

Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○
○○○○○

The Phonology of the Unspeakable

Marc van Oostendorp

Meertens Instituut / KNAW, Amsterdam, and Leiden University

January 24, 2009

Grammaticality

○○○○○

When nobody has anything to say

○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say

○○○○○○○
○○○○

Relative judgements

○○○○○○○○○
○○○○○

The Phonology of the Unspeakable

Grammaticality

When nobody has anything to say

Things children do not say

Relative judgements

The central problem

- The focus of grammatical theory is (or should be) the notion of grammaticality; a grammar for a language \mathcal{L} should define which forms are grammatical and which forms are ungrammatical in \mathcal{L}
- The usual definition in generative grammar is:
 - Those forms are grammatical which can be generated by the grammar; all other forms are ungrammatical.

Notice that this definition implies that we consider judgements to be the central topic of inquiry, not presence in a corpus or a dictionary, although we can use the latter as indirect evidence for grammaticality.

OT as a theory of grammaticality

- One of the advantages of (Classical, monostratal, parallel) OT is that it gives a very precise locus for *language-specific* ungrammaticality:
 - A surface form is ungrammatical in \mathcal{L} because there is at least one competing form which satisfies the constraint hierarchy of \mathcal{L} in a better way
- This should be qualified, since in practice any OT analysis assumes restrictions on Gen, so that there is a different locus for (universal) ungrammaticality: universals in Gen, particulars in Eval

OT as a superior theory of grammaticality

- Most other theories of 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 MSC underlyingly, or a process would always change this form.
- This is the source of a well-known problem: duplication
- Further, there is not always a clear distinction between language-specific and universal ungrammaticality (even in true P&P models it is not always clear why one thing is a principle and another thing a parameter)

The unspeakable in OT

- A specific problem for OT is that of *absolute ungrammaticality*, the non-existence of certain forms (e.g. contributions in Rice to appear)
- This is a specific problem for OT, since it has the special property that an output is assigned to every input (something is always generated)
- Other theories do not seem to have this problem, since they can e.g. have a crashing derivation, or 'hard' constraints acting as filters

A specific problem for OT?

- I will make three points
 - First, cases of absolute ungrammaticality are not exactly trivial outside of OT either
 - Second, there is evidence that children sometimes refuse to produce forms which they arguably know
 - Third, we have to deal with ‘relative ungrammaticality’ (some forms are more grammatical than others)
- All of this points to a necessary refinement of the notion of grammaticality which can be formalized in OT

Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○○
○○○○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○○
○○○○○

The Phonology of the Unspeakable

Grammaticality

When nobody has anything to say

Things children do not say

Relative judgements

The diminutive

		Diminutive
v[u:]s	'fist'	v[y:]s-ke
b[o:]k	'book'	b[ø:]k-ske
kl[uə]stər	'cloister'	kl[yə]stər-ke
compj[u:]tər	'computer'	compj[y:]tər-ke

The umlauted vowel must be stressed

We only find umlaut if the final full vowel is stressed:

kant[uə]r	‘office’		kant[yə]r-ke
CD-r[ɔ]m	‘CD rom’		CD-r[œ]m-ke
kan[a:]l	‘channel’		kan[ɛ:]l-ke
matr[ɑ]s	‘mattress’		matr[æ]s-ke

Words without a diminutive

j[u:]d[a]s	‘rotter’
p[u]m[a]	‘puma’
h[o:]m[o]	‘gay person’
[o:]m[a]	‘grandma’
[o:]p[a]	‘grandpa’
k[a]s[a]	‘check-out point’
[o:]per[a]	‘opera’
c[a:]mer[a]	‘camera’

This gap has also been observed for German by a number of authors; e.g. Fanselow and Féry 2003

