A Developmental Grammar for Syllable Structure in the Production of Child Language

Clara C. Levelt

HIL/Free University of Amsterdam, Amsterdam, The Netherlands

Niels O. Schiller

Harvard University

and

Willem J. M. Levelt

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

The order of acquisition of Dutch syllable types by first language learners is analyzed as following from an initial ranking and subsequent rerankings of constraints in an optimality theoretic grammar. Initially, structural constraints are all ranked above faithfulness constraints, leading to core syllable (CV) productions only. Subsequently, faithfulness gradually rises to the highest position in the ranking, allowing more and more marked syllable types to appear in production. Local conjunctions of Structural constraints allow for a more detailed analysis. © 1999 Academic Press

Key Words: phonological acquisition; syllabic development; optimality theory; local conjunction.

Languages have different syllable inventories. In terms of CV-structure, some languages have only the core syllable C(onsonant) V(owel), while other languages have an extensive inventory which includes CV, CVC, V, CCVC, and CVCC.

It is obvious that the form of the active vocabulary in child language is partly determined by the syllable inventory available to them through the state of their grammar; if the inventory of language learners is restricted compared to the inventory of the language to be acquired, some learners turn out to restrict their vocabulary to words that fit this inventory, while others

instead adjust their productions to the available syllable types. The question that we address here is how children acquire the different syllable types in their production and what system underlies the course of development. Except for Fikkert (1994), who treats the development of onsets and rhymes separately, no previous literature addresses this issue. A developmental Optimality Theoretic grammar (Prince & Smolensky, 1993; McCarthy & Prince, 1993) is presented to account for the developmental facts.

SYLLABLE STRUCTURE IN DUTCH

Dutch has an extensive syllable inventory. Except for the core syllable, CV, the Dutch syllable can contain a coda or lack an onset. Both onsets and codas can be branching. The inventory can be captured as (C) (C)V(C) (C), whereby the maximum syllable shape is CCVCC. Clusters of three or more consonants exist, but part of these consonants are in extra-syllabic appendix positions to the word.

MATERIALS AND METHODS

The developmental data come from 12 children acquiring Dutch as their first language (Fikkert, 1994; Levelt, 1994). At the outset of the data-collecting period the children ranged in age from 1;0–1;11 years. Spontaneous speech data were collected every other week for a period of 6 to 13 months. The corpus contains approximately 20,000 utterances and can be found in CHILDES as the CLPF corpus.

All the spontaneously produced utterances of this corpus formed the input to a syllabification algorithm developed by Schiller (Schiller, Meyer, Baayen & Levelt, 1996). Syllabification applied based on the Sonority Sequencing Generalization and Onset Maximization (cf. Selkirk, 1982). The program is supplied with a list of possible syllable onsets in Dutch in order to check the syllabifications. The token frequencies of the resulting syllable types were calculated, and then the syllable structures—in terms of sequences of C(onsonant) and V(owel)—were determined.[1] For this study we focused on the development of primary stressed syllables.

A developmental order of acquisition of syllable types was then deduced by aligning the CV-structures on a Guttman scale at four different points in time: initial recording, first three recordings, first six recordings, and all recordings. Guttman scaling is a procedure for obtaining an order and for seeing to what extent an order is followed (cf. Stouffer et al., 1950). Figure 1 shows the Guttman scale for all the primary stressed syllables found in the initial recordings of all children. The rows and columns are arranged in such a way that a line can be drawn from top left to bottom right, which captures as many of the Xs as possible. The fewer Xs appear to the right of the line, and the fewer empty boxes to the left of the line, the better the fit is.

Figure 2 shows the order of development that could be deduced from the Guttman scaling procedure. There is a general developmental order for the syllable structures CV (1), CVC (2), V (3), and VC (4). Then the learners divide into two groups, one group acquiring onset clusters before coda clusters, the other acquiring coda clusters before onset clusters. The final general acquisition is CCVCC.

---

[1] We thus abstracted away from the quality of segments in the C or V positions (see Fikkert 1994 on how the quality of segments can affect the order of acquisition of onsets and rhymes).

Group A

	cv	cvc	v	vc	cvcc	vcc	ccv	ccvc
Noortje	X							
Elke	X	X	X	X				
Robin	X	X	X	X				
Leonie	X	X	X	X				
Eva	X	X	X	X				
David	X	X	X	X	X			
Catootje	X	X	X	X	X			
Enzo	X	X	X	X	X		X	X
Leon	X	X	X	X	X	X	X	X

Group B

	cv	cvc	v	vc	ccv	ccvc	cvcc	vcc
Jarmo	X	X						
Tom	X	X	X					
Elke	X	X	X	X				

FIG. 1. Guttman scale of syllable types in initial recording sessions.

