Input, Intake, and Phonological Development;

The case of Consonant Harmony

0. Introduction

A puzzling observation about Consonant Harmony in Dutch- and English child language is that while for English both Labial and Dorsal Harmony is reported, Dutch data exhibit almost exclusively Labial Harmony. Consonant Harmony (CH) is usually described as a process in which two non-adjacent consonants within a word come to share one or more phonological features. In the most frequently occurring – and most commented on - type of CH the two non-adjacent consonants share their primary Place of Articulation (PoA) feature. Examples from PoA-Harmony in English and Dutch child language are in (1).

(1) PoA Consonant Harmony

a.	duck /dʌk/	[gʌk]	Dorsal Harmony
b.	take /teĭk/	[geĭk]	Dorsal Harmony
c.	soep /sup/ *soup*	[fup]	Labial Harmony
d.	slapen /slapə/ *sleep*	[fapə]	Labial Harmony

From a purely theoretical, in casu, Optimality theoretical perspective, this inter-linguistic difference is not expected; If child language data reflect an initial Markedness >> Faithfulness ordering of universal constraints, and CH forms are considered to be unmarked output forms, resulting from some highly ranked universal Markedness constraint, then we would expect similar CH data for both types of language learners, at least in a particular developmental stage. However, it is not the case that Dutch children either come from, or go to a grammatical stage in which Dorsal Harmony is present, nor do English children start out with only Labial Harmony, and add Dorsal Harmony later.

Our explanation appears simple: the observed difference results from different distributions of Place of Articulation features in the two languages. Lately there has been a growing interest in extra-grammatical influences on phonological acquisition (among others De Boysson-Bardies & Vihman, 1991; Fikkert & Levelt, 2002; Levelt & van de Vijver, to appear; Roark & Demuth, 2000). Several studies have demonstrated that the child is sensitive to frequency information in the language surrounding him (Saffran, Newport & Aslin, 1996). There is, however, no direct mapping between frequency information and acquisition. Levelt & van de Vijver showed that the distribution of syllable types in a language cannot directly predict the order of acquisition of these syllable types. Only in combination with a solid grammatical analysis can frequency information be indicative of order of acquisition. An explanation for acquisition in terms of frequency information therefore only seems simple. In the remainder of this paper we will show how we reached this conclusion. This involves a specific analysis of Consonant Harmony, as proposed in Levelt (1994) and Fikkert & Levelt (2002), the hypothesis that the child's own output (at least partly) determines his/her intake from child directed speech, and calculations in three child directed speech corpora, one Dutch corpus collected by Joost van de Weijer, and two English corpora, one from Bernstein-Ratner (1984) and one from Brent (Brent, M. & J. Siskind, 2001). Below in §1 we start out with a concise introduction to what can be called the general view of phonological acquisition from the perspective of Optimality Theory. In §2 follows the analysis of Consonant Harmony assumed here. In §3 we turn to the child directed speech databases. An initial round of calculations will reveal that a refinement of the analysis is needed. A new round of calculations shows that our refinements lead to a tentative but promising insight into the observed difference between the English - and Dutch child language. Our conclusions can be found in §4.

1. Phonological Acquisition in Optimality Theory

1.1. The General Picture

In most work on acquisition within the framework of OT it is assumed that there is an innate constraint grammar. This grammar comes with an initial ranking where Markedness constraints outrank Faithfulness constraints (a.o. Gnanadesikan, 1995; Levelt, 1995; Pater, 1997), as in (2):

(2) Markedness>>Faithfulness

Such a grammar leads to outputs that are unmarked, rather than faithful to their underlying representations.[1] This complies with numerous observations in child language data that the segmental, syllabic and prosodic structures are all restricted compared to the language to be acquired, and that they are restricted to the unmarked phonological elements. Development consists essentially of reranking the constraints

[1] In OT the underlying representation is referred to as the *input*. However, since this paper deals with child directed speech, i.e. input in the original sense, we will stick to the classic term *underlying representation*, and use *input* to refer to the language surrounding the language learner.

in the grammar until the grammar can faithfully output the types of marked structure that are allowed in the language to be acquired. The underlying representations of the language acquiring child are in general assumed to be similar to the adult output, the motivation being that research by people like Eimas et al. (1971) and Jusczyk (1997) has shown that speech perception is already well-developed by the time phonological acquisition starts.

