Ungrammaticality in Language Variation and Acquisition

(submitted)

May 19, 2012

Abstract

Absolute ungrammaticality (sometimes called ineffability) is usually seen as a problem that is specific for Optimality Theory. Based on possibly surprising instances of ungrammaticality from language acquisition and dialect geography, this paper shows, first, that dealing with cases of absolute ungrammaticality is not easy for other phonological theories either, in spite of appearances. Secondly, there is evidence that children sometimes refuse to produce forms which they arguably know, showing the reality of ineffability in a domain very different from pardigm gaps. Thirdly, a suprising case of 'relative ungrammaticality' is discussed — some forms are more grammatical than others — which cannot easily be accounted for in other models of grammar, which tend to be 'all or nothing'. All of these point to a refinement of the notion of grammaticality which can be formalized in a version of OT that has similarities to OT syntax and semantics.

1 Grammaticality

One of the most important concepts of grammatical theory is *grammaticality* (sometimes called *grammaticalness*) and its inverse *ungrammaticality*.[1] A grammar for a language $\mathscr{L}$ should define which forms are grammatical and which forms are ungrammatical in $\mathscr{L}$. Inversely, classical generative grammar held that those forms are grammatical which can be generated by the grammar; all other forms are ungrammatical (Chomsky, 1965, p. 11).

Taking grammaticality as central implies that we consider grammaticality judgements to be a crucial empirical topic of inquiry, next to or even as opposed to presence and frequency in a corpus or a dictionary — which are indications of actual use and of relative probability —, although we can use the latter as indirect evidence for grammaticality. This position is sometimes contested, but I hope to show in this paper that it has its merits. I discuss

[1] Acknowledgements.

several 'real-life' phenomena which do not seem directly reducible to corpus frequency and for which we need a concept of grammaticality.

One of the advantages of 'classical', i.e. monostratal and parallel, Optimality Theory (OT Prince & Smolensky, 1993) is that it determines a very precise locus for ungrammaticality:[2]

(1) A surface form is ungrammatical in $\mathscr{L}$ because there is at least one competing form which satisfies the constraint hierarchy of $\mathscr{L}$ in a better way

I argue in this paper that OT is as a matter of fact a *superior* theory of grammaticality in phonology. One reason for this is that most other frameworks of phonology, such as Lexical Phonology or Government Phonology, are not very explicit on the locus of (un)grammaticality. In these theories, forms can be ungrammatical because they violate some hard constraint on the surface, or some morpheme structure constraint underlyingly, or because they would always be turned into something else by some phonological process. This is the source of a well-known problem: the duplication problem, a problem which OT avoids (Prince & Smolensky, 1993).

The explicit stance on grammaticality is sometimes taken as a disadvantage of OT by opponents. For instance, the constraint NOCROSSING, against line crossing, has been put forward as an example of a probably inviolable constraint, and this as an argument against OT (Scheer, 2004; Hale & Reiss, 2008).

However, the problem of the difference between hard constraints such as NOCROSSING and inviolable ones (or parameters) is not avoided by not making it explicit: what is the property of line crossing that makes it different from e.g. autosegmental one-to-many relations? The former are indeed universally inviolable, whereas the latter seem avoided only for some tiers in some languages, but not universally (i.e. they are subject to parametric variation). The question why this is the case is simply not solved by calling one a principle and the other a parameter.[3]

This is not to say that the notion of grammaticality is completely unproblematic for OT. A specific problem for this theory is that of *ineffability* or *absolute ungrammaticality*, the non-existence of certain forms (Féry & Fanselow, 2003; Orgun & Sprouse, 1999; Raffelsiefen, 2002; Rice, 2005; Rice & Blaho, 2009). Ineffability is a specific problem for OT, since that theory has the special property that an output is assigned to every input — something is always gen-

[2]This definition should be qualified, since in practice any serious OT analysis assumes restrictions on Gen (there are no analyses in which one candidate is a spectogram whereas another candidate is a Government Phonology representation), so that there is a different locus for (certain types of) universal ungrammaticality, viz. in Gen, than for language-specific ungrammaticality.

[3]The reason for the difference presumably is (Sagey, 1986; Bird & Klein, 1990; Coleman & Local, 1991) that a structure violating NOCROSSING is a logical inconsistency given some plausbile interpretation of association lines, whereas one-to-many relations are merely inconvenient for phonetic implementation. This insight can be built into any theory of universality and variation, but it is not easily formalizable in any of them.

erated given an input, no matter what that input is. The question then arises why certain paradims seem to have gaps: forms which one would expect to be there but which simply do not arise. At first sight, other grammatical theories do not suffer from this problem, since they can for instance have a crashing derivation, or 'hard' constraints acting as filters, banning certain combinations of words or morphemes to ever surface. The term *ineffability* seems to have arisen first within the OT literature (Pesetsky, 1998).

Based on possibly surprising instances of ungrammaticality from acquisition and dialectology, I will show, first, that dealing with cases of absolute ungrammaticality is not exactly trivial in theories outside of OT either, in spite of appearances. Secondly, there is evidence that children sometimes refuse to produce forms which they arguably know, showing the reality of ineffability in a domain very different from pardigm gaps. Thirdly, I show a suprising case of 'relative ungrammaticality' — some forms are more grammatical than others — which cannot easily be accounted for in other models of grammar, which tend to be 'all or nothing'. All of these point to a refinement of the notion of grammaticality which can be formalized in OT.

The issue becomes even more pressing if we consider that several recent theories of phonology do not seem to take the notion 'grammaticality' into account at all. For instance, in Exemplar Theory (Bybee, 2001), there is no grammar and hence no real concept of grammaticality different from, for instance, frequency or statistical probability. I will return to this in the last section of this paper.

2 Absolute ungrammaticality

2.1 The diminutive in Limburgian

Let us first look at a relatively well-known case of ineffability, viz. the interaction between umlaut and diminutive formation. In Limburgian dialects (West-Germanic, spoken in the Dutch and Belgian provinces of Limburg), diminutives are formed by adding a suffix *-k*ə (*-sk*ə after stems ending in velar obstruents) and turning the stressed full vowel of the stem into a front vowel:[4]

(2)

		Diminutive
v[uː]s	'fist'	v[yː]s-kə
b[oː]k	'book'	b[øː]k-skə
kl[uə]stər	'cloister'	kl[yə]stər-kə
kompj[uː]tər	'computer'	kompj[yː]tər-kə

The process even applies to loanwords such as *computer*, indicating its productivity. Furthermore, we observe that schwas are skipped. Umlaut is attracted

[4]These data are the dialect of Maasbracht; they reflect the judgements of Ben Hermans, a native speaker and phonologist.

to the stressed vowel. In the example in (3), the final (stressed) vowel is umlauted, but not the vowels preceding it.

(3)			
	kant[uə]r	'office'	kant[yə]r-ke
	kan[aː]l	'channel'	kan[ɛː]l-ke
	matr[ɑ]s	'mattress'	matr[æ]s-ke

Certain words lack a diminutive, viz. those words which end in an unstressed full vowel:

(4)		
	j[ˈuː]d[ɑ]s	'rotter'
	p[ˈu]m[a]	'puma'
	h[ˈoː]m[o]	'gay person'
	[ˈoː]m[a]	'grandma'
	[ˈoː]p[a]	'grandpa'
	k[ˈɑ]s[a]	'check-out point'
	[ˈoː]pər[a]	'opera'
	k[ˈaː]mər[a]	'camera'

A similar gap has been observed for German by a number of authors; e.g. Féry & Fanselow (2003) (the German diminutive pattern seems less productive in general, however). The issue is that any realisation of these nouns with the diminutive suffix *ke* is judged ungrammatical; it does not matter whether we just umlaut the last vowel (A), or both vowels (B), or only the stressed vowel (C), or no vowel at all (D).

(5)	A	B	C	D
	*j[uː]d[æ]s-kə	*j[yː]d[æ]s-kə	*j[yː]d[ɑ]s-kə	*j[uː]d[ɑ]s-kə
	*p[u]m[ɛ]-kə	*p[y]m[ɛ]-kə	*p[y]m[a]-kə	*p[u]m[a]-kə
	*h[oː]m[ø]-kə	*h[øː]m[ø]-kə	*h[øː]m[o]-kə	*h[oː]m[o]-kə

This is a classical ineffability problem: given the input /judɑs+kə/, we would expect one of these forms to be 'optimal', i.e. satisfying the relevant constraints better than the rest. It is a mystery, then, why this form does not surface as the winner.

