[lab]	✓
[dist]→[cont]	✓
[lab]→[voice]	✓
[nas]→[lab]	✓
[voice]→[lab]	✓
*[apprx, dist]	✓
*[apprx, dors]	✓
*[apprx, liq]	✓
*[cont, dors]	✓
*[cont, liq]	✓
*[dors, dist]	✓
*[dors, lab]	✓
*[dors, liq]	✓
*[dors, nas]	✓
*[dors, voice]	✓
*[lab, apprx]	✓
*[lab, cont]	✓
*[lab, dist]	✓
*[lab, liq]	✓
*[lab, nas]	✓
*[liq, apprx]	✓
*[liq, dist]	✓
*[nas, apprx]	✓
*[nas, cont]	✓
*[nas, dist]	✓
*[nas, liq]	✓
*[nas, voice]	✓
*[voice, apprx]	✓
*[voice, cont]	✓
*[voice, dist]	✓
*[voice, liq]	✓

Table 4.7: Constraints active in the current study

thing, and even though the procedure by which constraints are derived results in some redundant constraints, the number of constraints remains very manageable. In the rest of this chapter, we will go deeper into some of the findings illustrated by Noortje above.

4.4 General observations

Based on the discussions above (the data for every child are given in full in appendix C), a number of general observations can be made concerning the assumptions and predictions discussed in the introduction to this chapter.

4.4.1 Continuity

A fundamental prediction we made at the beginning of this chapter is that the inventory at each stage of development should correspond to a possible adult inventory. This does not necessarily have to hold at the surface level (especially since early child inventories are often smaller than most adult inventories), but the underlyingly, the structure should follow the same principles. These principles, we assumed, are modeled in terms of monovalent features, i-constraints and c-constraints. We have seen that this prediction holds true with only a very limited number of exceptions.

4.4.2 Order of acquisition

Since a segment that has no features (an empty root node, for example) cannot be banned by Feature Co-occurrence Constraints, it was predicted that /t/ is present from the very first stage on. This prediction is borne out with the exception of Noortje, whose first stage contains only /m/. Furthermore, related to the discussion on the Continuity Hypothesis above, the Feature Co-occurrence Constraint can rule out certain impossible orders of acquisition (e.g., those leading to conditional gaps. The theory makes no claim as to determining the order of acquisition of individual features, but given the acquisition of those features, some combinations are ruled out.

4.4.3 Timing

In the introduction to this chapter, the assumption was adopted that constraints would emerge exactly at the point in time at which the second of the two features to which it refers is acquired. This assumption is proves to be tenable; no counterexamples were encountered. This raises the question whether what we are dealing with is in fact not the confirmation of an independent assumption, but rather an artefact resulting from some mechanism in the methodology. Such an artefact could arise if the algorithm by which constraints are uncovered had built-in to it some device to explicitly ban the introduction of constraints at any other stage than the one at which the second of both features is introduced, or, alternatively, if it had some way of knowing what constraints would be necessary at later stages.

The important point is that the algorithm has no memory beyond a single stage; each stage is treated as if it were the inventory of an individual language. Rather than evaluating what is different, the algorithm approaches each individual inventory anew. This approach ensures that the algorithm has no knowledge of what constraints were active at previous stages, and hence, it cannot know whether a constraint it activates is new or not. This, in turn, means that the timing of the activation of the constraints is not an artefact of the algorithm. An additional implication of the fact that no new constraints are created without the introduction of a new feature is a further reinforcement of the robustness of the FCC approach.

4.4.4 Converging stages

In this chapter, we approached the acquisition of the segment inventory as the successive application of the constraint generating algorithm introduced in chapter 3. This means that the model diverges from actual acquisition, as children do not likely reanalyse the structure of their inventory each time a new segment (feature combination) is added (but see section 5.2.3). At the same time, if the model converges on the same constraints for parts of the inventory that do not change form one stage to the next, this reinforces the claims it makes further. Indeed, we have seen that convergence is the norm without exception.

4.4.5 Progression of stages

Before considering individual stages and their content, it is important to consider how stages might differ from each other. There are three general ways in which a stage may develop into the next stage: first, a new feature may be acquired, secondly, a constraint is revoked, and third, no change occurs in either the number of features or constraints; in this case, a segment that was previously predicted but not attested is now in the inventory. We will come back to this phenomenon (which we will call ‘overprediction’) in detail. In each case, the inventory grows; this is trivial, as stages are defined as containing at least one additional segment in the inventory. The three different stage progressions are given schematically below.

(50) Stage progressions

	Feature	*Constraint*	
A	New	i.	New
		ii.	No new
B	No new		Lifted/Revoked
C	No new		No change

Situation A is a straightforward way to expand the inventory: to add a new feature to the system. Such a feature may be immediately subject to co-occurrence restrictions (A-i), or it may occur freely (A-ii). The latter situation is rare, and usually found in earlier stages, when there are few features to be combined. In practice, we only encounter situations in which a feature is free to combine with others in the first stages. Consider for example Eva, whose stage 1 has the inventory in (51a):

(51) Unconstrained feature combinations in Eva's first stage

a. Inventory

Stage	*Inventory*			
1	p	b	t	d

b. Features

Stage	*Features*
1	[labial], [voice]

This inventory is the perfectly symmetrical result of the features in (51b) (remember that there is no feature 'coronal'), and hence, no FCCs are introduced.

Situation Ai is default, and exemplified by Eva's Actual productions when she transitions from stage 1 to stage 2, as can be seen in example (52a).

(52) Eva's early stages

a. Stages

Stage	*Inventory*				
1	p	b	t	d	
2	p	b	t	d	n

b. Constraints

Stage	1	2
Constraints		*[lab, nas] *[voice, nas]

The feature [nasal] becomes active, but it is subject to restrictions: it may not co-occur with [labial], nor with [voice] (the former requirement remaining active in the final state of the grammar). This is enforced by the constraints in (52b).

Examples of situation B, where the inventory grows by revoking/demoting a constraint, are found in the data from virtually all children. Let us consider Noortje's inventory, and zoom in on the progression from stage 2 to stage 3. Example (53a) lists Noortje's inventory at stages 2 and 3. The inventory grows by adding /n/. In (53b), we see that no new features are acquired; the set in stage 2 suffices to generate all segments in the inventory at stage 3.