1.2 OT and Consonant Harmony

In Consonant Harmony of primary Place of Articulation features, consonants in a word carry identical PoA features. Usually, it appears that the PoA feature of C_2, in a $C_1VC_2(V)$ structure, is copied by C_1, replacing the underlying PoA feature of this consonant. In general C_2 is either Labial or Dorsal, and C_1 is Coronal in the adult target. Examples in (3) are repeated from (1) above:

(3) CH examples

a.	duck /dʌk/	[gʌk]	Dorsal Harmony
b.	take /teĭk/	[geĭk]	Dorsal Harmony
c.	soep /sup/ *soup*	[fup]	Labial Harmony
d.	slapen /slapə/ *sleep*	[fapə]	Labial Harmony

As the term Consonant *Harmony* suggests, the phenomenon has been analyzed as a process in which one consonant harmonizes, that is assimilates, to another consonant (Stemberger & Stoel-Gammon, 1991). Levelt (1994), in OT compatible terminology, reanalyzes the phenomenon as resulting from a constraint that requires alignment of a feature, Labial, with a specific edge of a phonological domain, namely the left word edge. Using OT in a more mature state, Goad (1997) presents a similar Alignment analysis of CH. Below in § 2 we will discuss Levelt's Alignment account as elaborated in Fikkert & Levelt (2002). In the remainder of this paragraph we will focus on the account of Pater & Werle (2001) and Pater (2002).

The markedness constraint triggering harmony in Pater and Werle's analysis is not an Alignment constraint, but a constraint that is reminiscent of the assimilation accounts: AGREE. AGREE forces consonants in a certain domain to agree in a specific PoA feature. In their paper AGREE is refined from an initial "Consonants in domain D must agree in place specification" to AGREE-L/R: any consonant preceding/following a consonant with Dorsal place must be Dorsal. In Pater (2002) the analysis is further elaborated, and the AGREE constraints are restated in the following way:

(4) Markedness Constraints driving Harmony (from Pater 2002)

←DOR	Any consonant preceding a dorsal must be homorganic
←DOR→	Any consonant preceding or following a dorsal must be homorganic
←LABDOR	Any consonant preceding a labial or dorsal must be homorganic
←LABDOR→	Any consonant preceding or following a labial or dorsal must be homorganic

Depending on the ordering of these constraints with respect to Faithfulness constraints that require identity between underlying and output PoA features, the consonants of the winning output can end up agreeing in the feature Dorsal or Labial.

A distinction is made between leftward and rightward "agreement" indicated by the arrows. However, rightward 'spreading' is uncommon, and only present in the earliest stages of speech. According to Fikkert & Levelt (2001) such productions reflect a stage in which underlying representations are still unsegmentalized and highly underspecified. This will be elaborated below in §3. Therefore, only the constraints ←DOR and ←LABDOR are further considered here.

An additional assumption in Pater & Werle (2001) and Pater (2002) – but well-accepted in general – is that there is a fixed ranking of Faithfulness constraints pertaining to PoA, whereby FAITH(DOR) and FAITH(LAB) outrank FAITH(COR) (Kiparsky 1994). It is not clear whether there is a fixed ordering between FAITH(DOR) and FAITH(LAB), or whether this can be set on a language particular or individual basis. Pater (2002) follows De Lacy (2002), who assumes a markedness scale Dorsal > Labial > Coronal. Below in (4) are examples of the working of an OT grammar resulting in CH outputs. If both ←DOR and ←LABDOR outrank both FAITH(LAB) and FAITH(COR), this entails that (I) both underlying Labial and underlying Coronal consonants will not surface as such when an underlying Dorsal consonant follows (5a and 5b), and (II) underlying Coronal consonants will not surface as such when an underlying Labial consonant follows (5c). When a Dorsal consonant precedes either a Labial or Coronal consonant, however, no harmony results (5d). Summarizing, the grammar in (5) captures a stage in which both Labial and Dorsal harmony occur.

(5) Dorsal and Labial Harmony

a. duck /dʌk/ -> [gʌk]

/dʌk/	FAITH(DOR)	←DOR	←LABDOR	FAITH(LAB)	FAITH(COR)
☞[gʌk]					*

[dʌk]		*!	*		

b. bug /bʌg/ -> [gʌg]

/bʌg/	FAITH(DOR)	←DOR	←LABDOR	FAITH(LAB)	FAITH(COR)
☞[gʌg]				*	
[bʌg]		*!	*		

c. top /tɑp/ -> [bɑp]

/tɑp/	FAITH(DOR)	←DOR	←LABDOR	FAITH(LAB)	FAITH(COR)
☞[bɑp]					*
[tɑp]			*!		

d. AGREE: /kʌp/ -> [kʌp]

/kʌp/	FAITH(DOR)	←DOR	←LABDOR	FAITH(LAB)	FAITH(COR)
☞[kʌp]			*		
[bʌb]	*!				