Interestingly, absolute ungrammaticality also holds for words with underlying front vowels; as far as we are aware, this has not been noted before:

(6)				
	z[ˈeː]br[a]	'zebra'	*z[eː]br[ɛ]-kə	*z[eː]br[a]-kə
	T[ˈiː]n[a]	girl's name	*T[ˈiː]n[ɛ]-kə	*T[iː]n[a]-kə
	t[ˈy]b[a]	'tuba'	*t[y]b[ɛ]-kə	*t[y]b[a]-kə
	z[ˈeː]n[y]w	'nerve'	*z[eː]n[y]w-kə	

Because of these data, one might be tempted to think that the relevant restriction is independent of umlaut; that it it is a templatic restriction on the diminutive suffix itself (e.g. that it can only be attached to a 'true' trochee, and that somehow stressless full vowels are not part of a full trochee). However, geographically this type of ineffability is restricted to dialects which have umlaut in the Dutch language area. A search in the GTRP database of 612 dialects[5]

[5] http://www.meertens.knaw.nl/mand/

reveals no other dialect which has this restriction except from those in which there is active umlaut. There thus seems to be a strong correlation between umlaut and ineffability.

This is even true for speakers of this particular dialect, who have the option to switch to the Standard Dutch suffix *-jə/-tjə/-kjə/-pjə* (the affix has many allomorphs, which does not need to concern us here; see Trommelen (1982)), which doesn't trigger umlaut and also does not cause ineffability: *judasje, puma'tje, homo'tje*.

No matter what the actual source of ineffability is, these facts are clearly problematic for OT. The issue is whether they are substantially easier to solve in other theories. The question if of course difficult to answer because there are so many versions of so many different theories of phonology around that it is hard to not overlook at least one of them — perhaps the one with the definitive answer. Yet broadly speaking, we can go in two directions: we can either try to solve the problem derivationally, or we can try to do it representationally. However, it seems that both roads lead to difficulties.

First, we cannot derive these facts purely derivationally. Although there are not many purely derivational phonological theories, it might be worthwile to point out that the solution cannot be found in derivations. Purely derivational theories such as Chomsky & Halle (1968) share with OT the property that given an input we should always be able to derive some output: if we insert a noun and a suffix, some rules apply, and there is no reason why in these cases the rules would be so destructive that they would erase everything. In other words, the umlauting rule should be blocked in case of an intervening full vowel, but it is not clear why and how this blocking would lead to zero.[6]

We can alternatively posit a 'hard' well-formedness constraint (for instance, a subcategorisation frame) on the Limburgian suffix:

(7) The suffix attaches to a 'perfect' trochee, i.e. a monosyllabic foot or a foot ending in schwa

The problem is, as we pointed out, that this requirement goes together with the umlauting behaviour and that connection is not made in this way. Obviously, we could try to build the umlauting into our constraint:

(8) The umlaut factor (a floating feature) needs to be linked on the surface to a stressed vowel, while it cannot cross full vowels. Also, we cannot delink the features of those intermediate full vowels (while we can delink the features of the stressed full vowel).

This constraint however mixes derivational and representational views in a way which can probably be considered unacceptable to most analysts. Again,

[6] One idea would be to implement the idea developed in this paper in a derivational framework, allowing morphological parsing rules to sometimes be blocked or alternatively specify a form as [minlexical insertion] (Halle, 1973). The problem in developing such an implementation will be that such blocking will have to have the property of global look ahead in the derivation: a form should not be parsed by a morphological rule if this will lead to phonological problems later on.

it is impossible to prove that other theories cannot make the link between umlauting and absolute ungrammaticality, but at least this does not seem so straightforward that cases of ineffability such as this one should make us decide against theories of violable constraints.

2.2 Dealing with ineffability within OT

With an analysis of the phenomenon in other frameworks pending, let us see how we can deal with it within OT. van Oostendorp (2009) argues that there are broadly speaking four different lines of treating ineffability in the OT literature.

1. *The 'paradigmatic' solution*: the Generator function does not generate an individual form, but a paradigm. Ineffability of an individual form means that this particular form is not generated within the paradigm (Raffelsiefen, 2002; Rice, 2005). This is problematic when there is no real inflectional paradigm (for instance in the case at hand, it is not clear that diminutives should always be generated in a morphological paradigm together with the underived nouns), but even if we were to include the diminutive inside a paradigm it would face the conceptual problem that we will need a derivation of at least two steps (the first generating the paradigm, the second selecting the relevant form and using it in a sentence of utterance), and it is unclear how the second step would in turn be blocked from e.g. choosing a different form in the paradigm (and even whether it is implementable in OT at all).
2. *The 'null parse' solution*: the Generator function generates a candidate in the phonology which does not have a phonetic interpretation, and this is selected as the winner in certain cases (Prince & Smolensky, 1993; McCarthy & Wolf, 2009) (see Ackema & Neeleman (1998) for a possible implementation in syntax). This is the classical solution which so far did not find a satisfactory formalisation. The problem is that in order for the 'null parse' to win (rather than some form which leaves only the offensive part of the input unparsed, say in our case, the unstressed full vowel), it needs to have a representation which is quite different from that of other candidates. It has turned out difficult to find independent justification for such a representation.
3. *The 'control' solution*: the Generator and Evaluator function conspire to create a (pronounceable) candidate, but a grammatical component outside of the standard OT system (called 'Control') subsequently blocks this candidate (Orgun & Sprouse, 1999). In our case, e.g. the winning candidate could be one with umlaut spreading through the unstressed vowel, and the Control component then would make this structure crash. The 'control' solution to the problem has not received a lot of support in the literature — it is not clear why the language faculty would have a design with two different constraint components —, and we will not pursue it here. (Note that in this particular case one could

also wonder whether it would be possible to avoid the problem with the constraint in (8) above.)

4. *The 'divergent meaning' solution*: we generate a phonologically well-formed form, but one which does not have the intended semantics; for instance, in this case, the form is not a diminutive at all, but just the underived form of the noun. This solution is standard in OT syntax/semantics (Keer & Bakovic (2001); Vogel (2001); Wilson (2001); Legendre (2009); see Müller (To appear) for an overview), and has been defended in van Oostendorp (2009) for phonology; we will follow this solution here.

In order for a divergent meaning solution to work, we need to assume that the input to the phonology does not consist of sets of phonological representations, as is standardly done. Rather, the input of phonology is an unstructured set of abstract morphemes, i.e. bundles of morphosyntactic and possibly semantic features without any phonological information, and lexical insertion — finding the right phonological objects corresponding to these features — is the job of the phonological Generator function (Wolf, 2008). The output consists of a two-dimensional structure, consisting of a morphological word and a phonological word. The former will be subject to semantic interpretation and the latter to phonetic implementation.

These assumptions have significant implications for our theory of faithfulness. In the first place, we need to make sure that the input is preserved in a slightly different way from how this happens in theories in which the input is a phonological object itself. Instead of this, we need to make sure that the underlying morphosyntactic information is preserved, in other words that all morphemes will be represented. We need a constraint (M)PARSE, requiring that individual morphemes are part of the morphological structure, to take care of this:

(9) MPARSE(M): Every morpheme M has to be parsed into a morphological word.

This is (possibly) slightly different from the general MPARSE of Prince & Smolensky (1993, 52), in which the whole input is left morphologically unparsed. We cannot be completely sure, because Prince & Smolensky do not provide a formal definition of their constraint — (10) is what comes closest:

(10) MPARSE (Prince & Smolensky (1993), p.52): "the constraint which requires the structural realization of morphological properties".

The sample analyses concern hypothetical monomorphemic forms (Latin **rĕ*) which are not integrated into the morphological structure for phonological reasons (in this case, because words have to be minimally bimoraic):

> Failure to achieve morphological parsing is fatal. An unparsed item has no morphological category, and cannot be interpreted, either semantically or in terms of higher morphological structure. (Prince & Smolensky (1993), p.52)

Prince & Smolensky do not discuss the case of an input which consists of more than one morpheme: does the whole form become unparsed, or merely the 'offensive' morpheme — and this is where the slightly higher precision of the new definition in (9) becomes relevant: it can be violated by *every* morpheme in the input individually if it isn't parsed. That means that it will be more profitable to still parse as many morphemes as possible into the overall structure.

When we are trying to form an illegal diminutive, it will thus be better to leave only the diminutive suffix unparsed than to not pronounce anything at all. A non-phonological example may serve to illustrate this point. No variety of Dutch allows diminutives of prepositions, as far as I am aware; yet nothing excludes such a sequence from being posited underlyingly. For this reason, we introduce the *ad hoc* constraint in (11):

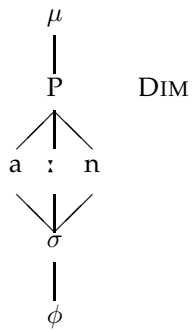
(11) LEXDIM: There is no diminutive of function words.

(12)

a:n 'to' Dim	LEXDIM	MPARSE(DIM)
☞ a:n $< +Dim >$		1
a. ∼ a:ntjə	$_1$ W	$_0$ L

The winner in this tableau should be imagined to look as follows, where μ abbreviates the full morphological (or morphosyntactic) tree, and ϕ the phonological structure:

(13) *Non-parsing of diminutive morpheme*



In this structure the diminutive suffix is not part of the morphosyntactic tree at all; that is why it violates MPARSE(DIM). Phonological faithfulness constraints on the other hand are not violated: phonological material has never been retrieved from the lexicon, so that we cannot be unfaithful to it.