(53) Lifted constraints in Noortje's inventory

a. Inventory

Stage	*Inventory*				
1	m				
2	m	p	t	k	
3	m	p	t	k	n

b. Features

Stage	*Features*
1	[nasal], [labial]
2	[nasal], [labial], [dorsal]
3	[nasal], [labial], [dorsal]

c. Constraints

Stage	1	2	3
Constraints	[lab]→[nas]	[lab]→[nas]	
	[nas]→[lab]		
		*[lab, dors]	*[lab, dors]
		*[nas, dors]	*[nas, dors]

The inventory at stage 3 is generated simply by abandoning the requirement that nasals should be labial.

We see that a subset of the inventory in stage 2, leaving /k/ (and thus [dorsal]) out of the equation, also exemplifies situation B: the nasal inventory grows not by adding a new feature, but by lifting the constraint that restricts [nasal] to only co-occur with [labial].

The final manner in which stages may progress is listed in 50C, where the inventory grows despite the fact that no new features are added, and no constraints are revoked. It may seem contradictory at first that the inventory can change without a change in the system generating it, i.e., the inventory changes whilst there is no change in the grammar. The context in which this occurs is one where, at a given stage, the combination of features and FCCs predicts a larger inventory than is attested. In a next stage, attested and predicted are aligned once more. Remember that there are no overpredicted segments in the system of features and FCCs generating the adult inventory (see chapter 3). We will discuss the properties and implication of overpredictions in section 4.5, but before doing this, let us clarify what overprediction is by an example.

For this, we will return to Eva's developing inventory. Example (54) gives every stage of Eva's inventory (54a), the features she employs (54b), and the segments that the system predicts, but which are not attested (54c). Finally, the constraints that govern her inventory are listed in table (4.4.5).

(54) a. Development of Eva's inventory

Stage	*Inventory*												
1	p	b	t	d									
2	p	b	t	d	n								
3	p	b	t	d	n	m	ʋ	j					
4	p	b	t	d	n	m	ʋ	j	l	s	z		
5	p	b	t	d	n	m	ʋ	j	l	s	z	f	
6	p	b	t	d	n	m	ʋ	j	l	s	z	f	k

b. Features

Stage	*Features*
1	[lab], [voice]
2	[lab], [voice], [nas]
3	[lab], [voice], [nas], [cont], [apprx]
4	[lab], [voice], [nas], [cont], [apprx], [liq]
5	[lab], [voice], [nas], [cont], [apprx], [liq]
6	[lab], [voice], [nas], [cont], [apprx], [liq], [dors]

c. Overpredicted segments

Stage	*Overpredicted*
1	
2	
3	
4	[lab, voice, cont] [lab, cont]
5	[lab, voice, cont]
6	[lab, voice, cont]

Two segments are overpredicted in Eva's onset inventory. First, in stage 4, the segment [lab, cont] (/f/) is legal according to the combination of features and constraints in her grammar, but it is not yet attested. Secondly, its voiced counterpart, /v/, remains unattested in her recordings. This situation arises because of the early acquisition of /ʋ/, consisting of [labial, continuant, approximant], at stage 3. A subset of this segment, namely [labial, continuant], makes up /f/, which is unattested at this time. This is accounted for by the constraint [continuant]→[approximant]. This constraint must be revoked at stage 4, however, because /s/ is now attested; clearly, [continuant] is no longer restricted to [approximant]. The combination [labial, continuant] can now no longer be ruled out, and hence /f/ is predicted to be in the inventory. Since it is not, it is overpredicted.

The other overpredicted segment, /v/, consists of [labial, continuant, voice]. This combination is predicted to occur because the subset combinations cannot be ruled out: we have already seen how [labial, continuant] is overpredicted. The combination [continuant, voice] is legal and attested, in /z/. Furthermore, [continuant, labial] is present in /ʋ/, and [labial, voice] in /b/. Indeed, we see that /f/ is attested in the next stage (although /v/ remains unattested).

Stage	1	2	3	4	5	6
Constraints			[apprx]→[cont]	[apprx]→[cont]	[apprx]→[cont]	[apprx]→[cont]
			[cont]→[apprx]			
						*[apprx, dors]
				*[apprx, liq]	*[apprx, liq]	*[apprx, liq]
						*[cont, dors]
				*[cont, liq]	*[cont, liq]	*[cont, liq]
		*[lab, nas]				
						*[lab, dors]
				*[lab, liq]	*[lab, liq]	*[lab, liq]
						*[liq, dors]
			*[nas, apprx]	*[nas, apprx]	*[nas, apprx]	*[nas, apprx]
			*[nas, cont]	*[nas, cont]	*[nas, cont]	*[nas, cont]
						*[nas, dors]
				*[nas, liq]	*[nas, liq]	*[nas, liq]
			*[voice, apprx]	*[voice, apprx]	*[voice, apprx]	*[voice, apprx]
			*[voice, cont]			
						*[voice, dors]
				*[voice, liq]	*[voice, liq]	*[voice, liq]
		*[voice, nas]	*[voice, nas]	*[voice, nas]	*[voice, nas]	*[voice, nas]

Table 4.8: Feature Co-occurrence Constraints in Eva's word onset productions

Segment	Catootje	Eva	Jarmo	Noortje	Robin	Tirza	Tom
	10	6	9	9	10	5	9
t			1	1			1
s	3						5
f	6-8	4		7			
z					5		
v	6	4-6					

Table 4.9: Overpredicted segments. Numbers indicate stages at which overprediction occurs.

4.5 Overpredictions

The previous section concluded with the observation that overpredictions occur, where the term denotes segments that are legal with respect to the state of the grammar at some stage, but not yet attested. This is indeed a situation that we encounter at some frequency, and in this section, we will investigate why this is so, when it happens, and what it means. We will conclude the section with a brief discussion of why the reverse situation, where the attested inventory is larger than can be generated, does not occur.

4.5.1 Overpredicted segments

As it turns out, the variety of overpredicted segments is smaller than one might expect. Over all children, only five different segments are ever overpredicted. These are listed in (55) below:

(55) Overpredicted segments

Segment	*Feature combination*
t	[ ]
s	[continuant]
f	[labial, continuant]
z	[voice, continuant]
v	[labial, voice, continuant]

Apart from the fact that only a limited subset of all possible segments is ever overpredicted, all six have in common that they are relatively simple. There is only one segment that requires three features, the others are less complex. This, as we will discuss in more detail below, is because one of the important contextual aspects of overprediction is that it is not possible, under certain circumstances, to rule out feature combinations that form a subset of other combinations.

Table 4.9 lists the five overpredicted segments, and indicate per child and per level of description in which stage(s) it is overpredicted. A number of observations can be made with respect to these tables, and we will discuss the time span between the stage at which the overpredicted segments become available and the stage at which they are attested, and the identity of the overpredicted segments.