In a subsequent stage in this analysis ←LABDOR is demoted below Faith(Cor). Now only Dorsal harmony occurs, as illustrated in (5):

(6) Only Dorsal Harmony

a. top /tɑp/ -> [tɑp]

/tɑp/	FAITH(DOR)	←DOR	FAITH(LAB)	FAITH(COR)	←LABDOR
[bɑp]				*!	
☞[tɑp]					*

b. duck /dʌk/ -> [gʌk]

/ dʌk /	FAITH(DOR)	←DOR	FAITH(LAB)	FAITH(COR)	←LABDOR
☞ [gʌk]				*!	
[dʌk]		*!			*

The grammar, containing the harmony triggering markedness constraints in (4), capture the developmental English data from Trevor (Compton & Streeter 1977) and Amahl (Smith 1973). Apparently, in the English data either both Labial and Dorsal Harmony are found, or only Dorsal Harmony. Pater (2002) deliberately does not include a markedness constraint ←LAB in the grammar, to exclude the possibility of a grammar that results in only Labial Harmony.

One of the original premises of OT is that the constraints that constitute the grammar are universal and innate. This entails that child language can only differ from adult language in the way adult languages differ from each other. These differences result from different rankings of the same set of constraints. The CH data are problematic in this respect, since consonant harmony in primary Place of Articulation features does not exist in adult languages. CH is a child specific phenomenon. What, then, is the status of the markedness constraints ←DOR and ←LABDOR, which appear to have, contrary to OT expectations, no effect whatsoever beyond the age of 3? Pater (2002) does link CH to a phenomenon in adult language, namely PoA assimilation between adjacent consonants. In Korean, for example (Pater 2002 and references therein), Labial and Coronal consonants assimilate to Dorsal consonants, and Coronal consonants assimilate to Labial consonants in clusters, replicating in a local way exactly the non-local CH pattern found in the English child language data. According to Pater, then, the constraints ←DOR and ←LABDOR do affect adult grammars. What changes is their application, which at some point becomes restricted to local situations. While this is a relatively satisfactory solution, two problems remain with this OT analysis: (1) the Dutch CH data are different and (2) CH is an emerging phenomenon in child language.

To start with the second problem, given the OT premise about constraints being universal and innate, and the one about the initial state of the grammar being Markedness >> Faithfulness, and the fact that the CH-triggering constraints are Markedness constraints, it is expected that CH data are present in child language from the onset of speech. This is not the case, however. Fikkert & Levelt (2002) observe in longitudinal data that had been collected from the onset of meaningful speech, from five children acquiring Dutch, that CH data only occur after a period in which the PoA structure of the attempted adult target words is rendered faithfully in the children's productions. Adult target words that result in CH productions in a later stage are not attempted in the earlier stages. While in child productions a lot of unmarked structure is indeed present from the start, in the CH case we find "the emergence of the unmarked." This is elaborated in §2.

The first problem relates to the specific constraints in the proposal. In the longitudinal Dutch data, only Labial Harmony is found. There is no Dorsal Harmony, and no combination of Dorsal and Labial Harmony. This state of affairs cannot be captured by the markedness constraints ←DOR and ←LABDOR. The grammar for Dutch child language appears to require a constraint ←LAB, then, instead of the constraints ←DOR and ←LABDOR.

Relating the two problems is a first step towards a solution: if certain constraints emerge, rather then being present from the start, then they could arise on a language-specific basis. While in the English child language grammar ←DOR and ←LABDOR emerge, in the grammar for Dutch child language it is ←LAB. The how and why will become clear in the next paragraphs.

2. CH resulting from an emerging Alignment constraint

2.1 The Fikkert & Levelt study

In her dissertation, Levelt (1994) stated that CH should not be treated as a phenomenon *an sich*, but that the CH data should be examined and interpreted in the wider context of developing Place of Articulation patterns in words. The following developmental pattern was found in data from children acquiring Dutch as their first language:

(7) Developmental PoA patterns in Words

Stage I: Consonants and Vowels in a word share the same PoA feature, Labial, Dorsal or Coronal, or the vowel is [+low] (/a/)

Examples: pop *doll* /pɔp/ -> [pɔp] (Labial word)

die *that one* /di/ -> [ti] (Coronal word)

koek *cookie* /kuk/ -> [kuk] (Dorsal word)

Stage II: Vowels can be separately specified from Consonants.

Examples: stoel *chair* /stul/-> [tu]

kikker *frog* /kɪkəɹ/ -> [kɪk]

Stage III: The anchor position for Labial is at the left edge of the word.

Dorsal is banned from the left edge and attaches to the right edge.

Examples: bad *bath* /bɑt/ -> [pɑt]

poes *cat* /pus/ -> [pus]

soep *soup* /sup/ -> [fup]

koek *cookie* /kuk/ -> [tuk]

kip *chicken* /kɪp/ -> [pɪp]

Stage IV: No restrictions

Consonant Harmony-like data was found in Stage I and in Stage III. However, the nature of the data was different in the two stages. In Stage I it appeared that the entire word formed the unit for PoA specification. Words consisted either of Coronal consonants and front, i.e. Coronal vowels, or of Labial consonants and round, i.e. Labial vowels, or Dorsal consonants and back, i.e. Dorsal vowels. In Stage III words are “segmentalized”: consonants and vowels can be separately specified for PoA features. In this stage, however, Labial specifications are directed towards the left edge of the word - in winning output structures – while Dorsal is specifically banned from this edge. Therefore, only Labial “harmony” is found in Stage III. Levelt stresses the fact that the CH-like data from Stage III do not result from a relation between two consonants in a word, but from a relation between a specific feature and a specific edge.

