

# Emergent Phonological Constraints

## The acquisition of \*COMPLEX in English

**Jeroen van de Weijer**

Shanghai International Studies University  
jeroen@shisu.edu.cn

**Abstract:** Emergent Phonology seeks to minimize the role of Universal Grammar in linguistics by investigating how units such as distinctive features, segments, words, morphemes, and syllables, and other aspects of grammar, such as phonological, morphological or syntactic rules and conditions, emerge in the course of acquisition and language use, rather than as part of an innate language capacity. An obvious candidate for being acquired rather than being innate are the phonological constraints that take a central place in Optimality Theory. In this paper I discuss whether, and if so how, a constraint like \*COMPLEX ‘No complex onsets’, which is assumed to be active in the acquisition of English and many other languages, could be acquired on the basis of the data to which the English language-learning child is exposed. If this constraint is acquired, it lessens the burden on any innate capacity, which is hypothesized to contain more general, cognitive strategies—perhaps not exclusive to linguistics.

**Keywords:** Emergent Phonology; acquisition; English; consonant clusters; Optimality Theory; innateness; constraints; Universal Grammar

### 1. Introduction

Emergent Phonology is a relatively recent paradigm in phonological theory, which proposes that units such as distinctive features, segments, words and elements of the prosodic hierarchy such as syllabic constituents and feet, **emerge** in the course of the acquisition process rather than being stipulated as part of Universal Grammar (UG), here understood as an innate capacity specifically for language (e.g., Bates et al. 1998; Beckman & Edwards 2000; Mielke 2008; Archangeli & Pulleyblank 2015; 2017; see also van de Weijer 2009; 2012; 2014; 2017; van de Weijer & Sloos 2013; van de Weijer & Tzakosta 2017, and many others). The question of whether Universal Grammar exists and, if so what it consists of, has had a long and controversial history in modern linguistics, starting with Chomsky (1965). I will not review this long history here (see, recently, Everett 2016, for instance, who takes a similar stance as that put forward below). In my

view, we should take an open, empirically-based approach to this issue, much like that advocated by Culicover (2016, ix):

“[W]e **must** try to show that there is an alternate explanation to any and all proposed specific aspects of ‘innate structure’, [...]. If [such attempts] are on the right track [...], they will have shed light on some mysteries about why languages have certain properties and how language gets into our heads. If they turn out in the end to be wrong about something (an unfortunate inevitability in our line of work, if we are willing to take risks), they may have lent additional credence to our understanding of what is likely to be part of the “innate structure.”  
(emphasis in the original)

That is, can elements of linguistic structure be explained as the result of the way language is acquired and used, rather than being stipulated as part of “innate structure” (i.e., UG)? The biggest advantage of this approach, in my view, is that it offers new, obvious possibilities of bridging the divides between “theoretical” phonology and “usage-based” linguistics (e.g., the very inspiring work by Bybee (2001; 2010), and so-called “cognitive” or “functional” approaches). That is, the framework offers a chance to unify various strands in linguistic research that were so far deemed diametrically opposed. In such a unified approach, linguistic units are shown to exist, i.e., to play a systemic role in language use, but not as a result of stipulation, but as a result of the way language functions.

This discussion is relevant to all sectors of linguistics, but here I will apply it to only a small, specific part of one particular phonological theory: the phonological constraints of Optimality Theory (Prince & Smolensky 1993). It is usually assumed that these constraints are part of Universal Grammar, so that the question how they might be acquired is not even raised, or receives only passing comment. There is some discussion in the literature that attempts to show why such constraints might be innate, and could not be acquired (Gnanadesikan 2004). In my view, however, the constraints can quite easily be shown to be acquirable on the basis of language data alone, obviating the need to postulate their presence in UG. This will be fleshed out below.