4.5.2 Possible causes

Overpredictions are taken to mean that segments that the child *should* be able to produce, are not encountered. This raises the question of what underlies overpredictions. Several possible causes exist:
1 – It could simply be the case that the system of Feature Co-occurrence Constraints as it is proposed here is too permissive. However, if this were the case, it would be difficult to account for the limited number of overpredicted feature combinations we find.
2 – Another possibility is that the child does in fact produce the overpredicted segments, but just happens not to do so during the recording session(s). Considering the generally unmarked status of the segments in (55), this seems an unlikely explanation. For example, van Severen, Molemans, van den Berg, and Gillis (2012) find that while chances of inclusion (avoiding false negatives) are related to sample size, this is true to a lesser degree for more frequent, less marked segments.
3 – A very similar explanation is that the segments are in the inventory, but do not reach criterium yet. The grammar of the child is constantly evolving, and the recordings take place at what are, essentially, random moments. Because of this, a 100% match between predicted and attested inventories is not to be expected in the first place. However, this explanation applies only to those overpredictions that are reasonably quickly resolved.
4 – Considering the limited variation in the table in (55), a possible cause could be that the overpredictions that are encountered are an artefact of the feature system that is used.

The latter option is certainly true of /t/, which cannot be ruled out, being devoid of featural content and thus immune to any FCCs. In fact, in every case in the current survey, /t/ is present at the first stage. In many cases it is in the attested inventory, but on those occasions where it is not, it is in the list of overpredicted segments: it cannot be ruled out and thus is predicted to always be in the inventory. In other words, the prediction is that it is acquired first. This is not always the case, but a look at table 4.9 reveals that where /t/ is overpredicted, this situation is usually resolved quickly.

If we go back to the raw data, before the filters listed in 4.2 are applied, we see clear evidence that the penultimate explanation (overpredicted segments are in the inventory but do not yet reach criterium) is also true. Of all the cases listed in table 4.9 above, only a handful turn out not to be in the inventory at all. These are listed below:

(56) Unattested overpredicted segments
/t/: Noortje, stage 1

/f/: Catootje, stages 6-8
/v/: Catootje, stage 6

This brings the list of overpredicted segments down to three: /t, f, v/. Of these, we know that the first cannot be ruled out, and tellingly, if it is unattested, it is only so in the first stage of Noortje's acquisition.

The other two segments, /f, v/ make up a considerable part of the fricative subset inventory of Dutch. There are two possible explanations why it should be these segments that we encounter here. First, it is a familiar observation that children prefer to avoid fricatives in onsets (Fikkert (1994), see also section 4.6.3 below on initial stopping). Secondly, each segment forms a subset of an approximant: /f, v/⊂/ʋ/. Indeed, in Catootje's stages 6-8, /f/ is not in the inventory, and in stage 6, /v/ is not in the inventory, whereas in these stages, /ʋ/ is.

Criteria and data inclusion

At this point, it is of interest to note that a clear prediction of the Feature Co-occurrence Constraints theory is that /t/ is present in the developing inventory from the start. Noortje's overprediction of /t/ in her stage 1 illustrates an important point. Stage 1 contains only /m/, but only because it minimally reaches criterion. Had it not, or had the criterion been slightly different, stage 1 would have extended over more recording sessions, and it would have included /t/.

Also of interest is that the majority of overpredictions in table 4.9 is only apparent in the sense that the segments are produced, but not in such a way as to reach criterium.

These observations illustrate an important proviso that must be acknowledged with respect to any study of acquisition: inclusion criteria are always somewhat arbitrary, and never without artifactual consequences. The criteria in 4.2 are no exception. They were chosen to resemble those in Levelt and van Oostendorp (2007), and also the criteria proposed in Ingram (1981).

It is, of course, possible to admit every instance of a segment into the data set. In fact, some studies do just this (Ferguson & Farwell, 1975). However, as pointed out by Ingram (1989), this makes any results extremely vulnerable to incidental variation, or even to non-linguistic utterances. In fact, this is illustrated by Noortje. Remember that her stage 1 is defined by containing only /m/. Before stage 1, she produces one word consistently (/mama/ [mama]) and some only once (/ku/ [ku] 'cow'). During the third recording, a new word enters her vocabulary: /χəmakt/ [mɑ] 'made'. Taken together with four instances of 'mama' (counted as a single instance for being tokens of one type), makes it that Noortje just reaches criterion for /m/. The problem is that it is difficult to ascertain whether /mama/ is a 'word' in the sense that it is generated by a phonological grammar, or whether it is a remnant from a previous stage of development.

If including every utterance is too permissive, Noortje's case illustrates that perhaps, our criteria still not restrictive enough. A case can be made to exclude items such as 'mama' and 'papa', onomatopoeia, and some other classes. In the current study, we have opted to restrict ourselves to objective, numerical inclusion criteria. These could have been stricter or laxer, but ultimately a choice must be made.

4.5.3 Context for overpredictions

In the previous section, we discussed possible causes of overprediction. We concluded by observing that many of the overpredicted segments are subset segments of others. Simply being an unattested subset segment of an attested segment is not enough to be overpredicted. This is because i-constraints can force the feature(s) comprising the subset segment to co-occur with another feature. Below, we will see that there are three formal contexts in which overpredictions occur:

(57) Segment (feature combination) A is overpredicted if
  a. A is unattested **and**
  b. A is empty

(58) Segment (feature combination) A is overpredicted if
  a. A is unattested **and**
  b. there is some attested feature combination B such that $A \subset B$ **and**
  c. there is an attested feature combination C such that at least one of the members of $A \in C$ **and**
  d. $C \neq B$

(59) Segment (feature combination) A is overpredicted if
  a. A is unattested **and**
  b. $|A| > 2$ **and**
  c. every subset $\{F,G\}$ such that $\{F,G\} \subset A$ is in some segment B, C, *etc.* **and**
  d. $B \neq A$ and $C \neq A$ *etc.*

The contexts in (57a) and (58a) are obvious: a segment can only be overpredicted if it is unattested. We have already discussed requirement (57): /t/ is either attested or overpredicted. The requirements in (58) treat cases of non-empty segments. Requirement (58b) demands a superset of the overpredicted segment be present, while requirement (58c) ensures that at least one of the features in A occurs in another combination than the one in B. This is to ensure that there is no i-constraint limiting the subset feature(s) to a single co-occurrence, as the following will illustrate. Take, for example, unattested segment A to be [F,G]. It is a subset of segment B, which is [F,G,H]. In this scenario, which corresponds to (58 a-b), it is still possible to describe the ungrammaticality of segment A by use of the constraint G→H. This possibility no longer exists, however, if segment C ([G,I]) exists, as the i-constraint is now violated by C and hence cannot be employed to rule out A.