The form *aantje* presumably does not exist for purely morphosyntactic reasons; but we can apply the same line of thought to the case at hand. In this case, we could say that we are not parsing the morphology for phonological

reasons. This would create the following table (λx : JUDAS(x) is a stand-in for the semantic and morphosyntactic properties of the name *Judas*):[7]

(14)

λx : judas(x), dim	PHON	UMLAUT	MPARSE(DIM)
☞ juːdɑs $< +dim >$			1
a. ∼ judɑskə		$_1$ W	$_0$ L
b. ∼ jydɛskə	$_1$ W		$_0$ L
c. ∼ judɛskə		$_1$ W	$_0$ L
d. ∼ jydɑskə	$_1$ W		$_0$ L

In this table, it is presupposed that the (morphological) constraint UMLAUT is violated whenever the stressed vowel is not hit by the umlaut factor on the suffix, whereas the cover constraint PHON is violated by either skipping the unstressed vowel or by umlauting it as well. The winning form in this table thus is *not* a diminutive and will not be treated as such by any outside module, such as the semantics or gender agreement. When a Limburg speaker tries to make a diminutive of *Judas* with the suffix *-ke*, he just produces something else: a non-diminutive. The option chosen here can thus be characterized as extreme neutralisation.

As we have mentioned above, it is also possible to use the non-umlauting, Standard Dutch, suffix *-je* in these contexts. This raises the question whether we should not include the form *judasje* in the table in (14): it would probably satisfy all relevant constraints and win. Although in many cases of ineffability one could rise this question — is English **intelligenter* a matter of ineffability or is it blocked by *more intelligent*? — we believe that in this case *-ke* and *-je* are two independent morphemes, for instance because in words with a different structure both are possible, so that they do not seem to compete in the same tableau.

2.3 Theoretical implications

What are the consequences of this view of the ineffability problem? In the first place, we have to crucially assume that the input consists of morphological and possibly semantic features, rather than of phonological strings. We assume, in other words, that lexical insertion is the job of the phonological Generator function. Lexical insertion and morphological evaluation work in

[7] Throughout this paper I use a tableau type introduced into the literature by Prince (2002) and known as a comparative tableau. On te furst line of such a tableau we find the winning candidate with the number of violations for each constraint indicated. Below this, the most important alternative candidates are presented, each also with a number of violations, as well as W and L markings. A cell gets a W marking if the actual winner fares better on this constraint than the candidate under consideration; it gets L if the latter would win in a direct competition with the winner on this constraint. In a well-formed tableau, every L should be preceded by an L.

parallel with the phonological evaluation. (Wolf, 2008, arrives at similar conclusions following a very different line of thought, viz. the study of allomorphy selection).[8]

This changes the implementation of the Richness of the Base principle of OT somewhat, but not (necessarily) its spirit. The principle now makes two more or less independent claims. First, it states that any bunch of morphemes can be underlying to the phonology (but the phonology may choose not to apply lexical insertion and the morphology may choose not to parse all morphemes). Secondly, there are no 'inherent' restrictions on the phonological structure of a lexical entry. We still have to account in our grammar for what happens to, say, a lexical entry /Θɜɸ/ in a language like English in which neither of these segments is allowed to surface.

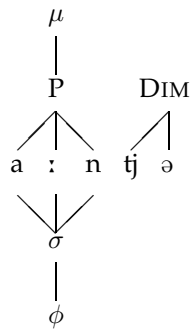
Also our view of phonological faithfulness should obviously change. The key idea is expressed in (15):

(15) Phonological faithfulness constraints relate the lexical representations of morphologically parsed material to their output counterpart.

In order to get a coherent linguistic representation, the 'dimensions', phonology and syntax, should be parallel at least to a large extent. A reasonable way of saying this is that phonological material should always be licensed by being part of the morphosyntax; if it is not, it should be counted as epenthetic (and therefore licensed by purely phonological considerations).

For instance, one candidate competing with (13) could be (16)[9]:

(16)



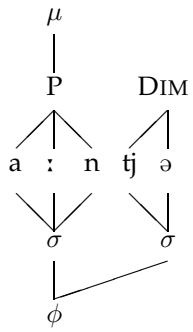
This candidate will violate MPARSE(DIM) in the same way as (13) (there is a morpheme which is not integrated into the morphosyntactic tree), but at the same time it will contain unparsed phonological material; it will therefore violate any constraints requiring phonological material to be integrated into

[8] See Kager (1996); Yip (1998); Zuraw (2000); Boersma (2001); Escudero (2005) for related ideas. In particular the ideas in Walker & Feng (2004) come close in the technical way in which certain analyses work, although they are implemented in Correspondence Theory.

[9] I put *tj* under one node since it is sometimes analysed as one segment. This is not directly relevant to the current analysis.

the phonological structure. If we do integrate the phonological material into the structure, however, we end up with the following representation:

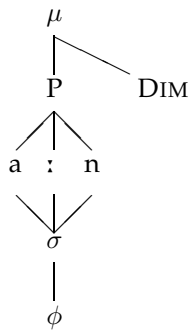
(17)



This representation still violates MPARSE(DIM), and furthermore it violates phonological faithfulness, since it contains segments which do not correspond to anything in the morphosyntactic tree — thus violating faithfulness constraints which have been formulated along the lines of (15). In other words, as far as the phonology is concerned, the string [tjə] counts as epenthetic. Since nothing in the phonology forces these segments to arise, they are considered illegal.

Finally, we should consider representations such as the following:

(18)



In this case, the diminutive suffix is parsed into the morphological structure, but lexical insertion has not taken place. The structure would violate our constraint against diminutive prepositions LEXDIM. What would happen if we would replace the preposition *aan* with the noun *maan*? It is obviously possible to make a diminutive, but the structure in (18) would still violate a constraint which enforces the phonological realization of morphemes, REALIZEMORPHEME (Kurisu, 2001; van Oostendorp, in press).

(19) REALIZEMORPHEME: morphological nodes should have an equivalent in the phonological representation.

Since REALIZEMORPHEME is an ordinary OT constraint, it can sometimes be violated so that we create structures such as (18). The result of this would for instance be zero allomorphs.

Faithfulness relations are thus threefold under this analysis:

(20) a. MPARSE constraints require every underlying morpheme to be part of the morphosyntactic tree.
b. REALIZEMORPHEME requires every part of the morphosyntactic tree to get phonological expression.
c. Phonological faithfulness constraints (e.g. MAX and DEP) require every phonological element (features, segments, etc.) to be organized in the phonological tree.

MPARSE constraints organize the morphosyntactic tree in the same way that MAX and DEP organize the phonological tree; REALIZEMORPHEME is responsible for mapping one tree on the other.

3 Selection in L1 acquisition

3.1 Avoidance of words

We have thus seen that classical ineffability can probably be solved under OT without abandoning its most standard principles, and furthermore that it is not always easily solved in other frameworks. We will now move to a different type of ineffability which to my knowledge has not received a lot of discussion in the literature so far: ineffability in child language acquisition.

We find reports in the literature (e.g. Leopold, 1949; Ferguson & Farwell, 1975; Kiparsky & Menn, 1977; Leonard *et al.*, 1981; Schwartz & Leonard, 1982; Schwartz *et al.*, 1986; Vihman, 1993; Yavas, 1995; Storkel, 2006) that children at a very young age (before having learnt more than 50 words) systematically avoid producing words which contain sounds they are not yet able to make.

For instance, in an early study, Leopold (1949) showed that his daughter Hildegard was only attempting oral and nasal plosives at the labial and apical place, and no other sounds. She did not produce all the sounds she was attempting correctly, so her avoidance of the other sounds could not possibly have been motivated by purely articulatory considerations. The child would try /n/, even though she was not successful, but she did not even attempt to produce /ŋ/.

Later experimental results confirm this finding. In an influential study, Schwartz & Leonard (1982) presented children with objects (for instance, a rubber bulb from a bike horn) that were named by nonsense words (for instance, *eshesh*). For every child, some of these words were in an 'IN' group: they consisted of phonemes that were present in words which the child already knew, and arranged in syllable structures that were also already demonstrably present in the child's phonology. For other children, they belonged to

an 'OUT' group, for instance because they consisted of sounds which did not yet occur in the child's vocabulary.

It was then observed whether and how children would repeat those words, immediately after they were named by the experimenter. The table in (21) repeats some of the examples of Schwartz & Leonard (1982)'s article (imitative productions were made in direct response to the experimenter, non-imitative productions after one of the particpants had produced some more utterances):

(21)

Target	Imitative	Non-imitative
bɛ (I)	mɛ*	mɛ*
oʃoʃ (O)	ozɪ*	NP
heb (I)	je*	hæ*
afaf (O)	aja.*	a:*
momo (I)	momo	mom
zivi (O)	NP	zi

I = IN for the child whose productions are presented. O = OUT for the child whose productions are presented. NP = not produced. * = Production scored as inaccurate.