Concretely, this paper focuses on the acquisition of English by native speakers. Infants acquiring their first language go through a long period of acquisition on their way to full competence, a path that is riddled with errors and involves a lot of variation. During the last fifty-odd years, it has become possible to record, process and analyse data concerning language acquisition in ever greater detail and in systematic ways, leading to much greater understanding of the facts, the differences and similarities between acquisition in different languages and the formation of a number of theories

regarding these processes. We can therefore now begin to focus on the question **why** children make the errors that they do. Two approaches emerge that seek to answer this question: one can be dubbed the “maturational” approach, which holds that children are simply not ready – anatomically, cognitively and/or neurologically – to process or produce the sounds and patterns that exist in adult language (cf. Hale & Reiss 2008, §2.7 for discussion). Since children develop in the same way around the globe, their patterns of development will also be roughly similar, *modulo* the precise language they are acquiring. The second “linguistic” or “grammatical” approach regards the errors that language acquirers make in the light of their developing grammatical system. Such approaches typically assume an “initial state”, given by Universal Grammar (UG), which forms the basis for the development of different L1 languages. An example of the latter approach is Optimality Theory (OT), which has been very successfully applied in analyses of acquisition and other data in many different languages during the last twenty years (see e.g., Johnson & Reimers 2010 and references cited there for specific applications to acquisition). The second approach also predicts that language acquisition is comparable across different languages, by virtue of the shared initial state and its shared pool of universal constraints.

In this paper I will not attempt to choose between the maturational approach and the linguistic approach, mainly because I think the choice here represents a false dichotomy, and should be rephrased in terms of the division of labour between maturation and the emerging grammar (see e.g., Wexler 1990). My main topic is the idea that the linguistic concepts that play a role in the acquisition process are not necessarily universal (part of “innate structure”). One key assumption in a UG-based theory like OT is that constraints, which form a key part of the theory, are universal and therefore do not need to be acquired. Constraints and constraint ranking give the theory a powerful typological outlook which has been verified for a considerable number of linguistic areas, also outside phonology. OT has therefore been much more successful than its rule-based predecessors. The assumption of universality is not a necessary one, however. It is not necessary to burden the child with an innate stockpile of sometimes highly specific constraints. Simpler cognitive strategies are available that enable us to derive constraints in a natural fashion, as I will show on the basis of data from English. For these strategies to work, it is necessary to assume that children are sensitive to statistical patterns in the linguistic data they are exposed to. Natural though this assumption may sound, this has not been common practice in generative frameworks at all, which have

related frequency considerations and other such notions to the realm of performance. To the extent that it is possible and even mandatory to take into account performance-related factors for a full understanding of developing competence, this study represents a step on the way towards the construction of a comprehensive theory of language, incorporating aspects of both competence and performance and assigning due value to both (as advocated explicitly by Jackendoff 2007).

This paper is organised as follows: in the next section I will present typical facts of English acquisition and raise the question how these should be analysed. After a brief discussion of the two main approaches to this problem (maturation vs. universalism), I will outline the OT approach. In section 3 I discuss the question of universality of OT constraints and argue in favour of an approach in which constraints are acquired, and not innate. In section 4 I discuss a number of implications of this approach and conclude.

## **2. Acquisition of English consonant clusters**

It is well known that children acquiring English (or any other language) (re)produce forms they are exposed to in novel, sometimes charming ways. Not all such changes should be characterised as simplifications, because in fact there are many strategies for the child to produce target (adult) outputs. In this paper I focus on the realisation of consonant clusters in such English child language, but the argument could be made for a number of other phenomena as well (both in the segmental realm, e.g., the realisation of specific vowels or consonants, and in the suprasegmental realm, e.g., the realisation of particular stress patterns or words with a specific number of syllables) and for any other language (see Johnson & Reimers 2010 for a cross-linguistic exploration). Consider the examples in (1) below, taken from Johnson & Reimers (2010, 20), whose data are based on previous studies by Smith (1973) and Gnanadesikan (2004). In these data the child's phonetic realisation is given first, followed by the adult forms in spelling:

(1)	Amahl (Smith 1973)	Gitanjali (Gnanadesikan 2004)
	[b̥ei] play	[kin] clean
	[b̥u] blue	[piz] please
	[ʃin] clean	[fen] friend
	[ʃai] sky	[dɔ] straw
	[b̥ot] sport	[gin] skin
	[b̥un] spoon	[bun] spoon

Both children are around two and a half years old at the time of these utterances, and their pattern is typical for children acquiring English as their first language: the consonant clusters in the adult data are simplified by “leaving out” one or more of the consonants in the consonant cluster. In the case of obstruent–sonorant clusters (*bl-*, *fr-*, *kl-*, etc.) the sonorant is omitted; in the case of *s* plus stop clusters (*sp-*, *sk-*, etc.) the *s* is omitted. These data, which could easily be multiplied for any child acquiring English or other languages with consonant clusters, are the focus of this paper.

There are two ways of approaching the data in (1). Why do children make “mistakes” like these? In the first approach, referred to as “maturation”, it is simply held that the child at 2.5 years old is not yet ready to produce clusters like these. The reasons for this could be in the development and training of its vocal organs – obviously, producing a consonant cluster requires more highly skilled control of the vocal apparatus than the production of a single consonant – or in a combination of control with neurological/cognitive factors: the child is not yet equipped to produce the clusters that will eventually be needed to be a fully competent speaker of the native language. Note that, as is also well known, the problem is not in perception, because children will not accept deliberately mispronounced words of the type [b̥ei] for ‘play’. Since clusters are found in the adult stage of the language, the child will eventually have to acquire them and all normally-developing children eventually will. This approach predicts that child language development will (at least initially) proceed along stages that are comparable across different languages, since maturation will be similar across children, and only be influenced by differences in the ambient languages as language acquisition proceeds.

A second approach is to relate the acquisition of phonological patterns, and the different stages therein, to the developing grammatical system of a language acquirer. Above, this was referred to as the “universalist” or “grammatical” approach. Since adult speakers have a grammar, which leads them to understand, produce, accept or reject utterances in their language, the question of language acquisition becomes equivalent to that of grammar acquisition, besides acquisition of the lexicon. Since adult speak-

ers of English produce consonant clusters and child speakers do not, these two “groups” of speakers must have different grammars. The question what the child grammar looks like and how it develops into the adult grammar becomes an obvious object of research. This is particularly apparent in Optimality Theory (Prince & Smolensky 1993), where grammars are characterised in terms of constraints (not transformations, or rules, as in earlier generative approaches).

The OT approach to language acquisition has been extremely influential in recent years (see e.g., contributions to Dekkers et al. 2000; Kager et al. 2004; Johnson & Reimers 2010, and many others). Its basic premise is that children start out with an inventory of universal constraints, which are basically of two types: markedness constraints and faithfulness constraints. Markedness constraints favour simple structures, e.g., basic vowels or simple prosodic structure. Faithfulness constraints inhibit modifying underlying forms, e.g., by deletion, insertion or segmental changes. The grammar of a language is characterised as a language-particular ranking (or hierarchy) of the language-independent constraints. For language acquisition, this means that the language variety of the child (e.g., Amahl in (1)) can be characterised as a particular constraint hierarchy  $C_1$  and that the language variety of an adult speaker can be characterised as a particular constraint hierarchy  $C_2$ . Language acquisition consists of the path (or rather paths, see Tzakosta 2004) between  $C_1$  and  $C_2$ . Apart from these hierarchies, OT also hypothesises an “initial state”,  $C_0$ , which reflects the initial setting of the constraint hierarchy, present at or before birth, which children start out with at the onset of language acquisition. In the OT perspective, both the content of the constraint set, as well as the initial ranking (or non-ranking) between the constraints is provided by Universal Grammar and therefore do not need to be acquired or learned (see e.g., Kager 1999; Tesar & Smolensky 2000).