Most cases of overprediction are covered under the definitions in (57) and (58), but there are exceptions. These are described in requirement 59. To illustrate, let us look at the overprediction of /v/ in Catootje's stage 6. Below, the inventory is given at the relevant stage and the ones preceding and following it.

(60) Overprediction in Catootje's stages

Stage	*Inventory*													
5	p	b	t	k	m	n	l	d	j	z	s	r		
6	p	b	t	k	m	n	l	d	j	z	s	r	ʋ	
7	p	b	t	k	m	n	l	d	j	z	s	r	ʋ	v

At stage 6, both /f/ and /v/ are predicted to be in the inventory, but are not actually attested. The situation is resolved at the next stage for /v/, but not for /f/ (incidentally, both segments are also not present in the unfiltered inventory – see above). Running /f/ through the requirements listed in the definition in (58), we see the following:

(61) Overprediction of /f/: [cont, lab] is

a. not attested and

b. there is some attested feature combination /ʋ/ [cont, lab, apprx] such that /f/⊂/ʋ/ and

c. there is an attested feature combination /m/ [lab, nas] such that [lab]∈[cont, lab] and

d. /m/ ≠ /ʋ/

In other words, the overprediction of /f/ falls neatly in the context described in (58). The same does not hold for /v/, however.

(62) Overprediction of /v/: [voice, cont, lab] is

a. not attested and

b. there is **no** attested feature combination X [voice, cont, lab, . . .] such that /v/⊂X and

c. there is an attested feature combination /m/ [lab, nas] such that [lab]∈[voice, cont, lab] and

d. /m/ ≠ X

There is no superset segment for /v/, and yet it is overpredicted. Looking at the inventory in (60), we see that every possible subset of /v/ is attested: [voice, continuant] in /z/, [continuant, labial (approximant)] in /ʋ/, and [labial, voice] in /b/. Thus, there cannot be a ban on the combination of [voice] and [continuant], there cannot be a ban on the combination of [continuant] and [labial], and there cannot be a ban on the combination of [labial] and [voice]. Since each of the three constituent features co-occurs with at least two others, i-constraints cannot be of help here, either. Hence, we need additional requirements to the ones in (57) and (58).

(63) Segment (feature combination) A is overpredicted if

a. A is unattested **and**

b. $|A| > 2$ **and**

c. every subset $\{F,G\}$ such that $\{F,G\} \subset A$ is in some segment B, C, *etc.* **and**

d. $B \neq A$ and $C \neq A$ *etc.*

With these three definitions, we can describe every case of overprediction in table 4.9, where definition (57) describes the overprediction of the zero-feature segment /t/, while definition (58) describes the overprediction of the one- and two-feature segments /s, f, z/, and definition (63) describes the overprediction of the three-feature segment /v/.

4.5.4 Underpredictions

Absent from the findings are underpredictions: feature combinations that are attested, yet not predicted by the set of features and constraints. FCCs allow *at least* the feature combinations that are actually attested. The absence of underpredicted segments follows directly from the procedure by which FCCs are constructed.

A segment can be underpredicted by virtue of either a c-constraint or an i-constraint. In the first case, a segment [F, G] is attested and yet a constraint *[FG] is activated. C-constraints, however, are generated based on a two-dimensional feature co-occurrence matrix where all possible combinations of two features are indicated as attested or unattested (see section 4.2 above). For every unattested feature combination in this matrix, a c-constraint is activated, but not for every attested combination of two features. For underprediction to occur, attested feature combinations must be ruled out. It is clear that the procedure used to derive c-constraints cannot do this.

The other possibility for underprediction is by an i-constraint. In this case, the constraint $F \rightarrow G$ is activated while F is attested to co-occur with at least one other feature H, where $G \neq H$.

The procedure for deriving i-constraints starts out from attested combinations of two features. Hence, it would seem that there is a higher risk for underprediction (as underprediction entails that an attested combination is ruled out). Every possible i-constraint that refers to an attested combination of two features is candidate for activation. Next, for each of these pairs, every other attested combination is checked. If the antecedent occurs in combination with any feature H where H≠G, the i-constraint F→G is no longer a candidate for activation. Hence, the procedure used to arrive at the set of activated i-constraints is unable to yield underprediction – just as the algorithm cannot derive underprediction by c-constraints.

4.5.5 Summary

Although the theory of Feature Co-occurrence Constraints largely yields correct results, we did encounter some examples of overpredicted segments – segments that are allowed by the set of features and constraints at some stage, but not attested. The set of overpredicted segments is very small, compared to the set of possible feature combinations. We discussed several possible causes for overpredictions, where the most important ones were a) a strong prediction that /t/ should be acquired at the first stages, and b) overpredicted segments are often subsets of attested segments (this is, of course, trivially true of /t/, as it is the interpretation of an empty segment). This brought us to a set of formal definitions of the contexts in which overprediction occurs. Finally, the reason that underpredictions do not occur follows from the constraint derivation algorithm. Before concluding the current chapter, let us consider the developmental origin of Feature Co-occurrence Constraints.

4.6 The origin of Feature Cooccurrence Constraints

The theory proposed in this thesis has two main ingredients: features and constraints. In section 2.3, we have investigated the question of the innateness of distinctive features from the perspective of categorisation, lexical storage and phonological rules and generalisations, and in each case, we have seen that the evidence supports the idea of innate features. In this section, we will discuss the innateness of constraints.

We must differentiate between different notions of innateness. One possibility is to claim that humans come into this world with a predetermined, universal and substantive set of constraints (CON). This is the position we find in the 'classical' OT literature (Prince & Smolensky, 1993/2004). On the other hand, it is possible that constraints are formed during acquisition. Given this position, two scenarios are possible: first, the child may posit whatever constraint she finds evidence for, and second, instead of an innate set of substantive constraint, the child comes into the world endowed with a limited set of innate constraint *templates*, that are actuated and/or populated based on experience with the surrounding language.

In this section, we will begin with an overview of the attitudes toward the developmental origin of constraints in the literature, focussing mainly on the most well-known constraint-based framework in phonology, Optimality Theory. Surprisingly little experimental research has been done with regards to this question, with Jusczyk et al. (2002) as a notable exception. We will consider their results and methodology in some detail. Finally, we will see that there are reasons to distinguish between constraints concerning melody, and those concerning prosody. The arguments for substantive innateness are much stronger in the case of the latter.