Group A: CVCC→ VCC → CCV → CCVC

CV → CVC → V→ VC

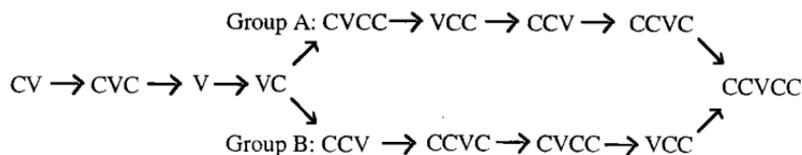Group B: CCV → CCVC →CVCC→ VCC

CCVCC

FIG. 2. Developmental order for syllable type.

OPTIMALITY THEORY

An OT grammar consists of a set of universal, violable constraints on output representations, ranked in a language-specific way. There are two sets of constraints: Faithfulness constraints, which require input and output to correspond to each other, and Structural constraints, which require output representations to be unmarked. These constraints are often in conflict with each other: in order to be faithful to an input representation a candidate will often be structurally marked. The other way around, a candidate can be unfaithful to its input in order to be structurally unmarked in some way. Constraints will thus often be violated. Depending on the language, violations of some constraints are regarded to be worse than violations of some other constraints, and this is expressed by the ranking of constraints in a language-specific constraint hierarchy. For every input a set of possible linguistic analyses—output candidates—is generated. These candidates are evaluated against the constraint hierarchy. The output candidate that is best evaluated is the one which least violates the hierarchy of constraints, and this winning candidate will form the actual output.

Child language is often unmarked compared to the language being acquired. In OT this can be expressed by assuming that constraints come with an initial ranking where Structural constraints outrank Faithfulness constraints (Demuth, 1995; Gnanadesikan, 1995; Levelt, 1995; Tesar & Smolensky, 1996). At this stage outputs will always be structurally unmarked, often at the cost of being faithful to their inputs, which are assumed here to be close to the adult output representations. Language learners, then, need to determine where the Faithfulness constraints are placed in the constraint hierarchy of their language. In other words, they need to discover in which ways the language they are acquiring can be structurally marked. This can be done by comparing the language surrounding them to their own language output. By promoting Faithfulness constraints to positions above Structural constraints in the hierarchy, or by demoting Structural constraints to positions below Faithfulness constraints, the outputs of learners can become more faithful to the input and more marked. Apart from determining the relative ranking of Faithfulness constraints and Structural constraints, learners also need to determine the ranking of Structural constraints among each other. However, here we focus on the relative ranking of Faithfulness with respect to Structural constraints. Reranking of constraints will stop when learners no longer detect differences between their own output and the language surrounding them (Tesar & Smolensky, 1996).

OT AND SYLLABIC DEVELOPMENT

There is a small set of well-motivated constraints on syllable structure in OT:
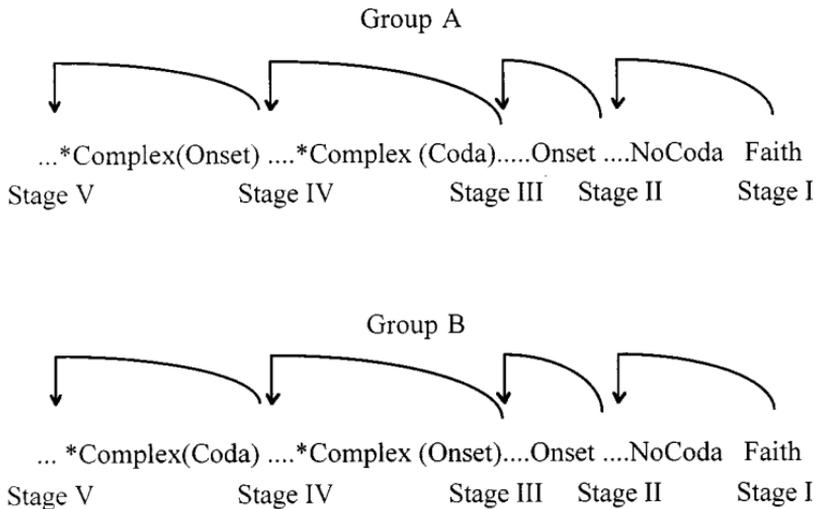
FIG. 3. Developmental OT grammar for syllable type.

ONSET:	Syllables should have an onset.
NO-CODA:	Syllables should not have a coda.
*COMPLEX(ONSET):	Syllables should not have complex (branching) onsets.
*COMPLEX(CODA):	Syllables should not have complex (branching) codas.

These Structural constraints interact with Faithfulness constraints, which we will simply refer to as a single constraint:

FAITH:	Input and output should correspond to each other.