In some cases, a word which as in 'IN'-group was in the 'OUT'-group for another child; in such cases, the experimenters always chose the grouping according to the child trying to produce the form in question. Schwartz & Leonard (1982), p. 329-330 conclude that their findings

> provide experimental confirmation of the observations that young children are selective in words which they attempt. Words which were consistent with the limited phonologies of the children studied were more likely to be acquired than words with phonological characteristics considered out of their phonologies. The same constraints appeared to govern the children's imitations of the experimental words. A greater number of words with characteristics in their phonologies were imitated than words with characteristics that were out of their phonologies.

Later studies have basically confirmed these results, at least for very young children: Storkel (2006) shows that older children, for whom most of the sounds are in the IN group, might actually pay more attention to OUT sounds.

In the quote just given, it seems that two different views of the phenomenon are entertained: it is first claimed that the children do not 'attempt' these words, probably meaning that they do not attempt to *produce* them. In the next sentence, on the other hand, it is claimed that the children did not *acquire* the OUT words.

Although it may seem difficult to distinguish empirically between a baby having acquired a word without producing it, and one not having acquired the word at all, it seems to me that the relevant evidence points in the direction of the former. In the first place, it sometimes happens that children did not use the form at first (in the imitative context), but then did use it later.

This indicates that the children must have known the words even when they didn't say them. In the second place, Schwartz & Leonard (1982) themselves point out in this quotation that direct imitation — in which presumably no acquisition is involved — is subjected to the same constraint. We therefore need to describe the situation anyway in which children do 'know' the form in some sense, but at the same time are not able to produce it.

In the following section, I will show that with careful study, we can also detect this pattern in a more abstract form in a corpus of spontaneous data. The point is going to be here that children do not just 'avoid' words with a single segment, but words with natural classes of segments. Since there is no reason to assume that their input will just miss words with such natural classes of sounds, the only reasonable conclusion we can draw from this is that the natural classes are in their learning behaviour. In section (3.3) I will give an ineffability analysis of this.

3.2 Evidence from spontaneous data

The following data are from the CLPF database, with data from 12 Dutch children aged between 1;0 and 2;0 at the start of a one-year period of data collection. Recordings were made (on average) bimonthly, at the children's homes; these recordings took between 30 and 45 minutes. The data were collected by Paula Fikkert and Clara Levelt, originally for their respective doctoral dissertations (Fikkert, 1994; Levelt, 1994)

Levelt & van Oostendorp (2007) use these data to demonstrate that we can describe the acquisition of segment inventories as feature-based if we also adapt a theory of very simple feature cooccurrence constraints. We can concentrate here on an aspect of Levelt & van Oostendorp's data that are not explicitly discussed in that paper: the fact that targeting a sound systematically precedes producing the same sound, even when seen from the point of view of features. We will concentrate on one child — Child 2 in Levelt & van Oostendorp (2007) — and refer to Levelt & van Oostendorp (2007) for full analysis of the other children.

(22) Child 2

P:	bdthp	m	s			fn		kx	l
T:	bdtmn*p*		s	zhk	x	fvr	l		
D:	529	540	554	588	602	615	629	643	766

Target>production: m, n, z, k, x, l, j
Production>target: h (p)

P: the segments which are succesfully (faithfully) produced by the child; T: the segments which are targeted by the child (i.e. which occur in adult pronunciations of the words which the child says); D: day in the life of the child.

For every day, we only note the *new* sounds in the inventory of the child; thus, at day 554 the child targets the sounds *b, d, t, m, n, p* and *s* of which she

succesfully produces *b, d, t, h, p, m* and *s*. Further, we only noted a sound in an inventory when it was produced at least twice at the stage which we investigated.

It becomes clear from this table that the child succesfully produces sounds for the first time either in the first recording in which the sound is targeted, or in some later recording. The exception is *h* which is sometimes produced in words which do not have that sound in the adult pronunciation before the child even attempts to say words with a real *h*. This may be an artefact of the transcription: a phonetically preaspirated vowel may have been transcribed in this way, for instance.[10]

The fact that the children first try certain sounds before they actually produce them correctly, may at first sight not seem very surprising. The interesting thing is that the avoidance systematically involves every word in which those sounds occur. The child would always have the option of producing the word without the sound in question — or of changing the sound just like she does in the later period.

Even more surprising is the fact that Levelt & van Oostendorp (2007) show that the produced inventories and the targeted inventories can be described in grammatical terms. They follow Jakobson (1942) in assuming that children do not acquire individual sounds, but distinctive features, plus a grammar which tells them how these features may be combined.

Levelt & van Oostendorp (2007) assume that such grammars can only consist of a limited number of feature cooccurrence constraints (*fcc*), viz. either *[F,G] (disallowing the combination of privative features F and G) or [F]⊃[G](requiring privative F to be accompanied by privative G). The claim is that more complicated constraints (e.g. regulating the cooccurrence of three or more features) are never necessary.

All segments that can logically be constructed on the basis of an acquired feature set, will be present, except when they are explicitly disallowed by a *fcc*. Furthermore a *fcc* linking [F] and [G] can only take effect at the moment when [F] and [G] have both become available, not later, although they can be retracted at a later stage.

Such a theory is quite restrictive, as can be seen when we study some of consonant inventories which are excluded by it:

(23) a. { t, d, p, b, s, z, f }
b. { t, p, z, f }

The inventory in (23b) lacks a voiced labial fricative, even though the two-way combinations of voicing, labiality and continuancy are all present; such an inventory cannot be described without a constraint disallowing the combination

[10] One further discrepancy is that the *p* was targeted only once during the recording session on day 529, even though it was produced more often in that same recording session (replacing other sounds). Strictly speaking, *p* should therefore appear under day 540, and this would be another instance of a sound being targeted only after having been produced.

of three features (*voi,lab,con). The inventory in (23b), on the other hand, has only one voiced sound, which is the coronal fricative; neither coronality on its own nor continuancy on its own induces the voicing; again we would need a constraint referring to at least three features (e.g. [cor,cont]⊃[voi]) to be able to properly describe this phenomenon.

I refer to Levelt & van Oostendorp (2007) for full analysis, but for the two children under study here the analysis would run as follows. First, Child 2 would acquire her sounds in the following order (we will disregard the *h* because of the analytical problems described above):

(24) Child 2: Acquisition of features

P:	lab cor voi	nas	cont					vel	lat
T:	lab cor voi nas		cont	vel			lat		
D:	529	540	554	588	602	615	629	643	766

Furthermore, the child has to activate the following *fcc*, for production and targeting:

(25) Child 2: Acquisition of *fcc*

	P:	T:
529		
540	(i) [nas]⊃[lab]	
554	(ii) *[cont,voice]	(ii) *[cont,voice]
	(iii) [cont]⊃[cor]	(iii) [cont]⊃[cor]
588		retract (ii)
602		retract (iii)
		(iv) *[lab,cont]
615	retract(i), retract(iii)	retract (iv)

The target inventories of segments can thus also be described in these neo-Jakobsonian terms, viz. as the development of a grammar. This grammar is independent of the sounds which can be produced at any given stage, even though those can be described within the same grammatical framework. I take this to be a strong indication that they are not random. The child seems to select the words it tries to produce according to the features and feature combinations which are available at this time.

Levelt & van Oostendorp (2007) show that the same analysis can be made for the other children in the corpus for which there is enough relevant data available. In all cases, we can describe both the production and the targeting in grammatical terms. This shows that avoidance of certain words *is* a

grammatical process. That in turn means that either the child only learns 'grammatical' words or she learns all words but only attempts to produce the grammatical ones.

3.3 An analysis of acquisition

There are thus three stages in the acquisition of a sound. First, the child knows words with the sound, but she avoids producing them. After this, the child attempts producing words with the sound, but fails to do so, and either deletes the sound or replaces it by something else. And finally, the child produces the sound correctly. Given the rather coarse nature of our data, we will of course not always find attestations for each three of these stages.

Let us take the (hypothetical) example of a child which first refuses to say a (hypothetical) word *sa*, then says *ta* (replacing the fricative by a corresponding plosive) and finally correctly produces *sa*. (The example is hypothetical only in that we do not look at the vowel.)

For the sake of the argument, I will assume that the child knows exactly the right underlying form in each stage. Given my analysis so far, there will be three relevant constraints:

(26) a. IDENT-[con] (Faithfulness: do not delete underlying [con])
b. *[con, cor] (Markedness)
c. MPARSE-SA

I adopt the fairly standard assumption that M≫F at the initial stage of acquisition (Tesar & Smolensky, 2000): acquisition is a matter of learning to demote markedness constraints below faithfulness constraints when they are violated in the target language.

As an extension to this, I assume that at the initial stage also F≫MPARSE; in other words, it is considered more important to try to speak a word well than it is to express the concept in the first place: a child will only start trying to pronounce a concept when it has enough evidence that it is necessary to say it.[11]

The three most important candidates get the following violation profiles for the constraints under discussion:

[11] Albright (2009) points out that there is a problem with the assumption that MPARSE is initially ranked low, which this approach shares with McCarthy & Wolf (2009). The problem is that this ranking predicts that newly acquired will never be allowed to undergo an alternation, unless there is any evidence for it. This will preclude using *wug* tests as a methodology. This is definitely true: one predicts that a speaker will always initially refuse to pronounce new words. We assume, however, that there will be sufficient pressure outside to system to produce the word as fast as possible and therefore to rerank MPARSE.