4.6.1 Ideas and Assumption in the Literature[10]

In the original conception of Optimality Theory, constraints were proposed to be both universal and innate. The *only* aspect of grammar that was thought to be language specific was the constraint ranking. In this respect, OT signaled a departure from generative phonology in at least two ways. First, as we shall see, constraints in earlier generative phonology (morpheme structure constraints, surface phonetic constraints) were language specific and learned. Secondly, constraints were originally invoked to *limit* the generative potential of the grammar. In OT, constraints *define* the generative potential.

Prince and Smolensky (1993/2004) are very clear in stating that the goal of the Optimality Theoretic program to provide a set of universal constraints. This ambition is made possible by the architecture of the theory; in previous thought, a conflict existed between the ambition to make universal statements and the requirement of these statements to be 'surface- or level true'. In Optimality Theory, the latter requirement no longer plays a role, because the theory provides a constraint-external mechanism of conflict resolution: strict domination. In Optimality Theory, constraints can be universal because they are *always* true: if a constraint X does not correspond to some output structure in language A, this does not mean that X is falsified by facts about A, but rather that in A some constraint Y takes precedence. The upshot is that if the theory is going to make predictions, it *must* assume substantively universal constraints; if language-specific constraints are admitted, overgeneration and post-hoc argumentation take over. Prince and Smolensky (1993/2004) do not specifically mention innateness of constraints but seem to imply that they are; the acquisition of grammar in OT is equal to ranking constraints, and a learnability problem arises if constraints must be learned, and then ranked – especially it they are to be ranked low (low ranking entails little influence and hence little overt evidence). Innateness of constraints (or constraint templates) is also the premise in later work on learning algorithms in OT (Boersma & Hayes, 2001; Tesar & Smolensky, 2000).

In addition, a subset principle violation poses a problem for emergent constraints. A well-known position in the learnability literature is that children learn only from positive evidence – that they make no generalisations based on the *absence* of structure. Hence, the most conservative grammar must corre-

---

[10] I would like to express my gratitude to Edoardo Cavirani, for parts of the following section originated in discussions with him and in attending his presentation (Cavirani, 2012)

spond to the initial state. The development of the grammar is informed only by positive evidence. This idea, roughly sketched here, is known as the 'subset principle' (first formalised in phonology by Dell (1981)). It is easy to see that the subset principle is difficult to combine with emergent constraints (that is, non-innate constraints that are emergent in the child's developing grammar). Constraints are, generally speaking, negative statements, banning structures that are not allowed, like NoCoda, which disallows codas. If the language surrounding the learner has only CV syllables (at the surface level), there is no way for the constraint to emerge in the child's grammar – even though the evidence favouring it is abundant. A grammar without NoCoda is perfectly capable of generating CV syllables, and the lack of coda consonants in the child's input or uptake could just as well be an accident. Conversely, consider the same learner but now surrounded by a language in which CVC syllables are quite frequent. Here, too, there is no reason for the child to posit a constraint NoCoda, because codas are allowed. See, for a similar argument, Hale and Reiss (2003), who explain the issue in clear terms. The problem is that lower-ranked constraints in OT can still be active sometimes, a process known as The Emergence of the Unmarked (TETU): unmarked structures sometimes arise in defiance of higher-ranked faithfulness constraints, if there is nothing to be faithful to (e.g., in reduplication, loan-word adaptation, etc.).[11]

For emergent constraints to work, then, there must be some external way to motivate their presence in the grammar. One way to do this is with a 'phonetic difficulty map', as has been proposed by Hayes (1999). Here, the learner maps the articulatory effort in phonetic space, and assigns a numerical value to each phoneme, corresponding with its 'difficulty'. Next, potential feature co-occurrence constraints[12] are profusely hypothesised by the learner, and then evaluated with respect to the segments on the map. So, for a constraint C and any two segments X and Y, four possibilities exist:

(64) a. Both X and Y violate C
b. Both X and Y satisfy C
c. X violates C, Y satisfies C
d. X satisfies C, Y violates C

The first two cases are uninteresting, as the outcome is the same for both segments. The important comparison is in the latter two situations. Roughly

---

[11] A way out in Optimality Theory could be to resort to the principle of Richness of the Base: since any conceivable output form is a candidate for any given input, GEN will propose forms that have a coda. The fact that these never surface could be interpreted as a form of positive evidence. Note that for this argument to hold the candidates with coda must not be harmonically bound by the forms without in any other way, as this would destroy the evidence for NoCoda. Another issue is that such a proposal to work still needs to refer to innateness, if not in CON, but in GEN: the learner must still consider the candidate with coda as a candidate, and hence, in the face of no surface evidence, know that there is such a thing as a 'coda' – as per the *innateness of primitives principle* (see section 2.3.

[12] Hayes (1999) proposes both paradigmatic and syntagmatic feature co-occurrence constraints.

speaking, if X violates C and Y satisfies C, C is a ‘good’ constraint if X is more difficult then Y. Thus, in this example, if the situation in (64c) holds, *and* Y is easier then X, potential constraint C makes a correct prediction. However, if X is easier then Y, the prediction is incorrect. For any pair of two segments, predictions are made and subsequently, the ‘goodness’ of constraints is determined by the following metric:

(65)

$$Effectiveness = \frac{Correct\ predictions}{Correct\ predictions + Incorrect\ predictions}$$

This yields an index of ‘groundedness’ for all constraints. All potential constraints are mapped in a constraint space, whereby constraints whose structural description differs minimally are neighbours. In this multi-dimensional space, local maxima of constraint effectiveness are determined, and the corresponding potential constraints are those that will constitute the grammar. In this way, Hayes (1999) sketches how constraints can be emergent. The necessary ingredients (or presupposed knowledge) for his learner are a) knowledge of phonological features, b) their phonetic content and c) the (relative) difficulty involved in expressing these features in different contexts. Hayes’ proposal thus presents us with a less severe version of the duplication problem: even though articulatory knowledge is not directly encoded in the grammar, it *must* be directly encoded in the algorithm that gives the grammar substance. The subset problem is circumvented, however, because the evidence for constraints is not the absence of certain pattern (i.e., negative evidence), but stems from the knowledge the learner has about relative phonetic difficulty.[13]

The constraints proposed by Fikkert and Levelt (2008), discussed in section 2.3.2, are not proposed to be innate either, but in contrast to Hayes’ inductive grounding, they rather emerge as generalisations over the lexicon. This is illustrated with the production patterns of Noortje, who has initial dorsals initially, then substitutes them for coronals, and only later reverts back to faithful productions of dorsals in C_1. These generalisations, being generalisations over the lexicon of words that the child has stored after being exposed to them, reflect input frequency. Thus, the authors are fairly explicit in the way in which they allow input frequency to play a role in acquisition. As is often the case when invoking input frequency, however, the argument runs the risk of being circular: if learners posit grammatical devices (constraints) to formalise generalisations over the input, it could be expected that these devices will leave a trace in the adult language. In fact, first language acquisition is often hypothesised to be the mechanism of language change!