No solution: absolute ungrammaticality

A	B	C	D
*j[u:]d[æ]s-ke	*j[y:]d[æ]s-ke	*j[y:]d[a]s-ke	*j[u:]d[a]s-ke
*p[u]m[ɛ]-ke	*p[y]m[ɛ]-ke	*p[y]m[a]-ke	*p[u]m[a]-ke
*h[o:]m[∅]-ke	*h[∅:]m[∅]-ke	*h[∅:]m[o]-ke	*h[o:]m[o]-ke
*[o:]m[ɛ]-ke	*[∅:]m[ɛ]-ke	*[∅:]m[a]-ke	*[o:]m[a]-ke
*[o:]p[ɛ]-ke	*[∅:]p[ɛ]-ke	*[∅:]p[a]-ke	*[o:]p[a]-ke
*k[a]s[ɛ]-ke	*k[æ]s[ɛ]-ke	*k[æ]s[a]-ke	*k[æ]s[a]-ke
*[o:]per[ɛ]-ke	*[∅:]per[ɛ]-ke	*[∅:]per[a]-ke	*[o:]per[a]-ke

Absolute ungrammaticality also holds for underlying front vowels

z[e:]br[a]	'zebra'	*z[e:]br[ɛ]-ke	*z[e:]br[a]-ke
T[i:]n[a]	girl's name	*T[i:]n[ɛ]-ke	*T[i:]n[a]-ke
t[y]b[a]	'tuba'	*t[y]b[ɛ]-ke	*t[y]b[a]-ke
z[e:]n[y]w	'nerve'	*z[e:]n[y]w-ke	

Connection between umlaut and template

- This type of ineffability is restricted to dialects which have umlaut
- Speakers of the particular dialect can switch to the Standard Dutch suffix, which doesn't trigger umlaut:
judasje.

Grammaticality
○○○○○

When nobody has anything to say
○○○○○●○○○
○○○○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○○○
○○○○○

Absolute ungrammaticality

- These facts are clearly problematic for OT
- But how about other theories? We can go in two directions:
 - Try to do it derivationally
 - Try to do it representationally

It cannot be done derivationally

- Although there are not many purely derivational phonologists anymore, it might be worthwhile to point out that the solution cannot be found in derivations
- Purely derivational theories share the property of OT that given an input we should always be able to do something
- 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 a crash (try to formalize this in SPE!)

Well-formedness constraint

- We can posit a hard well-formedness constraint (subcategorisation frame) on the Limburgian suffix:
 - The suffix attaches to a 'perfect' trochee, i.e. a monosyllabic foot or a foot ending in schwa
- The problem is that this requirement goes together with the umlauting behaviour and that connection is not made in this way

Other hard constraints

- Alternatively, we could claim that the umlaut factor (a floating feature) needs to be linked on the surface to a stressed vowel
- While it cannot cross full vowels
- Plus, we cannot delink the features of those intermediate full vowels (while we can delink the features of the stressed full vowel)
- ...
- This mixes derivational statements/parameters with representational statements/parameters (faithfulness and markedness)

Dealing with ineffability within OT (1)

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 (Rice 2005, 2006).
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 and Smolensky 1993, McCarthy and Wolfe 2006).

Dealing with ineffability within OT (2)

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 then blocks this candidate (Orgun & Sprouse 1999)
4. The 'divergent meaning' solution: we generate a phonologically well-formed form, but one which does not have the intended semantics; the form is therefore unusable. This solution is standard in OT syntax/semantics, and will be defended here for phonology.

The interface: Consistency of Exponence

“[Consistency of Exponence] means that the lexical specifications of a morpheme (segments, prosody, or whatever) can never be affected by Gen. In particular, epenthetic elements posited by Gen will have no morphological affiliation, even when they lie within or between strings with morphemic identity. Similarly, underparsing of segments — failure to endow them with syllable structure — will not change the make-up of a morpheme, though it will surely change how that morpheme is realized phonetically. Thus, any given morpheme’s phonological exponents must be identical in underlying and surface form.”