In Fikkert & Levelt (2002), henceforth F&L, the above study is replicated and further developed. Longitudinal, developmental data from 5 children acquiring Dutch as their first language were analyzed: Tom (1;0-2;22), Jarmo (1;4.18-2;4.1), Robin (1;4.14-2;4.28), Eva (1;4.12-1;11.8) and Noortje (1;7.14-2;11). These children were selected from a larger database containing data of 12 children – the CLPF corpus (Fikkert, 1994; Levelt 1994), available through Childes (MacWhinney & Snow, 1985) – because they were recorded from the onset of meaningful speech. A total of 8407 utterances were analyzed.

All the words in these utterances were coded for their PoA structure in the following way. Labial consonants were replaced by P, Coronal consonants by T and Dorsal consonants by K. Labial (round) vowels were O, Coronal (front) vowels I and Dorsal (low) vowels A. In words of more than one syllable the stressed syllable was coded. In the case of CVCV (where V = either long or short vowel) the CVC part was coded.

(8) Child Utterance Coding

Target	*Child Production*	*Coding*	*Result*
brood *bread*	bop	b= P o= O p = P	POP
snoep *candy*	fup	f = P u = O p = P	POP
paard *horse*	pat	p= P a = A t = T	PAT
trein *train*	tɛin	t = T ɛi = I n = T	TIT
lachen *laugh*	lɑxə	l =T ɑ = A x = K	TAK

The adult target words were coded in a similar way. In the case of consonant clusters, the PoA feature of the least sonorant consonant was coded. In the case of /sC/ clusters the PoA feature of the /C/ was coded.

(9) Adult Target Coding

Target	*Coding*	*Result*
brood /brot/	br = P	POT

	o = O t = T	
snoep /snup/	sn = T u = O p = P	TOP
paard /part/	p = P a = A rt = T	PAT
trein /trɛin/	tr = T ɛi = I n = T	TIT

In addition to the adult targets 914 words from a list of words that 4-year olds are supposed to know and use were coded. These sources were used to gain information about the PoA-structure of the intake of language learners.

In order to see whether a developmental order could be found for the distribution of PoA features over words, the PoA patterns of the child-utterances and those of the adult targets were aligned on – separate - Guttman scales. Guttman scaling is a procedure for obtaining an order in data, and for checking to what extent an order is followed (Torgerson, 1963). It turned out that the data could be aligned quite nicely, so it could be concluded that the PoA structures were acquired in a particular order over time.

Finally, the distribution of the different PoA patterns was calculated, both in the coded list of 914 words and in the set of attempted adult targets. This was done in order to check whether frequency in the input, or rather, intake, influenced the order of development in production.

The order of acquisition of the distribution of PoA features over words was analyzed in three types of data: (1) the children's productions (2) the children's faithful productions (3) the adult targets. A similar developmental order was found in the three types of data, namely the one in (10). The low vowels /A/ and /a/, coded A, can always appear with any combination of consonants, and is not included in (10). The small "v" stands for 'vowel', and is used when the PoA feature of the vowel is no longer an issue in the patterns. For some children the Stages IIIa and IIIb are collapsed into a single Stage III, and/or the Stages IVa and IVb are collapsed into a single Stage IV.

(10) Order of acquisition of PoA patterns in words

POP/TIT/KOK >	PIP/KIK/TOT >	PvT >	PvK, TvK >	KvT, KvK >	TvP, KvP
Stage I	Stage II	Stage IIIa	Stage IIIb	Stage IVa	Stage IVb

Although the developmental order was similar in the three types of data, the time course in which these developments took place differed. The most striking difference was found between the "produced" and "faithfully produced" data. It turned out that every PoA pattern was produced faithfully before it was produced unfaithfully, that is as a substitute for an adult target word with a different PoA pattern. This is illustrated in (11) with data from Robin:

(11) Robin's Faithful vs. Unfaithful use of PoA patterns

POA pattern in production	First Faithful use	First Unfaithful use for Targets:	
TIT	1;5.13	1;5.27 1;7.15 1;8.26 2;0.20	PIT TOT KAT KIT
PAP	1;5.13	1;6.25 1;8.26 1;11.9	PAT TAP KAP
POP	1;5.27	1;7.15	TOT