---

[13] It remains a question whether the proposal in Hayes (1999) is in fact more parsimonious then a theory in which all constraints are innate and ineffective constraints are quickly demoted, as every possible constraint must be generated and evaluated for its effectiveness. The difference is that for Hayes, only effective constraints are admitted to CON.

A more serious problem for emergent constraints of the type proposed by Fikkert and Levelt (2008), as discussed above, is that they are not learned from positive evidence: it is the absence of non-coronals in initial position that leads the child to posit [LABIAL, and later the absence of dorsals in that position to posit *[DORSAL. Superficially, it would appear that the same problem haunts the analysis proposed in the current thesis, but the learning datum for Feature Co-occurrence Constraints, is a new segment; if for that segment a new feature must be activated or acquired, it is accompanied by FCCs to keep the grammar as conservative as possible. In contrast, the learning datum for Fikkert and Levelt (2008) is introspection and generalisation over the lexicon.

4.6.2 Experimental evidence

One of the few studies that explicitly tackles the question of constraint innateness from an experimental point of view is Jusczyk et al. (2002). Here, the cross-linguistically highly frequent process of nasal cluster place assimilation was the subject of investigation. Children were exposed to triads of non-words of the form X... Y... XY, where XY is a concatenation of X and Y. X has a nasal coda, and Y has a obstruent onset. In this way, the effects of markedness (cluster assimilation) and faithfulness (no assimilation) could be tested. An example of a non-assimilating triad would be *on... pa... onpa*, an assimilating triad would be *on... pa... ompa*.

In the first experiment, 10 month olds were tested to see if they have a preference for either faithful or unmarked outputs. In this version, all stimuli were non-assimilating, but some concatenations were marked, and some were unmarked. An example of a marked triad is *un... ber... unber*, whereas its unmarked counterpart is *um... ber... umber*. Hence, although all triads represented faithful concatenations, half violated the markedness constraint agains non-homorganic nasal-obstruent sequences. The children showed a preference for the unmarked triads, indicating that they display evidence of a markedness constraint.

Experiment 2 sought to investigate whether 10 month old infants have faithfulness constraints, as well. To achieve this, the stimuli were manipulated so that half contained faithful concatenations (*um... ber... umber*), whereas the other half consisted of triads in which the latter member was phonologically unrelated to the former two: *um... ber... iŋgu*. An added benefit of this set-up was that it checked whether the results in experiment 1 were not due to a general novelty preference, and indeed, the children preferred the faithful sequence. Finally, in experiment 3, faithfulness and markedness were pitted against each other. One type of stimuli violated markedness but respected faithfulness (*un... ber... u**nb**er*), whereas the other violated faithfulness but respected markedness (*un... ber... u**mb**er*). As expected, the 10 month olds showed a preference for the alternating, unmarked sequences. In order to rule out the possibility that the results were due to native language knowledge (10 month olds are sensitive to phonotactic properties of the environment language

(Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993)), the same set of experiments was repeated with $4\frac{1}{2}$ month olds. The results were identical, indicating that native language experience may not have been the decisive factor.

According to Jusczyk et al. (2002), these results are consistent with a nativist OT account, in which constraints are substantively innate, and markedness constraints outrank faithfulness constraints in the initial state. With respect to the framework that we are developing, this implies that Feature Cooccurrence Constraints are innate too, and what is more, that they are all initially high-ranked.

However, the interpretation of Jusczyk et al. (2002) crucially relies on the assumption that the infants actually parse the stimuli as X... Y... XY, rather then X... Y... Z. In other words, it depends on whether children take the latter member of the triad to be derived from a concatenation of the earlier two remains uncertain. Hence, the results may not be as conclusive as they are presented to be.

4.6.3 Two types of constraints

In their discussion about emergent constraints, Fikkert and Levelt (2008) propose that some constraints might be innate, whereas others are not. The division, they propose, might coincide with the domains of (sub) segmental phonology on the one hand, and suprasegmental phonology on the other. That is, constraints concerning features are emergent (possibly as generalisations over the developing lexicon), whereas constraints governing prosodic structure might be innate. One possible argument for this is that cross-linguistically, a considerable variation exists in (the phonetic expression of) phonemic categories, while the typology of prosodic structures appears to be more restricted (to such a degree that some deny its existence *per se* (Scheer, 2004, for example) – but this goes beyond our present concerns).

This idea is also found in Inkelas and Rose (2008). Inkelas and Rose describe the phonological development of E., a child learning English. The two phenomena described are Positional Lateral Gliding and Positional Velar Fronting, two cases of what the authors call ‘strong merger’: neutralisation of a contrast in a prosodically strong position. This is quite unusual from the perspective of adult phonologies: both Positional Faithfulness (J. Beckman, 1998) and Positional Markedness (Lombardi, 1995; Zoll, 1998) are designed to capture the generalisation that positional effects in adult languages are Weak Merger (positional neutralisation in weak positions) and Strong Enhancement (making strong positions more salient). Weak Merger is taken to be the result of Positional Faithfulness: certain faithfulness conditions may only apply to strong positions, whereas Strong Enhancement is taken to be affected by Positional Markedness: certain contrasts may only appear in strong positions. Also, positional neutralisation rarely targets primary PoA. In short, if neutralisation is positionally limited, it is limited to weak positions. Hence, Strong Merger poses a challenge to the Continuity Hypothesis.