Let us briefly review how such an account would work for the case at hand. The child English forms in (1) differ from the adult forms in that the child forms do not permit consonant clusters while the adult forms do. That is, the child forms are less ‘marked’ than the adult forms. In other words, the child forms satisfy a markedness constraint that the adult forms violate. The constraint in question, \*COMPLEX, can be formalised as in (2) below. This is a widely recognised markedness constraint, which has also been applied to a wide variety of other languages (see e.g., Kager 1999, 97, among many others).

- (2) \*COMPLEX  
 \*CC (‘No word-initial consonant clusters’)

The children satisfy the markedness constraint \*COMPLEX at the expense of violating a faithfulness constraint against the deletion of consonants. This constraint, which is usually referred to as the correspondence constraint MAX (McCarthy & Prince 1995), is informally characterised in (3):

- (3) MAX  
 ‘No deletion’

We can now see that in the child grammar the ranking is markedness  $\gg$  faithfulness, and in the adult grammar the ranking is the other way around (at least for the specific constraints given in (2) and (3)). That is, in child language relatively unmarked forms are favoured as a result of the grammar that applies at that stage, whereas the adults have a different grammar. In fact, Optimality Theory claims that children start out with **all** markedness constraints being ranked high in the initial state, so that children’s first productions will always be relatively unmarked compared to the adult forms (Tesar & Smolensky 2000). In response to the adult forms they perceive, children will change their constraint hierarchy, promoting faithfulness constraints (or rather: demoting markedness constraints), in order to conform more and more to the adult grammar. A number of algorithms are available that describe the development of the hierarchy (or hierarchies); these will not concern us here (see Tesar & Smolensky 2000 for discussion and illustration of the so-called Constraint Demotion Algorithm, and Boersma & Hayes 2001 and Goldwater & Johnson 2003 for alternative algorithms, the Gradual Learning Algorithm and the Maximum Entropy Model, respectively). Note for now that the OT approach, too, predicts that children will go through comparable developmental stages on their way to full acquisition, since they start with identical initial states and the limits of variation are determined by the (universal) constraint set.

To sum up so far, we have considered two approaches to the problem of language acquisition and to the kinds of errors made by children on their way to full competence: the maturational approach and the linguistic approach. It should be noted that these two approaches are not necessarily incompatible with each other: rather, they should be regarded as two aspects of one and the same developmental process. As children mature, gain more control over their vocal organs and develop their general cognitive skills and memory capacity, their grammar will be formed. The exact contribution of maturational factors *vis-à-vis* the development of grammar needs further exploration. In the next section, I will discuss the burden that OT places on Universal Grammar.

### 3. OT constraints: innate or acquired?

The Optimality approach to language acquisition has been very successful because it outlines a very clear research programme and because it makes clear predictions which are, at least in principle, testable. For instance, it predicts that different constraint rankings correspond to different grammars. Different grammars should be potential grammars, so they should either be found in the languages of the world or qualify as learnable (see e.g., Levelt & van de Vijver 2004). Such typological predictions are at the heart of Optimality Theory (McCarthy 2002), and research in this area has enjoyed resounding success.

One aspect that could be questioned, however, concerns the status of constraints in this approach. In the standard OT approach, constraints are universal: all languages make the same generalisations – they just rank them differently (Ellison 2000). All constraints are present at birth and language acquisition can be basically characterised as a constraint (re)ranking process. The idea that such explicit linguistic-specific constraints as in (2) and (3) are present in our genetic make-up could be questioned, however. In recent years, it has become increasingly recognised that innate cognitive abilities should be as general as possible, applicable in a wide range of fields (e.g., Croft & Cruse 2004, 2). This is also underscored in the Emergent Phonology paradigm. The question of the innateness of constraints therefore deserves serious consideration. The OT viewpoint is that a constraint such as \*COMPLEX **could not** be acquired, because the speech data to which children are exposed (the adult forms) contain many violations of this constraint (e.g., words like *play*, *friend*, *spoon*: see (1)). Hence, since the constraint could not be acquired on the basis of the data, and since the child's grammar conforms to a grammar in which this constraint plays a decisive role, the constraint \*COMPLEX must, according to mainstream OT, be innate (Gnanadesikan 2004).