		1;7.15 1;8.12 1;9.26	KOT KOP TOP
POT	1;7.15	1;8.12 1;9.12	TOT KOT
TAT	1;7.15	1;7.15	KAT
PAT	1;7.29	1;7.29	TAP
TOT	1;8.26	1;9.12	KOT
TIK	1;9.26	1;10.9	KIK
PIP	1;10.9	1;10.9 1;10.9	TIP KIP
TOK	1;11.9*	1.10.23	KOK

It thus appears that a production pattern can be overgeneralized, just like the English past-tense morpheme –*ed* is overgeneralized during acquisition, and used with strong verbs for some time. The interpretation of forms like *sleeped, singed* and *goed* is that the past-tense rule has entered the grammar. F&L interpret the unfaithful production patterns in a similar way: the pattern is grammaticized. How is it grammaticized? By the emergence of a constraint in the grammar, high ranked, which forces the specific pattern to surface under certain circumstances. The next question is what exactly is overgeneralized? According to F&L the position of a certain PoA feature in the word is overgeneralized. This is illustrated by data from Robin. Around the age of 1;7.15 Robin starts to attempt more and more adult target words with a PvT structure, i.e. Labial consonant, any vowel, Coronal consonant. These targets are produced faithfully. One month later, in the recording at 1;8.12 the first cases of Labial CH appear. Except for TvT and PvP and a quickly disappearing KvK, no other patterns are produced, or even attempted with any frequency.

The idea in F&L is that as soon as words can be segmentalized, after an initial stage in which words are unanalyzed wholes, the learner analyzes his own vocabulary and concludes that Labial consonants are to be found at the left edge of the word. This leads to the emergence of a high-ranked constraint in his grammar, requiring Labial to align with the left edge of the word. As soon as the constraint forms part of the grammar, unfaithful optimal outputs can result: Labial now has to align with the left word-edge in the output, so a faithful analysis of an underlying structure, containing Labial in a non-initial position, cannot be the optimal output. An example of this new grammar, containing the emerged constraint [LABIAL – Labial should be aligned with the left edge - is in (12).

(12) Grammar with [LABIAL

a. poes *cat* /pus/

/pus/	[LABIAL	FAITH
☞pus		
puf		*!

b. soep soup /sup/

/sup/	[LABIAL	FAITH
☞fup		*
sup	*!	

In (a) it is shown that underlying PvT structures can be rendered faithfully. Any change is therefore a change for the worse. In (b), however, the underlying structure is TvP. Labial is not aligned with the left edge, and a faithful output TvP can therefore not be the optimal output. The other output candidate evaluated here, an apparent case of CH, does have Labial aligned with the left edge. It satisfies the high-ranked [LABIAL at the cost of faithfulness. In Robin's case, any TvP or KvP target would result in a PvP output for several months.

More or less at the same time [LABIAL emerges in the grammar, another constraint emerges which bans Dorsal from initial position, *[DORSAL. Before, initial Dorsal consonants could appear in KvK patterns. KvK output patterns then disappear for several months, and underlying KvK patterns are paired with TvK output patterns. Noortje provides a good example of this apparent U-shaped development:

(13) *Dorsal: KvK → TvK → KvK

KOK	koek	→	[kuk]	2;3.7
	klok	→	[kɑk]	2;5.23
KIK	kikker	→	[kɪk]	2;2.21
	kijk	→	[kɛik]	2;5.23
KOK	koek	→	[touk]	2;8.17
	klok	→	[dɔk]	2;8.17
KIK	kijk	→	[tɛik]	2;8.17
	kikker	→	[tikɑ]	2;9.1
KOK	kruk	→	[kyk]	2;9.29
	kuiken	→	[kœyk]	2;10.12

This constraint enters the grammar shortly after the vocabulary is enriched with PvK and TvK-shaped words. This can be illustrated with data from Robin again. From age 1;9.24 PvK and TvK words enter Robin's vocabulary. From the recording at 1;10.6 on KAK and KIK targets, which had, respectively, KAK and KIK outputs before, are paired with TAK and TIK outputs. At this point it appears that the marked Place of Articulation features, Labial and Dorsal each have their designated position in the word. The initial position is reserved for Labial, and Labial elsewhere in the word needs to be licensed by the left word edge. Dorsal is left with the final position. Unmarked Coronal can appear anywhere.

Summarizing F&L's theory so far, apparent cases of Labial consonant harmony, originate with the emergence of the constraint [LABIAL in the grammar. The constraint itself is a grammaticized observation in the active vocabulary of the language learner, which then consists mainly of PvT, TvT and PvP words, that Labial is, at least, attached to the left edge of the word. The constraint *[DORSAL emerges in a similar way. The PoA feature Dorsal is banned from the left edge, and this precludes Dorsal Harmony, the phenomenon that is often observed in Child English.