(McCarthy and Prince 1993, 1994)

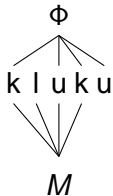
Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○○
○○●○○○○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○○
○○○○○

CoE Representation



Faithfulness constraints

- $\text{PARSE-}\phi(x)$: The morphological element x must be incorporated into the phonological structure. (No deletion.)
- $\text{PARSE-}\mu(x)$: The phonological element x must be incorporated into the morphological structure. (No insertion.)

Assumptions about morphology

- The input can be either an unstructured set of abstract morphemes, or a complex word consisting of morphemes arranged into some structure
- Lexical insertion is the job of Gen (Van Oostendorp 2007, Wolf 2007, 2008)
- The optimal output consists of a morphological word (just like the optimal output consists of a phonological word), because
- There are (M)PARSE constraints which require that individual morphemes should be part of the morphological structure.

Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○○
○○○○○○●○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○○
○○○○○

Relativized MPARSE

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

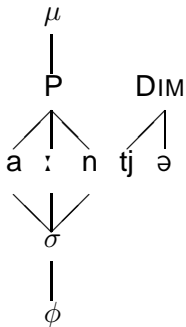
This is crucially different from the general MPARSE(M) of Prince and Smolensky 1993, in which the whole input is left morphologically unparsed.

Independent justification for MPARSE

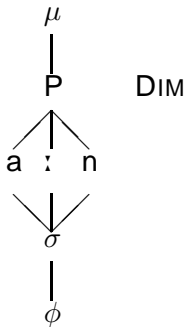
- LEXDIM: There is no diminutive of function words.

{ <i>aan</i> 'to', DIM }	LEXDIM	MPARSE(DIM)
a. <i>aantje</i> <+DIM>	*!	
☞ b. <i>aan</i>		*

Where is the phonology?



Where is the phonology?

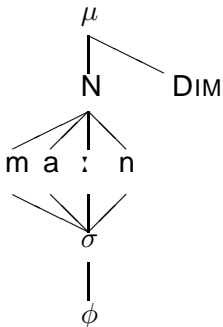


Theoretical implications

- We have to assume that the input is a bunch of morphological features
- Lexical insertion and morphological evaluation work in parallel with the phonological evaluation
- This changes the implementation of Richness of the Base, but not (necessarily) its spirit:
 - Any bunch of morphemes can be underlying
 - There are no 'inherent' restrictions on the structure of a morpheme.

Wolfe (2008) arrives at similar conclusions following a completely different line of thought.

Filling in the segments



REALIZEMORPHEME: morphological nodes should have an equivalent in the phonological representation

Not parsing the morphology for phonological reasons

$\{ \lambda x : \text{JUDAS}(x), \text{DIM} \}$	BINASS	UMLAUT	MPARSE(DIM)
☞ judas (DIM)			*
judaskə		*!	
jyɔɛskə	*!		

The winning form is *not* a diminutive and will not be treated as such by any outside module.

Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○○
○○○○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○○
○○○○○

The Phonology of the Unspeakable

Grammaticality

When nobody has anything to say

Things children do not say

Relative judgements

Avoidance of words

- It has been reported in the literature (e.g. Schwartz et al. 1986) that children seem to systematically avoid producing words which contain sounds they are not yet able to make
- even when they are able to perceive those words properly
- This means they must know certain words which they refuse to produce

R.G. Schwartz, 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.

Acquisition of features

- We assume monovalent features: [voice], [coronal], [velar], [labial], [continuant], [nasal], [lateral], [rhotic]
- Since these features are monovalent, absence of a feature gives a default interpretation
- Thus, the representation of /t/ is {[coronal]}; that of /m/ is {[labial],[nasal]}
- These seem to be learned in a specific order (mostly the same for all children)
- We are neutral on the issue of feature geometry

This is based on work done with Clara Levelt

Comparison

Child 2

P:	bdthp	m	s			wfn		kx	l
T:	bdtmnp		s	zhk	x	wfvr	l		
D:	529	540	554	588	602	615	629	643	766

Target > production: m, n, z, k, x, l, j

Production > target: h (p)