In the initial state, then, it is assumed that all the constraints on syllable structure outrank Faith. In the grammar of Dutch, however, with its highly marked syllable structure, Faith outranks all the constraints on syllable structure.[2]

From the Guttman scaling procedure we deduced a developmental order in the acquisition of syllable types. This means that the initial grammar develops into the final, language-specific grammar for Dutch via several intermediate grammatical stages. These intermediate grammars allow for gradually more marked output structures. Figure 3 shows how over time Faith gradu-

---

[2] We assume that Structural constraints are initially not ranked among each other. In the final grammar they might be ranked according to the order in which Faith has outranked them. However, since the focus of our study is the relative ranking between Faith and Structural constraints, we will actually not say anything about the ranking of Structural constraints among each other.

ally rises in the hierarchy, from the lowest to the highest position, outranking the structural constraints one by one. It can be seen that there are two developmental routes, both leading to the final state where Faith outranks all Structural constraints.

Below are some actual examples from different stages, with more marked structures appearing in every subsequent stage.

OT Stage I: CV

In the initial state of the grammar, all structural constraints outrank faithfulness. Such a grammar leads to outputs that are CV, whatever their input might look like, since only CV satisfies Onset, No-Coda, *Complex-Onset, and *Complex-Coda. The theory excludes the possibility of a grammar that allows only V syllables, or only CVC syllables in the output. This corresponds to our findings: initially only CV is available.

Input	*Output*	*Gloss*	*Subject*
/pus/	→ [pu]	cat	Jarmo (1;5.2)
/klar/	→ [ka]	ready	
/oto/	→ [toto]	car	
/api/	→ [tapi]	monkey	

*OT Stage II: CV, **CVC***

In the second stage both CV and CVC are available. A minimal reranking of the constraints has taken place: Faith has been promoted over the structural constraint No-Coda. With this grammar it is more important to be faithful to an input coda than to be structurally unmarked in this respect.

Input	*Output*	*Gloss*	*Subject*
/pus/	→ [pus]	cat	Jarmo (1;7.29)
/sxap/	→ [hap]	sheep	

*OT Stage III: CV, CVC, **V(C)***

Faith is now ranked above the structural constraint Onset too. From now on onsets are optional, and onsetless inputs will be rendered faithfully in the ouput.

Input	*Output*	*Gloss*	*Subject*
/oto/	→ [oto]	car	Jarmo (1;6.13)
/ap/	→ [ap]	monkey	Jarmo (1;7.15)

*OT Stage IV. Group A: CV, CVC, V(C), **(C)VCC**. Group B: CV, CVC, V(C), **CCV(C)***

At this point two different rerankings show up. For Group A, Faith now outranks *Complex-Coda, but it is still outranked by *Complex-Onset, while for Group B it is the other way around. Learners from Group A can faithfully produce complex codas, but not complex onsets, while the learners in Group B can faithfully produce complex onsets, but no complex codas.

Group A: Complex coda is retained, complex onset is reduced

Input   Output Gloss Subject

/plant/ → [bant]   plant from Group A: Cato (1;10.11)

Group B: Complex onset is retained, complex coda is reduced

Input   Output Gloss Subject

/erst/ → [it]   first   from Group B: Tirza (1;8.5)

/trein/ → [treinə] train   from Group B: Tirza (1;8.5)

GUTTMAN DATA VERSUS OT STAGES

Figure 4 shows that the OT analysis indeed captures the general developmental course deduced from the Guttman scale. However, not every stage deduced from the Guttman scale receives a separate OT analysis: in the OT analysis some stages are collapsed. As can be seen in Fig. 4, the OT grammar predicts a single stage III, which combines the stages V and VC from the Guttman scaling, a single stage IV, which combines either the Guttman stages CVCC and VCC (Group A) or the stages CCV and CCVC (Group B) and a single stage V, which combines either the deduced stages for CCV, CCVC and CCVCC (Group A) or CVCC, VCC and CCVCC (Group B). This same problem would arise for, for example, a parametric analysis of the facts. An OT analysis of every single stage is possible, however, when
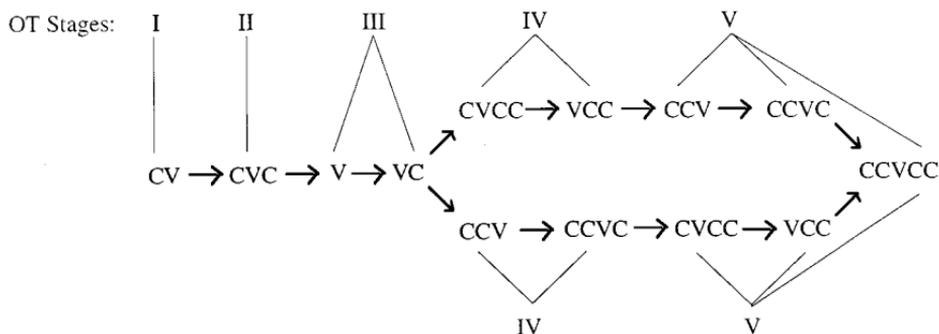