2.2 The Adult Input

The question that F&L posit next is: why do the Dutch children have the particular order of development of PoA patterns that is observed, and not some different order? Why are PvT words so early and TvP words so late? They turned to the adult input for an answer. It was found that as soon as consonants with different PoA features could be combined in production, at Stage III, the order of acquisition correlated very well with the distribution of the different PoA patterns in the surrounding language, i.e. Dutch. This is shown in (14) for the distribution found in the list of 914 words Dutch children are supposed to know and use at age 6:

(14) Correlation input-development I

List of required words			*Development*	
PvT	**233**	**25,49%**	Stage IIIa	PvT
TvT	143	15,65%		
TvK	**121**	**13,24%**	Stage IIIb	TvK PvK
PvK	**103**	**11,27%**		
KvT	**99**	**10,83%**	Stage IVa	KvT
TvP	**96**	**10,5%**	Stage IVb	TvP KvP
PvP	47	5,14%		
KvP	**47**	**5,14%**		
KvK	25	2,74%		

As can be seen, PvT is the most frequent PoA pattern in the set of 914 Dutch input words, and PvT is also the first pattern that is produced after the initial two "whole word" stages. The K-final patterns TvK and PvK have the next highest frequencies, and also occur next in production. Then follows the K-initial pattern KvT, both in frequency and in appearance. The P-final patterns TvP and KvP have the lowest frequencies and are also produced last.

If attempted adult targets are also an indication of adult input (de Boysson-Bardies & Vihman, 1991), then almost the same correlation input-frequency – developmental order is found there:

(15) Correlation input-development II

Attempted adult targets		*Development*	
PvT	**29%**	Stage IIIa	PvT
TvT	24%		
TvK	**11%**	Stage IIIb	TvK PvK
PvK	**8%**		
KvT	**10%**	Stage IVa	KvT
TvP	**6%**	Stage IVb	TvP KvP
PvP	6%		
KvK	4%		

KvP	**3%**		

As can be seen both in (14) and (15), input frequency does not correlate specifically well with PoA development in the earliest stages, Stage I and Stage II; Particularly KK and PP are of very low frequency, yet are produced very early. This supports the claim that in the initial two stages the language learner has a different, less detailed, lexical representation. However, input frequency does correlate with PoA development as soon as consonants can receive separate PoA feature specifications.

It was stated before that the constraint underlying the apparent cases of Labial consonant harmony, [LABIAL, was built into the grammar based on the child's lexicon. The word patterns in the early lexicon correlate with word patterns in the input. Indirectly, then, Labial harmony in Dutch child language can be traced back to the high frequency of Labial-initial words in the adult input. This, finally, leads to the hypothesis that will be tested in the remainder of our paper. If Labial Harmony in Dutch child language relates to the high frequency of Labial-initial words in the Dutch input, then Dorsal Harmony in English child language should be related to a highly frequent pattern of Dorsal-initial words in the English input. In other words, it should be possible to trace the different types of harmony in the two child languages back to different distributions of PoA in the respective adult input languages.

3. Dutch versus English

3.1 Method

The hypothesis was tested in three large samples of child directed speech, two English corpora and one Dutch (see below in 16). All the words in these samples were coded, using the same method as F&L.

(16) Corpora used in this study

English	*speech directed to*	*in the age of*	*size (word tokens)*
Brent & Siskind (2001)	15 children	0;9 – 1;3	477,679
Bernstein-Ratner (1984)	9 children	1;1 – 1;11	35,475
Dutch			
van de Weijer (1998)	1 child	0;6 – 0;9	173,752

Once the coding was done, we compared the distribution of PoA in the three corpora. If our hypothesis was correct, we expected to find a relatively high percentage of Labial segments in Dutch, compared to an additional relatively high percentage of Dorsal segments in English.

3.2 Initial Results

Initially, our expectations were not confirmed by the data. In (17) the percentages of PoA in initial position in the three corpora are shown. The category labeled “other” includes /h/, which does not have a PoA specification, and orthographic ‘r’, which has many different possible places of articulation depending on the speaker: Coronal /r/, Dorsal /ʀ/ or Retroflex /ɹ/. Numbers are rounded percentages. The overall pattern is similar across the three corpora, with Coronal being the most frequent, followed by Labial, and finally, Dorsal.

(17) PoA in Word-initial position

	Brent	**Bernstein**	**Weijer**
coronal	49	52	51
labial	30	28	25
dorsal	12	9	11
other	9	11	13

Contrary to our expectations, the percentage of Dorsal segments in initial position was not remarkably higher in English than in Dutch. Neither was the percentage of Labial segments higher in Dutch than in English. In fact, it was even slightly lower.

Table (18) shows PoA in the second consonant position, C_2. Again there were no large overall differences between the three corpora. There is a slightly higher preference for Coronal in English than in Dutch, and a slightly higher percentage for Dorsal in Dutch than in English. However, this does not immediately explain away the different harmony types in the two languages.