In Positional Velar Fronting (PLG), underlying velars are produced as coronals in the onset of words and stressed syllables. In Positional Lateral Gliding (PLG), all laterals are glided. Laterals in onsets of words and stressed syllables become palatal [j], laterals in weak positions become [w]. Some examples are given in (66) (representing a subset of the examples given in Inkelas and Rose (2008, ex. 1):

(66) Strong Merger in E.’s speech

a. Positional Velar Fronting:

[t^{h}ʌ]	‘cup’	1;09.23
[ˈtuwɔ]	‘cool	’ 1;11.02
[əˈdɪn]	‘again’	1;10.25

b. i. Positional Lateral Gliding: strong positions:

[jæmp]	‘lamp’	1;10.0
[jɪdiə]	‘Lydia’ (adult: [ˈlɪdiə])	2;1.8
[jivan	‘Livan’ (adult: [liˈvan])	2;8.19

ii. Positional Lateral Gliding: weak positions:

[hæwət^{h}ʌəkɚ]	‘helicopter’	1;11.10
[æwədɛɾɚ]	‘alligator’	2;1.18
[hɨwdə]	‘Hilda’	1;11.10
[bejgu]	‘bagel’	1;9.24

c. PVF and PLG interacting:

[ʧiːn]	‘clean’	1;11.0
[ˈʤæsəs]	‘glasses’	2;2.1

Positional Velar Fronting is a systematic part of E.’s speech for a fourteen-month period beginning early in his second year, whereas Positional Lateral Gliding starts eight months later. Hence, during some period, the two processes interact, providing further evidence for the hypothesis that the processes are productive parts of E.’s phonological grammar (see (66c)). What we see here is that in target /Kl/ onset clusters, PVF ensures that the dorsal is fronted, whereas PLG glides the lateral. The resulting /Tj/ cluster is then merged to a single affricate. Interaction and merger (palatalisation) does not occur in all cases; in roughly one third of the eligible clusters, the lateral is simply omitted.

The apparent contradiction to the Continuity Hypothesis posed by Strong Merger is solved by Inkelas and Rose (2008) by appealing to the notion of ‘phonologisation’. Phonologisation refers to two related phenomena: first, it can mean the diachronic adoption into the phonological grammar of a phenomenon that was already phonetically present; second, it can mean the adoption (into the phonological grammar of individual children) a phenomenon that was already phonetically present, applying rather in the time scale of individual language acquisition. In the case of E., what this means concretely is that he is aware of the fact that strong positions require enhancement of articulatory gestures (in other words, he is aware of ‘syntagmatic contrast’). Laterals contain the orchestrated movement of two articulators: tongue tip and tongue body. The relative prominence of these two movements is how clear (tongue tip) /l/ and dark (tongue body) /l/ are distinguished. In gliding his laterals, E. retains this double identity of laterals, resulting in a coronal (tongue tip) glide in the positions where clear /l/ is usually found, and a labio-velar (tongue body) glide in positions where we usually encounter dark /l/. These positions, of course, coincide with strong and weak positions respectively. We can see that E. has phonologised this pattern when we consider intervocalic positions: pretonic intervocalic consonants are glided to [j], whereas post-tonic laterals are glided to [w].

In PVF, E. also remains faithful to the articulatory gestures involved in producing velars. However, children's production of stops is less refined than adults' productions, and physiologically, their tongue body is much larger with respect to the oral cavity. These two things result in greater difficulty producing velars. In addition, in English, stops in strong position are articulated more extremely than stops in weak positions. Furthermore, velar stops in strong positions are somewhat more fronted. The combined result of these factors is that "... [i]n the context of imperfect articulatory control, bigger tongue size, when combined with a relatively shorter palate, implies that even a slight increase of vertical tongue movement, required in the enhanced articulations in prosodically strong positions, will have direct consequences for the childs production of target velars. The greater emphasis on the dorsal articulator expands tongue contact into the coronal region, yielding the coronal release that characterises fronted velars." (Inkelas & Rose, 2008, p. 724)

Similar cases have been reported. For example, Chiat (1989) reports on a child who stops fricatives word-initially and in the onset of strong syllables word-medially, but produces them faithfully elsewhere.[14] Similarly, in a report on two experimental studies on the consonantal substitution patterns in different positions by English learning children with delayed phonological development, Rvachew and Andrews (2002) note a general pattern that '... consonants that occur at the beginning of stressed, word-internal syllables will be produced in the same manner as consonants that occur at the beginning of words', that the syllable final position patterns alike regardless of position in the word, and, with respect to fricatives, for two of the three children the intervocalic (before an unstressed nucleus) position and the syllable final position patterned alike. Although Rvachew and Andrews (2002) note that substitution patterns are generally alike for syllable initial positions, regardless of the adjacency of a word boundary, and that the same holds for syllable final positions, they treat intervocalic consonants as a distinct category ('ambisyllabic'). In other words, the children's consonantal substitution patterns were influenced by foot structure.

In his study on truncation in child language, Pater (1997) underlines the importance of recognising the syllable as a unit in the developing phonology,

---

[14] The same child also exhibits positional velar fronting, see Chiat (1983).

next to the foot. The general truncation pattern for trisyllabic (and longer) words in English and Dutch (Fikkert, 1994; see Pater, 1997 for further references) is for the stressed and final syllable to survive. Together, they satisfy a minimal word requirement, where the minimal prosodic word is a trochaic foot. For example, Pater reports on examples where *cinnamon* truncates to [sɪmɛn], *Allison* to [æ:sʌn], and *museum* to [zi:ʌm]. However, there are forms that deviate from this pattern: *banana* can become both [nænə] and [bænə], *marakas* becomes [ma:kas], and *delicious* becomes [dɪʃəs]. The crucial point is that the onset of the stressed syllable is replaced by the onset of the syllable that is otherwise deleted, when the latter is of lower sonority. These cases, then, indicate that children are aware of the relation between sonority and syllable structure.

The crucial point in these studies is that the learners are very much aware of prosodic structure and are willing to risk sacrificing segmental contrast in order to express syntagmatic contrast (prosodic strength). Here, we see an early activity of prosodic well-formedness, at the expense of (sub)segmental faithfulness.

4.6.4 About the innateness of Feature Co-occurrence Constraints

The arguments reviewed so far are equivocal; on the one hand, there seems to be good reason to oppose the idea of innate Feature Co-occurrence Constraints – even if features themselves might be innate. A hint in this direction was made by Levelt and van Oostendorp (2007) and Fikkert and Levelt (2008). Substantial backing for the idea that segmental constraints , as opposed to prosodic constraints (Inkelas & Rose, 2008), need not be innate is provided by Hayes (1999). The two domains are very different, where the suprasegmental domain is more restricted then the sub-segmental domain. This intuition is also captured by Scheer (2004), who divides phonology in UPPER and LOWER. Relations in UPPER are much more restricted: government and licensing are both bound by directionality, for example. Although this difference in restrictedness is not in itself an argument about innateness or emergence, it does show that there is an important difference.

Emergent constraints do not fare well in light of the subset principle. Constraints are statements on what is *not* allowed, but by definition, there is no positive evidence for illegal structures. Note that this argument is independent of the argument about constraint ranking (or deactivation). As soon as there *are* constraints, positive evidence must be available to determine their place in the grammar. Both extreme positions about constraints, namely substantive innateness and absolute emergentism are not viable, then – at least for constraints governing sub-segmental phonology. Hence, the third option, of constraint templates, remains as the best option. Note that the issues posed by the subset principle do not categorically rule out emergent constraints, as long as some independent mechanism can be found to generate them. This is the route we will adopt here, because in light of the studies and arguments cited above, it would seem that there is no evidence for innate Feature Co-occurrence Constraints.