(27)

$\lambda x : \mathrm{sa}(x)$	*[con, cor]	IDENT-[con]	MPARSE-SA
a. $\emptyset$			*
b. sa	*		
c. ta		*	

Our assumptions give the initial ranking *[con, cor]≫IDENT-[con]≫MPARSE-SA, which will lead the child to not pronounce the word at all.

The moment she starts to realize that she has to produce the morpheme SA, she will have to rerank MP to a higher position. Reranking the two lowest constraints (the minimal reranking possible, given the requirements) will do the job; it will lead to *[con, cor]≫MPARSE-SA≫IDENT-[con], which will designate [ta] as the winner.

The next, and final step, will then be to observe that this is still not the desired output. In order to get there, we need another reranking. A reranking of the two constraints at the bottom will not lead to the desired result — as a matter of fact it will lead back to the initial stage of ineffability — so the child may try a reranking of the two constraints at the top of the hierarchy. This will lead to a ranking MPARSE-SA≫*[con, cor]≫IDENT-[con], which however still gives [ta] as the outcome. Because this is still not correct, the child reranks the bottom two constraints (again a minimal reranking) which leads to the right and final result, [sa].

Schematically, we can thus write down the development as follows (M=markedness, F=faithfulness, MP=morphological parsing):

(28)

First stage:	M≫F≫MP:	$\emptyset$
Second stage:	M≫MP ≫F:	ta
Intermediate stage:	MP≫M ≫F:	ta
Third stage:	MP ≫F ≫M:	sa

This seems a plausible scenario within OT, given our account of ineffability: ineffability is a stage in the acquisition of any morpheme — there will always be a resistance to pronouncing abstract morphemes to begin with. Note that it is not so clear how this avoidance can be implemented in other frameworks: why would it be the case that the existence of a bad segment in a word leads to a total crash of the derivation for children? Those other frameworks will thus here be confronted with the problem that is usually thought to be unique to OT. It may be a challenge to solve this problem which the proponents of those theories might find worth considering.

4 Relative judgements

Our final point will concern the issue of relative judgements: constructions which are considered to be not completely grammatical, but also not completely ungrammatical. Although this issue has always played a role in the

background in grammatical theory (e.g. Chomsky, 1955, Chapter 4), and many papers are full of relative grammaticality judgements such as '??', '?*' and '**', it is not clear how we should interpret these markings in actual practice. Clearly, an all or nothing theory of grammaticality is much too simplistic. At the same time, most existing theories are of an all-or-nothing nature.

Notice that this type of variation has been the topic of rather extensive theoretical discussion in the last decade (see Boersma, 1998; Anttila, 2002; Coetzee, 2009; Coetzee & Pater, 2011, and references provided there). One important characteristic of such work is that it aims to show that classical models of categorical grammaticality should be replaced by models which incorporate stochastic or other kinds of 'performance' information. My contribution to this debate is twofold. First, I aim to show that such 'gradient' data within a speaker can have a clear correlate in a speaker community as an implicational relation on a geographic map. I believe that this shows something new: that speakers may have the same kind of hierarchy of grammatical forms, but use a different cut-off point on whether or not they pronounce certain forms. Secondly, I show that such a combination of data can be succesfully accounted for in the model developed here.

In this section, we will show how the OT formalisation of grammaticality can shed light on this issue. The burden is, again, on prononents of alternative views of grammaticality to explain how the issue would work under those alternatives.

4.1 Two types of pronouns in Dutch

We will once more take our empirical material from Dutch. It is a well-established fact that (Standard) Dutch has two series of pronouns, so called 'strong' and 'weak' pronouns (Berendsen, 1986). The latter tend to be phonologically encliticized, whereas the former can bear independent stress:

(29)

strong	weak	
ɪk	ək	I
jɛi	jə	you (singular / informal)
y	y	you (singular / polite)
hɛi	di/ti/i	he
zɛi	zə	she
hɛt	ət	it
wɛi	wə	we
jœli	jə	you (plural)
zɛi	zə	they

At first sight, it may look as if the weak pronouns can be phonologically derived from the strong ones, but Berendsen (1986) has convincingly shown that this is not the case. In the first place, there is no general rule reducing diphthongs to schwa in Dutch, which would be necessary for many of these cases;

in some other cases there is no clear reduction relation at all (in particular this is true for /hɛi/–/di/). In the second place, some idioms can only have reduced pronouns and not weak ones:

(30) *daar kun jə/*jɛi donder op zeggen*
'you can bet your bottom dollar'

(31) *laat mə/*mɛi niet lachen*
'don't make me laugh'

Thirdly, weak pronouns do not always have the same meaning as strong pronouns: *jə* can have a general meaning ('people') which is not shared by *jɛi* (32), and zɛi usually only refers to people, but zə can also refer to inanimate objects (33):

(32)

jə wordt moe	jɛi wordt moe
(i) you get tired	(i) you get tired
(ii) people get tired	(ii) *people get tired

(33) Kees zegt dat de fietsen daar nog staan, maar zə/*zɛi zijn weg
K. claims that the bikes are still there, but they have gone

Hardly discussed in the literature, however, is the fact that next to the weak and strong pronouns mentioned before, there is also a third type, consisting of a consonant only:

(34)

k	I
j	you (singular / informal)
	you (singular / polite)
	he
z	she
t	it
w	we
j	you (plural)
z	they

There are two potential reasons why these are not discussed. The first is that they are quite informal, and many of them may be argued to not even be part of the standard language. The second is that in this case, phonology may play a role: the pronouns in (34) are equal those in (29), minus the schwa — that is then the reason why we do not find anything for the *vous* form of the second person singular or for the third person plural: those consist of only a vowel in the weak form to begin with.

I know exactly two studies which do discuss this phenomenon: de Vries (2001) and Noske (2005).

de Vries (2001) (a speaker of a northern variety of Dutch, something which will become relevant later) observes that single *k* occurs at the beginning of

a sentence, and that it can occur before any vowel or consonant. He cites examples such as the following:

(35) a. *'k Interesseer me daar niet meer voor* 'It doesn't interest me any more'
b. *'k Heb trek in iets hartigs* 'I am craving something salty'
c. *'k Wil er best heen* 'I wouldn't mind going there'
d. *'k Speel niet meer mee* 'I am no longer play along'

Furthermore, de Vries (2001) notes that there is voicing assimilation in examples such as: /ək vul/ 'I feel' → [kful]; /ək ben/ 'I am' → [gben] (the direction of assimilation is like that of voicing assimilation in Dutch more generally). He also claims that the vowelless form is impossible before a *k*-initial verb: *[kkɛik]. Similar observations hold for *t*, which is impossible before *t* and *d*, for instance.

Noske (2005), on the other hand, is a very short paper based on informal observation of differences between Flemish and Dutch speakers. Noske (2005) claims that "the neuter pronoun *het* [ət], e.g. can be cliticized and can loose its schwa more easily and in more positions in Southern Dutch [i.e. the varieties of Dutch spoken in Belgium] than in Northern Dutch [the varieties spoken in the Netherlands]. In both Northern and Southern Dutch, schwa can be deleted if it is followed by a vowel initial inflected verb, if the main sentential stress is not located on one of the words." Noske provides the following example:

(36) het is [tɪs]

Noske (2005) notes that a "more marked contrast between the North and the South (...) can be found if *het* is in a position following a tensed verb, e.g. in *was het* 'was it':

(37) *was het* Northern Dutch: [ʋasət]-[ʋazət] *[ʋast];

(38) *was het* Southern Dutch: [wast];

And finally, "also other unstressed pronouns, like *ik* 'I' (...) can loose their vowels much more easily in West-Flemish and other Southern variants than in Standard Northern Dutch (...)":

(39) ik hoor /ik hor/ West-Flemish: [kor], Northern Dutch [ɪk hor]

The difference between Northern and Southern Dutch according to Noske is one of the relative ranking of a ONSET requiring syllables to start with a consonant and ALIGN, disallowing syllables to cross word boundaries:

(40) a. Northern:

ɪk oːr	ALIGN	ONSET
(.ɪk.)(.oːr.)		2
a. ∼ (.koːr.)	2 W	L

b. Southern:

ɪk oːr	ONSET	ALIGN
(.koːr.)		2
a. ∼ (.ɪk.)(.oːr.)	$_2$ W	L

Noske (2005) observes that this difference may be related to influence of French on Southern Dutch, since according to van Oostendorp (2000a), French phonology allows processes to cross word boundaries much more freely. A further ranking argument for this is that we need a similar ranking for differences in compounding behaviour (*oneens* means 'in disagreement'):

(41) a. Northern:

ɔn+ens	ALIGN	ONSET
(.ɔn.)(.ens.)		2
a. ∼ (.ɔ.nens.)	$_1$ W	$_1$ L

b. Southern:

ɔn+ens	ONSET	ALIGN
(.ɔ.nens.)	1	1
a. ∼ (.ɔn.)(.ens.)	$_2$ W	L

Noske (2005)'s analysis is not without problems. For instance, it is not clear why ALIGN would be responsible for the difference between Northern [ʋasət]-[ʋazət] and Southern *[wast], since in both cases the boundary between the two words is blurred.