Child 4

P:	pbt dh	n			mw	s	flj	xk
T:	pbt dh kmn s f w	xl	j	z		r		
D:	479	511	546	565	590	617	643	657 703

Target > production: k, m, n, s, f, v, w, x, l, j, z, r

Child 10

P:	m	ptkh	d		n	j	b		s	wlx	f
T:	m	ptkhbdjns v		w	fl	z	x	r			
D:	589	777	797	811	827	840	854	929	985	1026	1065

Target > production: b, d, j, n, s, v, w, f, l, z, x, r

Comparison: Child 2 / Onset

P:	bdthp	m	s			wfn		kx	l
T:	bdtmnp		s	zhk	x	wfvr	l		
D:	529	540	554	588	602	615	629	643	766

Comparison: Child 10 / Onset

P:	lab nas	cor vel	voi					cont	lat
T:	lab nas	cor vel cont voi			lat			rho	
D:	589	777	797	811	840	929	985	1026	1065

Issues

- The acquisition of target segments and of produced segments can both be understood in a classical Jakobsonian way as the acquisition of features
- In particular, children systematically start targeting features just before they are actually able to produce them.
- This is a (new) indication for active selection: at some point children decide to start targeting a new sound

Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say
○○○○○○○
●○○○

Relative judgements
○○○○○○○○○
○○○○○

Three stages

- There are thus three stages in the acquisition of a sound:
 - The child knows the sound, but doesn't produce it
 - The child attempts producing the sound, but replaces it by something else
 - The child produces the sound correctly

An analysis of acquisition

- Let us take the (abstract) instance of a child which first refuses to say *sa*, then says *ta* and finally correctly produces *sa*.
- There are three relevant constraints:
 - IDENT-[continuant] (Faithfulness)
 - *[CONTINUANT] (Markedness)
 - MPARSE-SA
- We adopt the standard assumption that $M \gg F$ at the initial stage of acquisition
- And further that $F \gg \text{MPARSE}$ (the unmarkedness of silence)

Modeling the three stages

- Let us take the (abstract) instance of a child which first refuses to say *sa*, then says *ta* and finally correctly produces *sa*.
- There are three relevant constraints:
 - IDENT-[continuant] (Faithfulness)
 - *[CONTINUANT] (Markedness)
 - MPARSE-SA
- We adopt the standard assumption that $M \gg F$ at the initial stage of acquisition
- And further that $M \gg \text{MPARSE}$ (the unmarkedness of silence)

This gives us exactly the three stages

sa	M	F	MP
a. \emptyset			*
b. sa	*		
c. ta		*	

- First stage: $M \gg F$, $MP: \emptyset$
- Second stage: $MP \gg M \gg F$: ta
- Third stage: $MP \gg F \gg M$: sa

Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○○

Relative judgements
○○○○○○○○○
○○○○○

The Phonology of the Unspeakable

Grammaticality

When nobody has anything to say

Things children do not say

Relative judgements

Two types of pronouns in Dutch

strong	weak	
ɪk	ək	I
jɛi	jə	you (Du)
y	y	you (Sie)
ɦɛi	di/ti/i	he
zɛi	zə	she
het	ət	it
wɛi	wə	we
jœli	jə	you (Ihr)
zɛi	zə	they

A third type of pronoun?

- Next to the weak and strong pronouns mentioned before, there is also a third type, consisting of a consonant only:

☞	k		I
	j		you (Du)
			you (Sie)
			he
	z		she
☞	t		it
	w		we
	j		you (Ihr)
	z		they