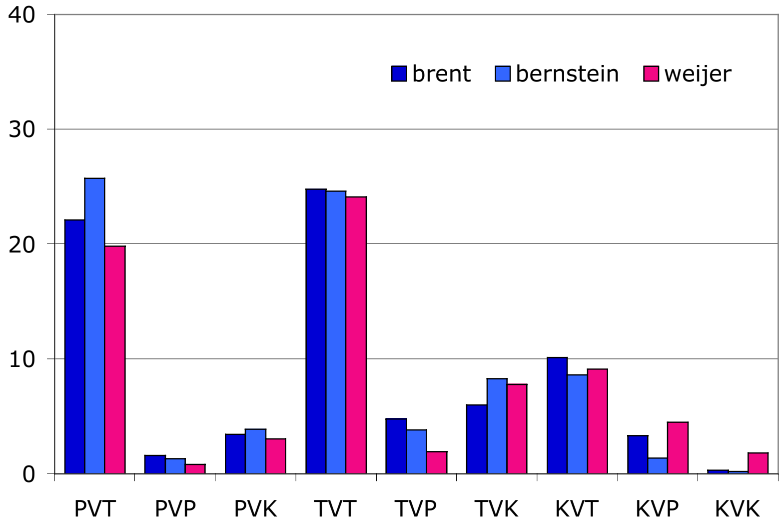
(18) PoA in Word-final position

	Brent	**Bernstein**	**Weijer**
coronal	66	70	58

labial	12	9	13
dorsal	11	12	16
other	11	9	13

The overall similarity between the languages is further illustrated in (19), which shows how evenly distributed each word form is in the three corpora.

(19) PoA patterns in Dutch and English words



Based on these results, and given the hypothesis, we would expect exclusively Labial harmony in English child language, just like in Dutch child language. According to Pater & Werle (2000), however, Labial Harmony does not occur independently of Dorsal Harmony in English child language.

3.3 Plosives, fricatives, nasals

Are there other language specific aspects of the segment inventories of the two languages that could be taken into account? There are some asymmetries in the Dutch and English consonantal inventories:

- English has no Dorsal fricative /x/, whereas Dutch has
- Dutch has no voiced Dorsal plosive /g/, whereas English has
- English has more types of Coronal consonants

The manner aspect of sound might influence the children's intake from the input. This becomes even more likely when we take into account the characteristics of their own productions again. In the early stages, initial consonants are plosives and nasals, fricatives appearing - sometimes much – later. In medial and final position, both stops and fricatives are present from early on, while nasals can appear somewhat later (Fikkert, 1994). It could thus be that children focus on Place of Articulation characteristics of plosives and nasals in initial position, while for the second consonant position they take into account plosives and fricatives.

In order to test this hypothesis, we focused on plosives, nasals, and fricatives in our material. Now clear differences between the two languages started to appear. In (20) are the PoA distributions for initial plosives and nasals. Numbers are rounded percentages:

(20) PoA in C_1 Plosives + Nasals

	Brent	**Bernstein**	**Weijer**

coronal	35	41	**56**
labial	**38**	**35**	30
dorsal	**27**	24	14

One striking difference in the distributions for plosives + nasals is that in Dutch there are far more initial Coronal plosives+nasals than there are initial Labial or Dorsal plosives+nasals, while for English the distribution over the three categories is more or less even. Dutch has a relatively high percentage of Labials in initial position, but in English both marked types of segments, Labial and Dorsal, are well attested in initial position.

For the second consonant position, C_2, either medial in CVCV words, or final in CVC words, we hypothesized, based on the early productions of children, that both plosives and fricatives are early intake. The results for plosives+fricatives in the input are in (21):

(21) PoA in C_2 position: plosives+fricatives

	Brent	**Bernstein**	**Weijer**
coronal	77	79	59
labial	11	7	14
dorsal	11	14	27

Here we find another striking difference between the two languages. We find a salient preference for Coronal in C_2 plosives, more extreme in English than in Dutch. In

Dutch the high frequency of the Coronals apparently goes mainly at the cost of the frequency of Labials, while in English both Labial and Dorsal have low frequencies.

Summarizing the results, specifically in English there is a salient preference for unmarked, Coronal in C_2 position. Both Labial and Dorsal occur with higher frequencies in C_1 than in C_2 position. In Dutch, Coronal does not appear to have a designated position. Dorsal has a higher frequency in C_2 position, while Labial has a higher frequency in C_1 position.

3.4 Discussion of the results

Are these findings the clue to the how and why of the different types of CH in Dutch and English child language? We think this could be the case. It turns out that in English there is a very strong preference for Coronal in C_2 position. The language learner could conclude that C_2 is the designated unmarked position, with respect to PoA. Therefore, the C_1 position is the preferred position for both marked PoA specifications, Labial and Dorsal. During development the language learner can overgeneralize this observation, inferring that Labial and Dorsal should always be licensed by the left edge, C_1 position. In his or her grammar a constraint to this end will emerge. This constraint could very well have the form proposed by Pater, ←LabDor, which leads to both Labial and Dorsal harmony in English child language.