4.2 Relative judgement and geographical pattern

It is interesting to observe that a relative judgement from a speaker of the northern parts of the Netherlands (de Vries) is attributed to a geographical pattern by another author (Noske).

It is actually not difficult to replicate these findings. In a small online experiment, we asked 47 native speakers of Dutch to pairwise rank the following word combinations *'k heb* 'I have', *'k hoor* 'I hear', *'k zal* 'I will', *'k zeg* 'I say', *ik ben* 'I am' and *'k beloofde* 'I believed'. These six forms can be paired in 15 different ways; the participants had to listen to these forms, spoken by a speaker of Standard Dutch of the Netherlands, and say which of the two they found more acceptable (this was a forced-choice task).

It turned out that there was exactly one ranking which fitted the judgements almost perfectly:

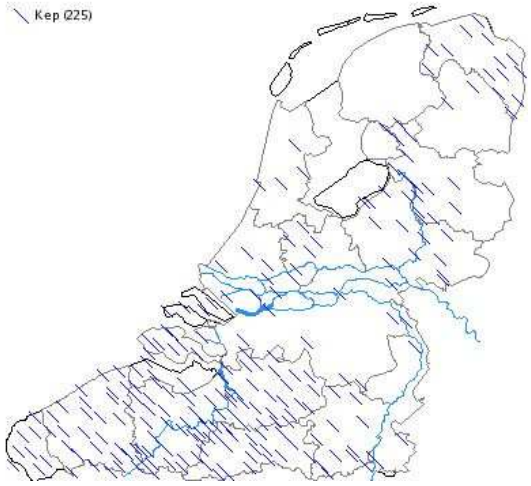
(42) kɛp 'I have' ≻
kor 'I hear' ≻
ksɑl 'I will' ≻
ksɛx 'I say' ≻
gbɛn 'I am' ≻
gbəlofdə 'I believed'

In this example, A $\succ$ B means that A was judged better than B more often. The relation can be considered transitive. Altogether then, (42) represents 15 $\succ$ relations; in all cases there was a majority of speakers who considered that particular relation to hold rather than the opposite (there were never more than 12 speakers who held an opposite relation, furthermore, no speaker had ever more than 3 relations reversed with respect to the standard).

Two parameters seem at stake: first, the pairwise distinction between *have-hear*, *will-say*, *am-believe* can be understood as the difference between auxiliaries and lexical verbs. Secondly the difference between [kɛp, ksɑl, gbɛn] and [kor, ksɛx, gbələfdə] seems phonotactic in nature: a simple onset is better than an onset consisting of a plosive followed by a fricative which in turn is better than an onset made up of two plosives.[12]

Interestingly, most of these sequences can also be found in the GTRP database. Here, for instance, is a map of the form *ik heb* 'I have'.

(43) *Ik heb* = I have (AUX)

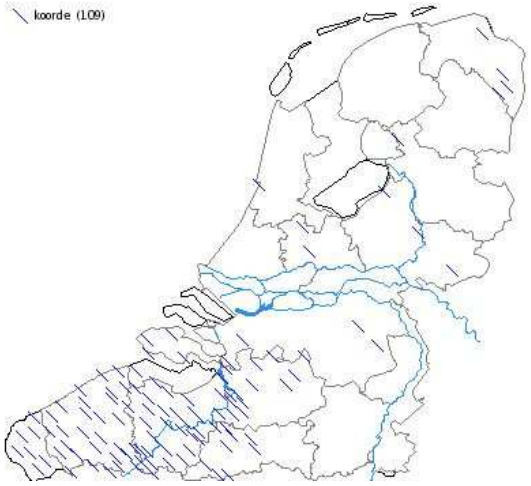


It is important to realize what had been asked here. Informants got a questionnaire in which they had to translate a sentence starting with *Ik heb* into their own dialect. No external indication had been provided as to whether this needed to be a strong form or a weak form, let alone whether a weak form would have to contain a schwa; if anything, the orthography of the Standard Dutch form would have suggested a strong form. In spite of this, many people everywhere in the language speaking area chose a weak form without a schwa. Notice that this is a situation which is very similar to the selection behaviour of children.

[12] In many varieties of Dutch, *heb* and *hoor* start with an *h* which often gets deleted in the case at hand. In the maps below, we counted transcriptions such as [khɛp] and [kɛp] as belonging to the same category.

This becomes already different when we consider the next form in our table:

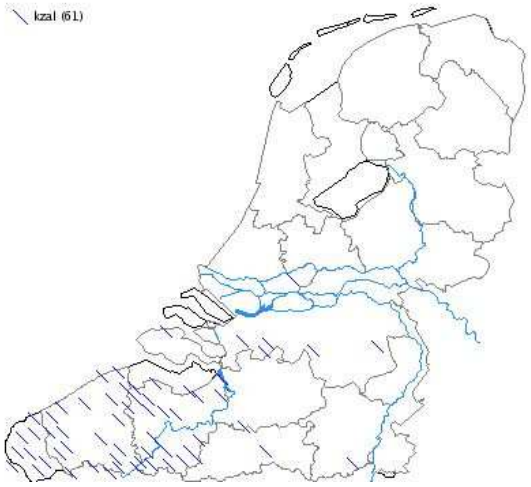
(44) *Ik hoorde* = I heard



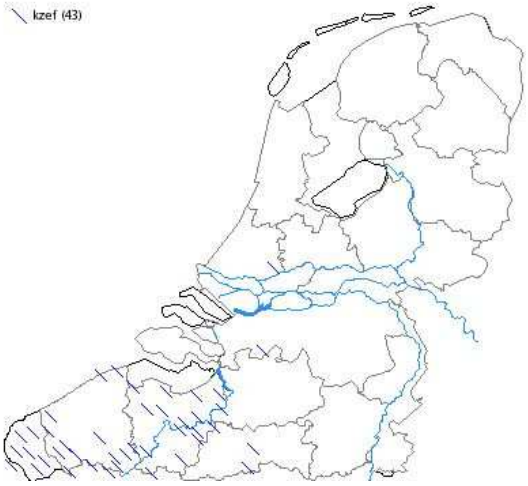
This map confirms the claims of Noske (2005): when we are not dealing with vowel-initial auxiliaries, the phenomenon becomes almost exclusive to the south(-west); in particular, we only find it in the Belgian provinces of West- and East-Flanders, and parts of Antwerp and Flemish Brabant. The south-westernmost tip is French Flanders, the area close to Dunkerque.

The following three maps show that also the rest of the hierarchy is confirmed by the geographical data:

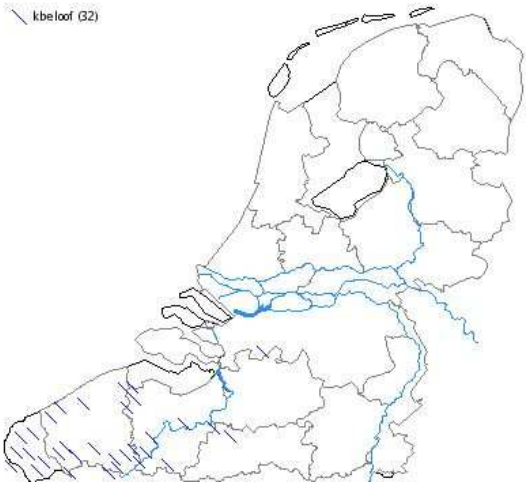
(45) *Ik zal* = I will



(46) *Ik zeg* = I say



(47) *Ik beloof* = I promise



The number of locations in which people offer a translation without a vowel becomes smaller and smaller for each form (unfortunately, *Ik ben* 'I am' was missing from the questionnaire). With the lexical verb starting with a plosive, *beloof*, the form can basically only found in French Flanders, an area where we may suspect that forms with vowels have been completely eroded.

Altogether, we can establish the following implicational hierarchy for these data:

(48)

A	B	C	D	E	
-	kɛp	kɛp	kɛp	kɛp	*have*
-	-	kor	kor	kor	*hear*
-	-	-	ksɑl	ksɑl	*will*
-	-	-	ksɛx	ksɛx	*say*
-	-	-	-	gbələfdə	*promise*

The following observations are important. First, some forms are better than some other forms, there is no absolute grammaticality. Second, the relative judgements from speakers from the Netherlands are reflected by a geographical pattern in Belgium.

These two observations are in need of an explanation. It seems plausible that all speakers have the same grammatical hierarchy, but a different cut-off point defining at what moment they are (still) willing to give an ultraweak form in response to a questionnaire such as GTRP. The question is how we can define such hierarchies. They are difficult to capture in any classical (=categorical) generative theory of grammaticality, and at first sight OT does not seem to fare any better.