If this argument worked, it would be strong evidence for the innateness of constraints, and hence in favour of a UG approach. However, it is a little too early for such a conclusion. In my view, the constraint \*COMPLEX could quite easily be acquired on the basis of the data to which the child is exposed. One crucial assumption is necessary, though: it is vital to take lexical frequency into consideration. Let us consider the words to which children who are acquiring English as their first language are exposed. The 150 most frequent words in child-directed English, taken from the CHILDES corpus (MacWhinney 2000), are given in (4):



- (4) you, the, it, a, and, to, I, that, oh, on, what, in, we, is, do, are, no, there, that's, your, it's, don't, of, one, going, have, this, well, can, not, right, he, yes, like, now, got, think, go, with, look, put, then, for, you're, they, just, was, all, want, up, some, she, see, Thomas, at, get, be, me, them, because, what's, isn't, did, know, out, but, come, here, little, so, her, if, very, when, you've, there's, I'm, down, didn't, shall, yeah, okay, Mummy, big, nice, good, back, bit, he's, can't, about, where, off, would, say, had, my, were, or, his, two, does, they're, more, him, doing, please, I'll, doesn't, we've, these, haven't, aren't, has, as, dear, why, let's, again, have to, over, we'll, those, make, she's, from, gone, will, play, need, take, really, an, how, I've, mhm, car, other, another, Daddy, round, been, who, sit, eat, where's, time, something, alright, too

The data show that there are no words with onset clusters among the 100 most frequent words in child-directed speech, and only three among the 150 most frequent ones. So, surprisingly perhaps, words with onset clusters are very uncommon in everyday English speech. We know that children are sensitive to statistical skewing in the distribution of words (and any other phenomena) to which they are exposed (e.g., Saffran 2001; Dawson & Gerken 2011, etc.). In view of the overwhelming presence of words without initial consonant clusters, it is not surprising at all that children would generalise across data such as that in (4), i.e., assume that *no* words have initial clusters, and structure their own production accordingly (reinforced by the obvious phonetic fact that consonant clusters are simply harder to produce than singleton consonants). In this way we are led to the conclusion that a constraint like \*COMPLEX can be learned on the basis of the available data, if we take into account the (token) frequency of different words in English. It is not necessary to assume that constraints like \*COMPLEX are innate; rather, such constraints may be acquired as initial hypotheses about the language, made by the child purely on the basis of the data it is exposed to. The same argument can be made for many other markedness constraints, e.g., against nasal vowels, coda consonants, “marked” vowels and other structures in English and other languages (e.g., van de Weijer & Sloos 2013, who demonstrate similar facts for several markedness constraints in French).

In the light of the evidence in favour of the idea that grammars are best described as constraint hierarchies, I assume that these “initial hypotheses”, or generalizations across (frequent) data, become grammaticalised, in the sense that they will form the content of the grammar. At a later stage, the child will learn that there are, in fact, initial consonant clusters in the language it is acquiring, and thus be led to revise its initial hypothesis. This could be reflected in the grammar as a shift in the constraint ranking between the markedness constraint and the faithfulness constraint MAX. In the next section I will discuss the implications of this approach.

#### 4. Implications and conclusion

In the previous section I showed that a constraint like \*COMPLEX can be learned by L1 speakers of English if the frequency of words in English is taken into consideration. This shifts a lot of the burden of Universal Grammar: instead of containing formally explicit constraints specific for linguistics or phonology, the only ability that is needed is to make generalizations across a pool of data – a general cognitive strategy that is operative in any field of human experience. It relates a usage-based, performance factor such as frequency of occurrence to grammatical competence, and thus represents a step on the way to a comprehensive theory of language use and linguistic competence. In yet other words, it provides an approach to see how exactly competence (the child's knowledge of grammar) is based on performance (the language input to which the language-learning child is exposed).