For Dutch there is no such designated unmarked position. There is, therefore, no designated position for marked specifications either, like in English. For Dutch, then, the distribution of Labial and Dorsal over the two positions becomes an important factor, as described in F&L. Labial occurs far more often in C_1 position than in C_2 position: in the Van de Weijer corpus, no less than 78% of all the Labial segments occur in C1 position. Dorsal occurs more often in C_2 position than in C_1 position: 57% of the Dorsals is in C_2 position. Of the Coronals, 47% is found in C1 position, and 53% in C2 position, confirming again the non-biased position for this feature.

4. Conclusions

In this paper we have shown that the different types of CH attested in Dutch and English child language can be traced back to different PoA distributions in the two adult languages. Taking into account the characteristics of the child's own lexicon plays an important role in this finding. For one thing, during development the child builds constraints into his or her grammar on the basis this lexicon. The learner's lexicon is shaped by what he or she takes in from the input, and it turns out that frequencies can determine what is taken in. Fikkert & Levelt (2002) showed that the highly frequent pattern PVT in Dutch leads to the early inclusion of this pattern in the lexicon of Dutch children. Subsequently, a constraint [Labial emerges in the developmental grammar. But the learner's own production abilities also determine the intake from the input: plosive production, for example, leads to intake of information from plosives in the language input. English and Dutch, which appeared to be very similar with respect to the distribution of PoA in words, turned out to be very different when the learner's intake from the input was considered.

References

Bernstein-Ratner (1984). Phonological rule usage in mother-child speech. *Journal of Phonetics*, *12*, 245-254.

Brent, M. & J. Siskind (2001). The role of exposure to isolated words in early vocabulary development. *Cognition* 81, 31-44.

de Boysson-Bardies, B. & M. Vihman (1991). Adaptation to language: evidence from babbling and first words in four languages. *Language* 67, 297-319.

de Lacy, P. (2002). *The formal expression of markedness*. Doctoral dissertation, University of Massachusetts, Amherst.

Compton, A. & M. Streeter (1977). Child phonology: Data collection and preliminary analyses. Ms. Stanford University.

Eimas, P., Siqueland, E., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303-3-6.

Fikkert, P. (1994). *On the acquisition of prosodic structure*, PhD dissertation, Leiden University, HIL Dissertation Series 6.

Fikkert, P. & C. Levelt (2002). Putting Place into Place. Paper presented at the GLOW Workshop on Phonological Acquisition. Utrecht, April 12-13, 2002.

Gnanadesikan, A. (1995, to appear). Markedness and Faithfulness constraints in child phonology. ROA-67. To appear in: Kager, R. , J. Pater & W. Zonneveld (eds.)*Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge University Press.

Goad, H. (1997). Consonant harmony in child language: an Optimality-Theoretic account. In Hannahs, S-J, & M. Young-Scholten (eds.) *Focus on Phonological Acquisition*, 113-142. Amsterdam: John Benjamins.

Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge MA: The MIT

Press.

Jusczyk, P., P. Luce & J. Charles-Luce (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33, 630-645.

Kiparsky, P. (1994). Remarks on markedness. Handout from TREND 2, Stanford University.

Kuhl, P. et al. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science* 277, 684-686.

Levelt, C. (1994). *On the acquisition of place*. PhD dissertation Leiden University, HIL Dissertation Series 8.

Levelt, C. & R. van de Vijver (to appear). Syllable types in cross-linguistic and developmental grammars. In Kager, R., J. Pater & W. Zonneveld (eds.). *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge University Press.

MacWhinney, B. & C. Snow (1985). The child language exchange system. *Journal of Child Language* 12, 271-296.

Pater, J. (1997). Minimal Violation and Phonological Development, *Language Acquisition* 6, 201–253.

Pater, J. (2002). Form and substance in phonological development. In Mikkelson, L. & C. Potts (eds.). *WCFFL 21 Proceedings*, 348- 372.Somerville, MA: Cascadilla Press.

Pater, J. & A. Werle (2001). Typology and Variation in Child Consonant Harmony. In: Féry, C, A. Dubach Green & R. van de Vijver (eds.) *Proceedings of HILP5*, 119-139. University of Potsdam.

Roark, B. & K. Demuth (2000). Prosodic constraints and the learner's environment: a corpus study. In *Proceedings of the 24th annual Boston University Conference of*

Language Development, 597-608.

Saffran, J. Newport, E. & R. Aslin (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, 35, 606-621.

Torgerson, W. (1963). *Theory and methods for scaling*. Wiley, New York.

Weijer, J. van de (1999). *Language input for word discovery*. PhD dissertation, Nijmegen University, Nijmegen, The Netherlands.