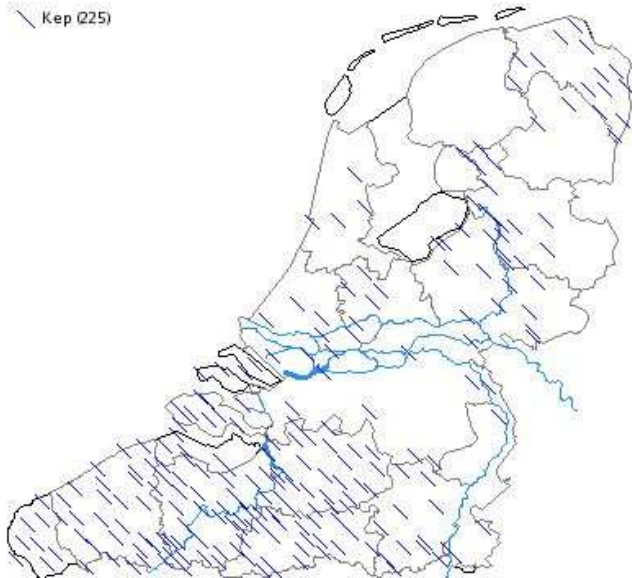
Implicational hierarchy

- κερ 'I have' γ
kor 'I hear' γ
ksal 'I will' γ
kseχ 'I say' γ
gbɛn 'I am' γ
gbəlofdə 'I believed'
- In my speech, there is no competition with a schwa version, because I cannot have that at the beginning of the phrase either; weak pronouns are simply impossible

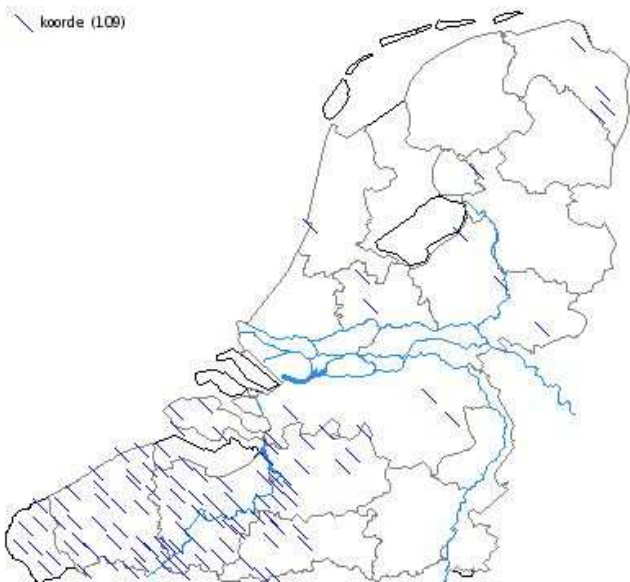
Data are from work done with Gertjan Postma

Ik heb = I have (Aux)

Key (225)



Ik hoorde = I heard



Grammaticality
○○○○○

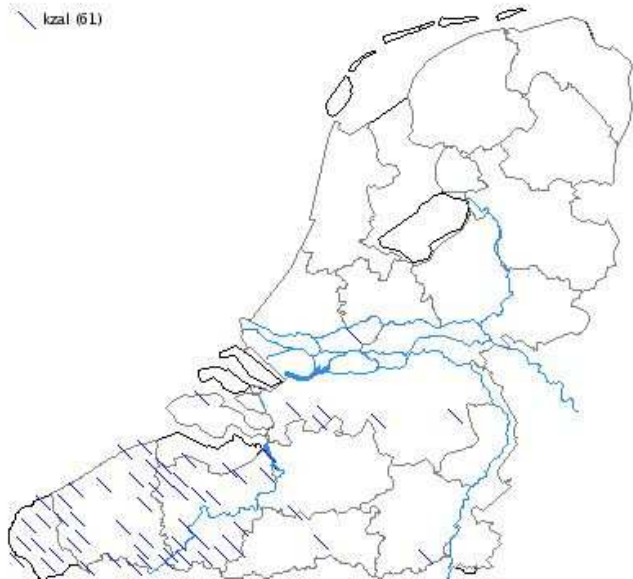
When nobody has anything to say
○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○

Relative judgements
○○○○○●○○○○○
○○○○○

Ik zal = I will

\\ kzal (61)



Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○

Relative judgements
○○○○○○●○○○
○○○○○

Ik zeg = I say

Ik zeg (43)