If the set of Feature Co-occurrence Constraints is not innate, then, we must somehow account for the limitations of their form. Only constraints referring to two features are allowed, and only two connectives occur. The motivation of these limitations were discussed in chapter 3, but for now, I propose that rather than an innate set of substantive constraints, the child is equipped with a much smaller set of innate constraint templates.

4.7 Summary

The theory of Feature Co-occurrence Constraints and the segment inventory as developed in chapter 3 is shown to account for the acquisition of the segment inventories of different children in the present chapter. Acquisition is modeled as the succession of stages, defined by an addition to the inventory. Although these stages are not independent from the perspective of the features that are acquired (i.e., features are acquired in a monotonically increasing manner), they are independently approached by the constraint derivation procedure. Even though the procedure has no memory of what constraints it generated at a previous stage, we saw that the assumption that the activation time of constraints is severely limited is borne out fully: all constraints are activated no later than the moment at which the second feature that they refer to is acquired. More importantly, the algorithm converges on those parts that do not change from one stage from the next: a gap described by constraint C remains a gap described by constraint C.

The demonstration of Noortje's development showed that the theory is largely correct in describing the inventory in terms of features and constraints on feature combinations. However, the theory overshoots at some points: a total of six different segments were found to be overpredicted in the data from the children included in the study. One of these, /t/, cannot be excluded, and hence, it must be present from the first stage; either as an attested segment, or as an overprediction. For the other segments, /p, f, s, v, z/, the contexts in which they can be overpredicted were defined. However, overpredictions were found to be largely an artefact of the inclusion criteria for segments to be accepted as 'acquired'; a look at the unfiltered initial data revealed that only a number of cases of overpredictions actually concern segments that are not produced during the recording sessions. Underpredictions are not attested, which was attributed to properties of the constraint derivation procedure.

Finally, we turned to the question of innateness of constraints, a matter which is separate from the question of innateness of features (see section 2.3). We found that there are good reasons to assume that constraint templates, but not substantive constraints, are innate.

All in all, the model was able to account for the attested inventories, for each child and for each stage. This means that an important prediction, namely that Continuity holds at the level of the structure underlying the inventory (constraints and features), is borne out.

4.7.1 Final thoughts

The question of how the acquisition of the segment inventory can be modeled in a satisfactory way has divided researchers for a long time. Approaches that are based on a strong continuity with the adult grammar and utilise distinctive features have great appeal. However, some feature-driven theories of phonological acquisition are too restrictive. Jakobson (1941/1968), for example, proposes a universal order of acquisition based on oppositions between features. Such a universal order has, however, never been uncovered. On the contrary, considerable variation in order of acquisition exists both between and within languages. At the same time, as pointed out by Levelt and van Oostendorp (2007), some theories are too permissive. One example is Beers (1995), who describes the order of acquisition in terms of the unfolding of a feature geometry, where mother nodes are acquired before daughter nodes, and variation is restricted to the relative order of acquisition of sister nodes. Importantly, however, the theory makes no attempt at restricting the combinatorics of features once they have been acquired: every feature is free to combine with others once it is part of the child's system. Gaps in the segment inventory remain unaccounted for, so at best, the theory is incomplete.

Non-feature-driven theories have been put forward, in part as answer to the difficulties described above. Some eschew the notion of distinctive features altogether (exemplar theory), others hold that phonological features are emergent properties of an ever more densely populated lexicon (lexical diffusion theories). Each make specific predictions with respect to the acquisition path.

Exemplar Theory (M. Beckman, Yoneyama, & Edwards, 2003; Zamuner, Gerken, & Hammond, 2005) holds that children acquire their language based on whole forms – words, segments – rather than abstract categories such as distinctive features (various exponents of Exemplar Theory differ in the degree to which they allow abstraction to take place later in development). An important aspect of the theory is that acquisition proceeds by general (acoustic, statistical) processing of the input; i.e., there is no language specific competence. This entails that input frequency is an important, if not the sole, predictor of the path of acquisition. For example, Gonzalez-Gomez, Poltrock, and Nazzi (2013) find that children acquiring French learn CVCV sequences at an earlier stage when C1 is labial and C2 is coronal, than when the PoA specifications are the reverse. This, they argue, is due to the higher frequency of labial-coronal than coronal-labial words in French.

Other studies, too, have found evidence in favour of input frequency effects on phonological acquisition, but many have failed to do so. Most important to our present subject, Levelt and van Oostendorp (2007) found no correlation between the order of acquisition of Dutch consonants and the relative frequency in a corpus of Dutch child-directed speech (van de Weijer, 1998).

Apart from a lack of empirical evidence, the frequency approach suffers from a number of other problems. First of all, it is entirely unclear what the precise relation is between input frequency and acquisition. Even if a correlation is found, cause and effect are rarely scrutinised. Does a higher frequency cause earlier acquisition, or is an independent learning bias the cause of both early acquisition and higher frequency in the adult language (see, for a rare comment on this issue, Fikkert and Levelt (2008))? What is the relevant measure of frequency? What constitutes a relevant corpus? Furthermore, the frequency approach appears ill equipped to deal with individual variation, such as the variation in acquisition order encountered in the current study. The tacit assumption is that large corpora provide an accurate representation of the input frequency for a given language community. However, in order to account for individual paths of acquisition, individual input corpora should be used (see again Fikkert and Levelt (2008)).

In theories of lexical diffusion, again the unit of acquisition is the word, rather then the feature. Lexical items are subject to finer degrees of specification (word → segment → feature), the more the lexicon becomes populated with ‘neighbours’: words that are highly similar (see, for example, M. Beckman and Edwards (2000)). Again, Levelt and van Oostendorp (2007) found no evidence for lexical diffusion, in the sense that the transition to correct production of a segment is independent of the word it is in.

The current theory adopts the classical, feature-based, Jakobsonian view of a high degree of continuity from child to adult language (see also chapter 3. Whereas earlier theories focused on the order of acquisition of segments, contrasts or features (Jakobson, 1941/1968; Beers, 1995), and failed in the sense that they were either too stringent (Jakobson) or too lenient (Beers), the primacy of the *order* of acquisition has been abandoned here – even thought it is possible to predict impossible orders. Rather, by focussing on the *mechanism* employed by learners of a language, we have developed a theory that accurately describes the segment inventory at every stage, for different individual children, maintaining continuity at both the level of the material (features) and the mechanism (constraints).