Of course, it remains to be seen whether this approach is on the right track. More so than the universalist Optimality approach, the usage-based approach is empirical, because it predicts that statistical patterns in the adult data will be reflected in the child's productions. Consider the words in (4) once again. Many words have coda consonants, although these are also marked from a cross-linguistic view: Optimality Theory proposes a constraint NOCODA which explicitly penalises coda consonants (Prince & Smolensky 1993). OT is agnostic about whether children will first learn to produce initial consonant clusters or coda consonants. The usage-based approach is more explicit in this respect: in view of the higher frequency of coda consonants in frequent words, coda consonants are predicted, all things being equal, to be acquired earlier (which may turn out to be correct, as, for instance, the data in (1) suggest). Note, of course, that all things are often *not* equal, so these predictions need to be carefully checked. Research on order of acquisition in French (van de Weijer & Sloos 2013) suggests that this approach is at least on the right track.

Another implication of this approach is that some constraints will not be learned. Consider a language like Mandarin Chinese which has no consonant clusters at all. Will children acquiring Chinese learn a constraint against consonant clusters? This is not a necessary assumption: faithfulness to the adult forms is all that is required. In this sense, constraints are not universal, as in mainstream OT; rather, they are acquired on an as-needed basis. (Note that Chinese children will have difficulties with consonant clusters when they start learning English as a foreign language; this means that the patterns in their own language interfere with those in the L2.)

To conclude, I have discussed a very small sample of acquisition data with the aim of showing how current linguistic theory seeks to explain the kinds of errors language acquirers make. From a linguistic viewpoint, these are not “errors” but reflect the grammar-in-progress that the child has built up on the basis of the data that are available to them. Since grammars are considered to consist of a set of ranked constraints in Optimality Theory, the question is how these grammars start out and how they develop. As Everett (2016) puts it, “The question needing to be answered is ‘Where did the phonological knowledge come from?’”. OT answers this question relying on a rather large dose of innateness in its standard account: both the content (i.e., the constraints) and the organisation (i.e., the initial state and the algorithm for reranking) are assumed to be universal. I showed, however, that this is not a necessary assumption. If frequency of use is taken into account, constraints come from the only source that should be available in an empirical approach to language: from the language the children are exposed to. On this view, the constraints can be considered as “working hypotheses” that the child construes on the basis of the majority data and reranking reflects the adjustments to these hypotheses. In this way, usage of language forms the basis of grammatical principles and paves the way for a further integration of performance and competence in language. Obviously, this paper has only explored this idea for only a tiny portion of acquisition-related data, and the strong position taken here remains to be verified for other patterns of acquisition data in English and other languages.

### Acknowledgements

I am very grateful to two *Acta Linguistica Academica* reviewers for helpful and insightful comments, which considerably improved this paper. I also thank Marjoleine Sloos and Marina Tzakosta who commented on an earlier version. The usual disclaimers apply.

### References

- Archangeli, Diana and Douglas Pulleyblank. 2015. Phonology without Universal Grammar. *Frontiers of Psychology* 6. 1229.
- Archangeli, Diana and Douglas Pulleyblank. 2017. Phonology as an emergent system. In S. J. Hannahs and A. R. K. Bosch (eds.) *The Routledge handbook of phonological theory*. London: Routledge.