FIG. 4. Comparison between OT stages and developmental stages.

some machinery is added to the grammar. It has been proposed that OT allow for so-called ''Local Conjunctions'' of constraints (Smolensky, 1993; Kirchner, 1996; Ito & Mester, 1998). Two (or more) constraints can form a conjoined constraint which is ranked above the individual constraints it consists of in the hierarchy. A conjoined constraint is violated only if all the constraints in the conjunction are violated. With this mechanism an account for every stage is possible. For example, a conjoined constraint ONSET & NO-CODA would be invoked to account for the stage in which only the types CV, CVC and V occur, to the exclusion of VC. FAITH would, at this stage, outrank both the individual constraints ONSET and NO-CODA, but would still be outranked by ONSET & NO-CODA. Both the structures V and CVC satisfy this high-ranked constraint: although they either violate ONSET (V) or NO-CODA (CVC), they do not violate both ONSET and NO-CODA at the same time. The structure VC, however does combine the two vices of having a coda and lacking an onset, and thus violates the conjoined constraint ONSET & NO-CODA. Since this Constraint is, at this stage, ranked above FAITH, this violation is fatal.

With the help of Local Conjunction, it is thus possible to present a more detailed analysis of the facts. However, it is a powerful device, and conditions on exactly what can be combined need to be formulated. What needs to be investigated next is whether the more specific stages deduced from the Guttman scales, requiring Local Conjunctions, are grammatically real and have a counterpart in any of the world's languages, or in other child languages.

SUMMARY AND CONCLUSIONS

A developing OT grammar is able to account for the order of development of syllable structure by children acquiring their first language. In this grammar Structural constraints concerning syllable well-formedness initially outrank Faithfulness constraints. This leads to rigorously unmarked structure, namely the syllable CV, as the only possible output of the grammar. By subsequent promotions of Faithfulness in the ranking, outranking the syllable structure constraints one by one, more and more marked structures can appear in the output. Five of the developmental stages and two developmental routes, deduced from a Guttman scaling procedure, could be accounted for grammatically this way. An account of all developmental stages was possible with the help of Local Conjunctions of Structural constraints.

REFERENCES

Demuth, K. 1995. Markedness and the development of prosodic structure. In J. Beckman (Ed.), *Proceedings of the North East Linguistic Society,* Graduate Linguistic Student Association, University of Massachusetts, Amherst.

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

Gnanadesikan, A. 1995. *Markedness and faithfulness constraints in child phonology.* Unpublished manuscript, University of Massachusetts, Amherst. [ROA-67][3]

Ito, J., & Mester, A. 1998. *Markedness and word structure: OCP effects in Japanese.* Unpublished manuscript, University of California, Santa Cruz. [ROA-255]

Kirchner, R. 1996. Synchronic chain shifts in optimality theory. *Linguistic Inquiry,* **27,** 341–350.

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

Levelt, C. 1995. *Unfaithful kids: Place of articulation patterns in early child language.* Paper presented at the Department of Cognitive Science, Johns Hopkins University, Baltimore, and at the Department of Linguistics, MIT, Cambridge, MA.

Prince, A., & Smolensky, P. 1993. *Optimality theory: Constraint interaction in generative grammar.* Unpublished manuscript, Rutgers University, New Brunswick and University of Colorado, Boulder.

Schiller, N., Meyer, A., Baayen, H., & Levelt, W. 1996. A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics,* **3,** 8–28.

Selkirk, E. 1982. The syllable. In H. van der Hulst and N. Smith (Eds.), *The structure of phonological representations II.* Foris: Dordrecht.

Smolensky, P. 1993. *Harmony, markedness, and phonological activity.,* Paper presented at Rutgers Optimality Workshop-1, Rutgers University, New Brunswick.

Stouffer, S., Guttman, L., Suchman, E., Lazarsfeld, P., Star, S., & Clausen, J. 1950. *Measurement and prediction.* Princeton, NJ: Princeton University Press.

Tesar, B., & Smolensky, P. 1996. *Learnability in optimality theory.* Technical Report JHU-CogSci-96-3, Johns Hopkins University, Baltimore.

[3] ROA papers can be downloaded from http://ruccs.rutgers.edu/roa.html.