4.3 Variation in OT

I claim that the problem of grammaticality hierarchies can be solved in OT if we assume a theory of ineffability such as the one described in this article, in combination with a theory of variation.

There are several theories on variation in OT; as far as I can see, they could all be helpful. I use the model of Anttila (1997a,b, 2007) as an example. In this model, it is assumed that grammars can consist of strata of unranked constraints. When constraints are unranked in the grammar, any of the possible rankings can be chosen at any point with equal likelihood. For instance, suppose that we have three unranked constraints A, B, C. These can be ranked in $3! = 6$ different ways. Suppose 4 out of 6 possible grammars give output α for input x, and 2 grammars give output β for the same input. A plausible interpretation of this is that β will be judged more questionable as a realization of x than α.

In the case of the vowelless pronouncs, it is not so clear what the competing alternatives are, since the only alternative form which can be used is the strong pronoun, which however is most probably a separate lexical entry with even a different meaning, as we have discussed above. This means that these forms will not enter in the same tableaux (we assume that it is permissible to manipulate morphosyntactic features but not the meaning). On the other hand, the weak forms with a schwa are also not acceptable in this position, since words (or phrases) do not start with a schwa (van Oostendorp, 2000b).

Given a sufficient high ranking of the constraint (which is inviolable) against schwa-initial words or phrases, the two only really interesting candidates are

one with a full vowel, and one with a null parse.[13] I propose that the following constraints are relevant:

(49) a. ALIGN: Syllables should nog cross the boundary of a lexical word.
b. PHON: Onsets should have a rising sonority slope that is sufficiently steep.

Both constraints are deliberately kept general and vague, since the focus here is on constraint interaction rather than the actual constraints. If we include (relativized) MPARSE, we get the following violation profiles for all the relevant forms (we disregard the forms starting with a plosive, since we do not have sufficient data on these):

(50)

	ALIGN	PHON	MPARSE
kɛp 'I have'	✓	✓	✓
kor 'I hear'	*	✓	✓
ksɑl 'I will'	✓	*	✓
ksɛx 'I say'	*	*	✓
∅	✓	✓	*

Even with this very simplistic model, consisting of an extremely small set of too simple constraints, we already get an interesting result. In (51), I listed the six possible rankings for our three constraints. In the columns next to those, I indicate whether the faithful form is chosen (✓) or the null parse (∅).

(51)

	k+ɛp	k+or	k+sɑl	k+sɛx
A≫P≫M	✓	∅	∅	∅
A≫M≫P	✓	∅	✓	∅
M≫A≫P	✓	✓	✓	✓
P≫A≫M	✓	∅	∅	∅
P≫M≫A	✓	✓	∅	∅
M≫P≫A	✓	✓	✓	✓
	6	3	3	2

Thus, [kɛp] is selected by all six possible rankings, whereas [ksɛx] is prefred over the null parse only two times. Furthermore all grammars which select e.g. [kor] would also select [ksɛx], but not vice versa.

As can be seen, the hierarchy follows from this at least partly: the best form is generated by the largest number of rankings, and the worst form by the smallest number of rankings. In this system, there is no difference between [kor] and [ksal], but such a difference would follow if we would assume that in

[13] Notice that this analysis necessarily hinges on the option that the weak forms with a schwa will be derived from the consonant-only forms by epenthesis of a schwa.

Dutch dialects, PHON≫ALIGN always holds: we then have only three grammars (the bottom three): three generating [kɛp], two generating [kor] and one generating [ksɑl] and [ksɛx].

We can now interpret the correlation between this variation and the geometrical pattern in the following way. We assume that all speakers of (varieties of) Dutch have these three constraints unranked (with maybe a preference for PHON≫ALIGN), and the cut-off point is a performance effect: in some regions people are more inclined to give some form different from the null parse in the southwest than they are elsewhere.

5 Conclusion

Absolute ungrammaticality at first sight presents itself as a problem which is unique for OT; the problem has hardly been discussed in other frameworks, which more generally do not seem to be too much concerned with this concept. On closer inspection, however, it turns out that most other frameworks just are not very explicit about their notions of grammaticality, which makes it difficult to evaluate them.

Further, we have seen that there is a way of formalising absolute ungrammaticality in OT which can also deal with children's selection behaviour during acquisition and relative grammaticality judgements. Neither of those have received any systematic treatment in other approaches, as far as I can see.

Since this is the case, it is extremely hard to compare the results which have been established here to what these data would look like from other perspectives. It seems to me that most alternative generative frameworks are demonstrably too absolute: they can only dinstinguish forms that are 'in' from forms that are 'out'.

This causes problems for all three of the topics we have discussed. In cases of simple absolute ungrammaticality such as diminutives it is not so clear at first sight what the filter is that has such devastating effects: actually many different kinds of things have to be excluded, in particular because the absolute ungrammaticality seems the result of an unfortunate interplay between the wish of the suffix to trigger umlaut and the unwillingness of full vowels to transmit this umlaut. The avoidance of certain words during acquisition causes similar problems: why would the derivation of words with certain sounds come to a halt at an early stage, and why would the filter relevant for this crash disappear later? Under an OT point of view, this is easy to understand, since all constraints are present at any stage of the derivation; but this is simply not true for other frameworks by definition. Finally, the relative judgements discussed in section 4 seem to fall completely outside the scope of most generative accounts of grammaticality.

The problems of frameworks like Exemplar Theory are quite different. At first, the theory seems to have a similar problem like OT with ineffability: it is not clear why certain new forms could not be produced at all, while others can. In other words, the diminutive of *homo* 'gay man' may be rare, but so was

the diminutive of *computer* at least at some point. Yet the latter has become acceptable, whereas the former has not.

What causes the absolute resistance against forming diminutives of the first type? The type of solution argued for in this article cannot be transferred immediately to Exemplar Theory. That is definitely also true for the acquisition problem, and maybe even more so: why would children systematically avoid certain words which carry just certain properties. They contain certain types of sounds for which maybe there are not so many exemplars yet — but that seems true for all sounds at the first stages of acquisition.

Further, the issue of relative judgement seems also difficult to resolve. It may seem straightforward at first sight: the forms with a small regional distribution must have a lower frequency, since they are used by fewer people. Hence, the fact that they are also judged less well-formed may be an indication that grammaticality judgements are really a derivative of frequency (Pereira, 2000). However, there is no indication at all that the speakers from the Netherlands, giving their judgements, could be in any sense aware of the dialect situation in Belgium. These are two countries and they are not very well known to each other; there is very little reason to believe that Dutch tourism or business contacts in West-Flanders are at such a level that this could have influenced Dutch speakers' judgements.

At the same time, it should be noted that two of the three datasets studied here do *not* involve classical grammaticality judgements, but two types of production: spontaneous speech in children, and answers to a dialectological questionnaire by adults. It turns out that we can understand these data better if we look at them through the lens of a theory of grammaticality

Bibliography

Ackema, Peter & Ad Neeleman (1998). 'Optimal questions'. *Natural Language and Linguistic Theory*, 16: 443–490.

Albright, Adam (2009). 'Lexical and morphological conditioning of paradigm gaps'. In: Blaho, Sylvia & Curt Rice (eds.), *Modeling ungrammaticality in Optimality Theory*. London: Equinox.

Anttila, Arto (1997a). 'Deriving variation from grammar'. In: Hinskens, Frans, Roeland van Hout, & Leo Wetzels (eds.), *Variation, Change and Phonological Theory*, pp. 35–68. Amsterdam: John Benjamins.

——— (1997b). *Variation in Finnish Phonology and Morphology*. Ph.D. thesis, Stanford University.

——— (2002). 'Variation and Phonological Theory'. In: Jack Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, pp. 206–243. Oxford: Blackwell.

——— (2007). 'Variation and optionality'. In: de Lacy, Paul (ed.), *The Cambridge Handbook of Phonology*, pp. 519–536. Cambridge, Mass.: Cambridge University Press.

Berendsen, Egon (1986). *The phonology of cliticization*. Dordrecht: Foris Publications.

Bird, Steven & Ewan Klein (1990). 'Phonological events'. *Journal of Linguistics*, 26: 33–56.

Boersma, Paul (1998). *Functional Phonology*. Ph.D. thesis, University of Amsterdam.

——— (2001). 'Phonology-semantics interaction in OT and its acquisition'. In: Kirchner, Robert, Joe Pater, & Wolf Wilkey (eds.), *Papers in Experimental and Theoretical Linguistics 6: Workshop on the Lexicon in Phonetics and Phonology*, vol. 6, pp. 24–35. Edmonton: Department of Linguistics, University of Alberta.

Bybee, Joan (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Chomsky, Noam (1955). *Transformational Analysis*. Ph.D. thesis, University of Pennsylvania.

——— (1965). *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press.

Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.

Coetzee, Andries W. (2009). 'An integrated grammatical/non-grammatical modelof phonological variation'. In: Kang, Young-Se, Jong-Yurl Yoon, Hyunkyung Yoo, Sze-Wing Tang, & Yong-Soon Kang (eds.), *Current Issues in Linguistic Interfaces*, vol. 2, pp. 267–294. Seoul: Hankookmunhwasa.