- Bates, Elizabeth, Jeffrey L. Elman, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi and Kim Plunkett. 1998. Innateness and emergentism. In W. Bechtel and G. Graham (eds.) *A companion to cognitive science*. Malden, MA & Oxford: Blackwell. 590–601.
- Beckman, Mary E. and Jan Edwards. 2000. The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development* 71. 240–249.
- Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32. 45–86.
- Bybee, Joan L. 2001. *Phonology and language use* (Cambridge Studies in Linguistics 94). Cambridge: Cambridge University Press.
- Bybee, Joan L. 2010. *Language, usage, and cognition*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Croft, William and David Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Culicover, Peter W. 2016. Foreword. In M. H. Christiansen and N. Chater (eds.) *Creating language – Integrating evolution, acquisition, and processing*. Cambridge, MA: MIT Press. vii–x.
- Dawson, Colin and LouAnn Gerken. 2011. When global structure “explains away” local grammar: A Bayesian account of rule-induction in tone sequences. *Cognition* 120. 350–359.
- Dekkers, Joost, Frank R. H. van der Leeuw and Jeroen van de Weijer (eds.). 2000. *Optimality theory: Phonology, syntax, and acquisition*. Oxford: Oxford University Press.
- Ellison, T. Mark. 2000. The universal constraint set: Convention, not fact. In Dekkers et al. (2000, 524–553).
- Everett, Daniel L. 2016. An evaluation of Universal Grammar and the phonological mind. *Frontiers in Psychology* 7. 15.
- Gnanadesikan, Amalia E. 2004. Markedness and faithfulness constraints in child phonology. In Kager et al. (2004, 73–108).
- Goldwater, Sharon and Mark Johnson. 2003. Learning of constraint rankings using a maximum entropy model. Paper presented at Proceedings of the Stockholm workshop on variation within Optimality Theory.
- Hale, Mark and Charles Reiss. 2008. *The phonological enterprise*. Cambridge: Cambridge University Press.
- Jackendoff, Ray. 2007. A whole lot of challenges for linguistics. *Journal of English Linguistics* 35. 253–262.
- Johnson, Wyn and Paula Reimers. 2010. *Patterns in child phonology*. Edinburgh: Edinburgh University Press.
- Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.
- Kager, René, Joe Pater and Wim Zonneveld (eds.). 2004. *Constraints in phonological acquisition*. Cambridge: Cambridge University Press.
- Levelt, Clara and Ruben van de Vijver. 2004. Syllable types in cross-linguistic and developmental grammars. In Kager et al. (2004, 204–218).

- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk* (Third edition). Mahwah, NJ: Lawrence Erlbaum.
- McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. In J. Beckman, L. W. Dickey and S. Urbanczyk (eds.) *Papers in Optimality Theory* (University of Massachusetts Occasional Papers in Linguistics 18). Amherst MA: Graduate Linguistic Student Association. 249–384.
- McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.
- Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.
- Prince, Alan and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar*. Technical Report TR-2, Center for Cognitive Science, Rutgers University, New Brunswick, N.J. and Technical Report CU-CS-697-93, Department of Computer Science, University of Colorado, Boulder.
- Saffran, Jenny R. 2001. Words in a sea of sounds: The output of statistical learning. *Cognition* 81. 149–69.
- Smith, Neil V. 1973. *The acquisition of phonology: A case study*. Cambridge: Cambridge University Press.
- Tesar, Bruce and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.
- Tzakosta, Marina. 2004. *Multiple parallel grammars in the acquisition of stress in Greek L1*. Doctoral dissertation. Leiden University.
- Weijer, Jeroen van de. 2009. Optimality Theory and Exemplar Theory. *Phonological Studies* 12. 117–124.
- Weijer, Jeroen van de. 2012. *Grammar as selection: Combining Optimality Theory and Exemplar Theory*. Nagoya: Kougaku Shuppan.
- Weijer, Jeroen van de. 2014. The origin of OT constraints. *Lingua* 142. 66–75.
- Weijer, Jeroen van de. 2017. Where now with Optimality Theory? In Q. Ma (ed.) *Frontier research in phonetics and phonology*. Beijing: Foreign Language Teaching and Research Press.
- Weijer, Jeroen van de and Marjoleine Sloos. 2013. Learning markedness constraints: The case of French. In S. Aalberse and A. Auer (eds.) *Linguistics in the Netherlands 2013*. Amsterdam & Philadelphia: John Benjamins. 188–200.
- Weijer, Jeroen van de and Marina Tzakosta. 2017. The status of \*Complex in Greek. In T. Georgakopoulos (ed.) *Proceedings of the Twelfth International Conference on Greek Linguistics*. Berlin: ICGL.
- Wexler, Kenneth. 1990. Innateness and maturation in linguistic development. *Developmental Psychobiology* 23. 645–660.