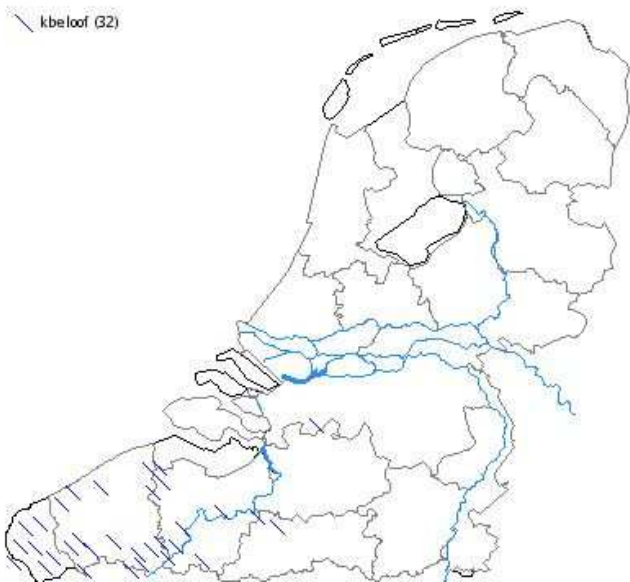
Grammaticality
○○○○○

When nobody has anything to say
○○○○○○○○○
○○○○○○○○○○○○○

Things children do not say
○○○○○○○
○○○

Relative judgements
○○○○○○●○○
○○○○○

Ik beloof = I promise



Implicational hierarchy

A	B	C	D	E	
-	kεp	kεp	kεp	kεp	<i>have</i>
-	-	kor	kor	kor	<i>hear</i>
-	-	-	ksal	ksal	<i>will</i>
-	-	-	kεx	kεx	<i>say</i>
-	-	-	-	gbəlofdə	<i>promise</i>

Two observations

- Some forms are better than some other forms, there is no absolute grammaticality
- The relative judgements from speakers from the Netherlands are reflected by a geographical pattern in Belgium
- It seems plausible that all speakers have the same hierarchy, but a different cut-off point
- Hierarchies of judgements are difficult to capture in any classical (=categorical) generative theory of grammaticality

Variation in OT

- There are several theories on variation in OT; each of these can be used as to describe relative grammaticality
- I use Anttila's model as an example
- Variation within a grammar is modeled by assuming grammars have strata of unranked constraints
- E.g. suppose we have three constraints A, B, C
- Suppose 4 out of 6 possible grammars give output α for input x , and 2 grammars give output β
- This means that β will be judged more questionable than α .

Unacceptable forms

- In this case, it is not so clear for Dutch speakers what the competing alternatives are, since the only form which can be used is the unrelated strong pronoun
- If we include (relativized) MPARSE, forms can also be in competition with ways of saying nothing
- We are dealing with phonotactic constraints as well as Alignment constraints on function words and lexical words

Comparison

	ALIGN	PHON	MPARSE
kεp 'I have'	✓	✓	✓
∅	✓	✓	*
kor 'I hear'	*	✓	✓
∅	✓	✓	*
ksal 'I will'	✓	*	✓
∅	✓	✓	*
ksex 'I say'	*	*	✓
∅	✓	✓	*

Rankings

	kεp	kor	ksal	kεx
A»P»M	✓	∅	∅	∅
A»M»P	✓	∅	✓	∅
P»A»M	✓	∅	∅	∅
P»M»A	✓	✓	∅	∅
M»P»A	✓	✓	✓	✓
M»A»P	✓	✓	✓	✓
	6	3	3	2

Conclusions

- Absolute ungrammaticality at first sight presents itself as a problem which is unique for OT
- On further inspection, it turns out that it is not a problem for other theories, since those are not very explicit about their notions of grammaticality
- Further, we have seen that there is a way of formalising absolute ungrammaticality in OT which can also deal with children's selection behaviour and relative grammaticality judgements.