Coetzee, Andries W. & Joe Pater (2011). 'The Place of Variation in Phonological Theory'. In: Goldsmith, John, Jason Riggle, & Alan C.L. Yu (eds.), *The Handbook of Phonological Theory: Second Edition*, chap. 13, pp. 234–256. London: Blackwell-Wiley.

Coleman, John & John Local (1991). 'The 'no crossing constraint' in autosegmental phonology'. *Linguistics and philosophy*, 14, 3: 295–338.

Escudero, Paola (2005). *The attainment of optimal perception in second-language acquisition*. Ph.D. thesis, Universiteit Utrecht.

Ferguson, C. & C. Farwell (1975). 'Words and sounds in early language acquisition'. *Language*, pp. 419–439.

Féry, Caroline & Gisbert Fanselow (2003). 'Ineffability in Grammar'. *Linguistische Berichte*, Sonderheft 11: 288–310.

Fikkert, Paula (1994). *On the acquisition of prosodic structure*. Ph.D. thesis, University of Leiden.

Hale, Mark & Charles Reiss (2008). *The phonological enterprise*. Oxford: Oxford University Press.

Halle, Morris (1973). 'Prolegomena to a theory of word formation'. *Linguistic Inquiry*, 4: 3–16.

Jakobson, Roman (1942). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Lundequist.

Kager, René (1996). 'On affix allomorphy and syllable counting'. In: Kleinhenz, Ursula (ed.), *Interfaces in Phonology*, pp. 155–171. Berlin: Akademie Verlag.

Keer, Ed & Eric Bakovic (2001). 'Optionality and Ineffability'. In: Legendre, Géraldine, Jane Grimshaw, & Sten Vikner (eds.), *Optimality-Theoretic Syntax*, pp. 97–112. Cambridge, Mass: MIT Press.

Kiparsky, Paul & Lise Menn (1977). 'On the acquisition of phonology'. In: Macnamara, John (ed.), *Perspectives in Neurolinguistics and psycholinguistics*, pp. 47–78. New York: Academic Press.

Kurisu, Kazutaka (2001). *The Phonology of Morpheme Realization*. Ph.D. thesis, University of California at Santa Cruz.

Legendre, Géraldine (2009). 'The neutralization approach to ineffability in syntax'. In: Rice, Curt & Sylvia Blaho (eds.), *Modeling ungrammaticality in Optimality Theory*, Advances in Optimality Theor. London: Equinox.

Leonard, L.B., R.G. Schwartz, B. Morris, & K. Chapman (1981). 'Factors influencing early lexical acquisition: Lexical orientation and phonological composition'. *Child Development*, 52: 882–887.

Leopold, W.F. (1949). *Speech development of a bilingual child. A linguist's record*. Evanston, Ill.: Northwestern University Press.

Levelt, Clara C. (1994). *On the aquisition of place*. Ph.D. thesis, University of Leiden.

Levelt, Clara C. & Marc van Oostendorp (2007). 'Feature co-occurrence constraints in L1 acquisition'. In: Los, Bettelou & Marjo van Koppen (eds.), *Linguistics in the Netherlands 2007*, pp. 162–172. Amsterdam and Philadelphia: John Benjamins.

McCarthy, John & Matthew Wolf (2009). 'Less than zero: Correspondence and the null output'. In: Rice, Curt & Sylvia Blaho (eds.), *Ineffability in Optimality Theory*. London: Equinox. URL `http://roa.rutgers.edu/view.php3?id=998`. Also at ROA 722.

Müller, Gereon (To appear). 'Optimality-theoretic syntax'. In: Alexiadou, Artemis & Tibor Kiss (eds.), *Syntax. An international handbook*. URL `http://roa.rutgers.edu/view.php3?id=1641`.

Noske, Roland (2005). 'A prosodic contrast between Northern and Southern Dutch: a result of a Flemish-French sprachbund'. In: Broekhuis, Hans, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, & Jan Koster (eds.), *Organizing Grammar. Linguistic Studies in Honor of Henk van Riemsdijk.*, pp. 474–482. Berlin: Mouton De Gruyter.

van Oostendorp, Marc (2000a). *Phonological Projection*. Berlin/New York: Mouton de Gruyter.

——— (2000b). 'Wieringse nasaalvelarisering'. *Taal en Tongval*, 52, 1.

——— (2009). 'Dutch Diminutives and the Question Mark'. In: Rice, Curt & Sylvia Blaho (eds.), *Ineffability in Optimality Theory*. Equinox.

——— (in press). 'An Exception to Final Devoicing'. In: van der Torre, Erik Jan & Jeroen van de Weijer (eds.), *Voicing in Dutch*. Amsterdam: John Benjamins.

Orgun, Cemil Orhan & Ronald Sprouse (1999). 'From MParse to Control: Deriving ungrammaticality'. *Phonology*, 16: 191–220.

Pereira, Fernando (2000). 'Formal grammar and information theory: Together again?' *Philosophical Transactions of the Royal Society*, 358: 1239–1253.

Pesetsky, David (1998). 'Some optimality principles of sentence pronunciation'. In: Barbosa, Pilar (ed.), *Is the best good enough?*, pp. 337–383. Cambridge, Mass: MIT Press.

Prince, Alan (2002). 'Arguing Optimality'. URL http://roa.rutgers.edu/view.php3?roa=562. Also available as ROA 562-1102.

Prince, Alan & Paul Smolensky (1993). 'Optimality Theory: Constraint Interaction in Generative Grammar'. Manuscript, Rutgers University and University of Colorado at Boulder. Available at ROA and published 2004 at Blackwell.

Raffelsiefen, Renate (2002). 'Gaps in Word Formation'. In: Kleinhenz, Ursula (ed.), *Interface in Phonology*. Berlin: Akademie Verlag.

Rice, Curt (2005). 'Optimal gaps in optimal paradigms'. *Catalan Journal of Linguistics*. URL http://ling.auf.net/lingBuzz/000055.

Rice, Curt & Sylvia Blaho (eds.) (2009). *Modeling ungrammaticality in Optimality Theory*. London: Equinox.

Sagey, Elizabeth (1986). *The representation of features and relations in nonlinear phonology*. Ph.D. thesis, MIT.

Scheer, Tobias (2004). *A Lateral Theory of Phonology. Volume I: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.

Schwartz, R.G. & L.B. Leonard (1982). 'Do children pick and choose?' *Journal of Child Language*.

Schwartz, R.G., L.B. Leonard, D. M. Frome Loeb, & L.A. Swanson (1986). 'Attempted sounds are sometimes not. An expanded view of phonological selection and avoidance'. *Journal of Child Language*, 14: 411–418.

Storkel, Holy L. (2006). 'Do children still pick and choose? The relationship between phonological knowledge and lexical acquisition beyond 50 words'. *Clinilcal Linguistics and Phonetics*, 20, 7-8: 523–529.

Tesar, Bruce & Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, Mass: MIT Press.

Trommelen, Mieke (1982). *The Syllable in Dutch: With Special Reference to Diminutive Formation*. Ph.D. thesis, University of Utrecht. Published at Dordrect: Foris, 1984.

Vihman, Marilyn M. (1993). 'Variable paths to early word production'. *Journal of Phonetics*, 21: 61–82.

Vogel, Ralf (2001). 'Case conflict in German free relative constructions'. In: Müller, Gereon & Wolfgang Sternefeld (eds.), *Competition in syntax*, pp. 341–375. Berlin: Mouton de Gruyter.

de Vries, Jelle (2001). *Onze Nederlandse spreektaal*. Den Haag: Sdu Uitgevers.

Walker, Rachel & Bella Feng (2004). 'A Ternary Model of Morphology-Phonology Correspondence'. In: B. Schmeiser, A. Kelleher, V. Chand & A. Rodriguez (eds.), *Proceedings of WCCFL 23*. Somerville, MA: Cascadilla Press. URL http://www-rcf.usc.edu/~rwalker/MPCorrespondence.pdf.

Wilson, Colin (2001). 'Bidirectional optimization and the theory of anaphora'. In: Legendre, Géraldine, Jane Grimshaw, & Sten Vikner (eds.), *Optimality-theoretic Syntax*, pp. 509–543. Cambridge, Mass: MIT Press.

Wolf, Matthew (2008). *Optimal interleaving: Serial phonology–morphology interaction in a constraint-based model*. Ph.D. thesis, University of Massachusetts at Amherst.

Yavas, Mehmet (1995). 'Phonological Selectivity in the First Fifty Words of a Bilingual Child'. *Language and Speech*, 38: 189–202.

Yip, Moira (1998). 'Identity avoidance in phonology and morphology'. In: Lapointe, Steven G., Diane K. Brentari, & Patrick M. Farrell (eds.), *Morphology and its relation to phonology and syntax*, pp. 216–246. Stanford: CSLI Publications.

Zuraw, Kie (2000). *Patterned exceptions in phonology*. Ph.D. thesis, University of California at Los Angeles.