

# Final devoicing, allomorphy, and Freedom of the Lexicon

An online learning approach to underlying  
representations

*Diana Apoussidou*

University of Massachusetts Amherst

University of Amsterdam

*dapou1@linguist.umass.edu*

# The problem

- How do you learn lexical/underlying representations (i.e. *hidden structure*) from surface forms?
- Cases: final devoicing, allomorphy

# Proposal

- Lexical representations are encoded in constraints → Direct interaction with the grammar
- Lexical constraints are ranked as other constraints → No specific learning mechanism for acquiring the lexicon
- Acquisition of lexicon and grammar proceed in an online fashion, simultaneously

# Dutch final devoicing

	Sg	Pl
<bed>		
UR	bɛd	bɛdən
SF	/bɛt/	/bɛdən/
<cap>		
UR	pɛt	pɛtən
SF	/pɛt/	/pɛtən/

# Conventional OT analysis

bɛd	NoSFV	ID	NoVoi	IVV
/.bɛd./	*!		*	
☞ /.bɛt./		*		

bɛd+ən	NoSFV	ID	NoVoi	IVV
☞ /.bɛ.dən./			*	
/.bɛ.tən./		*!		*

pɛt+ən	NoSFV	ID	NoVoi	IVV
/.pɛ.dən./		*!	*	
☞ /.pɛ.tən./				*

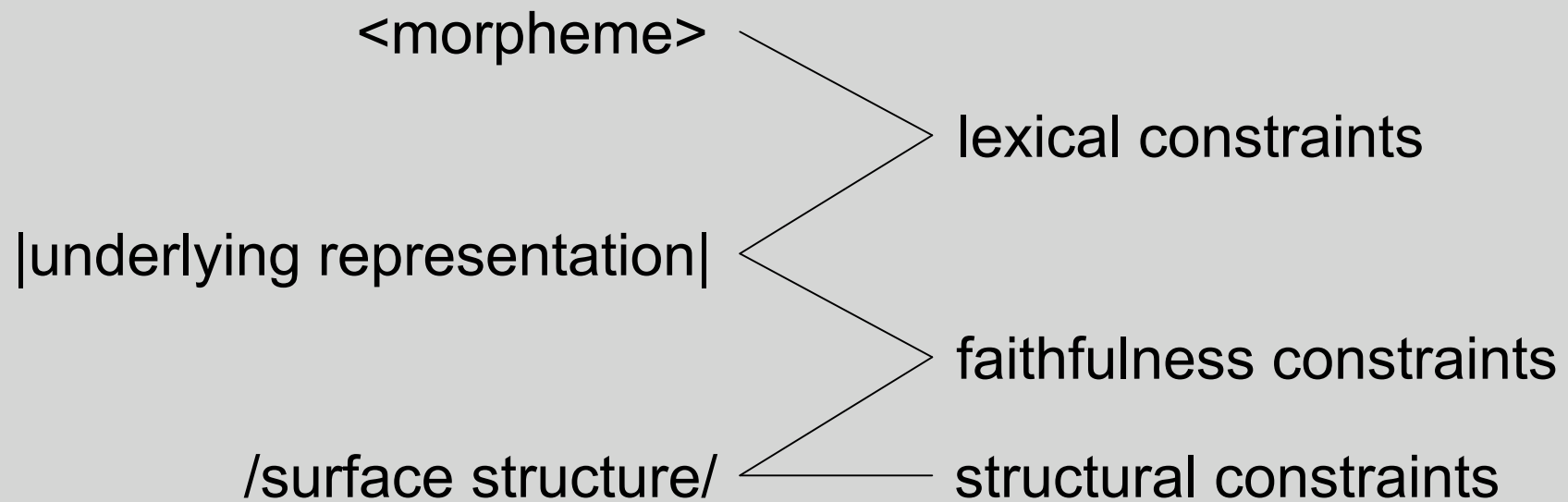
# Generating URs

Apoussidou (2006, 2007), cf. Wolf (2007)

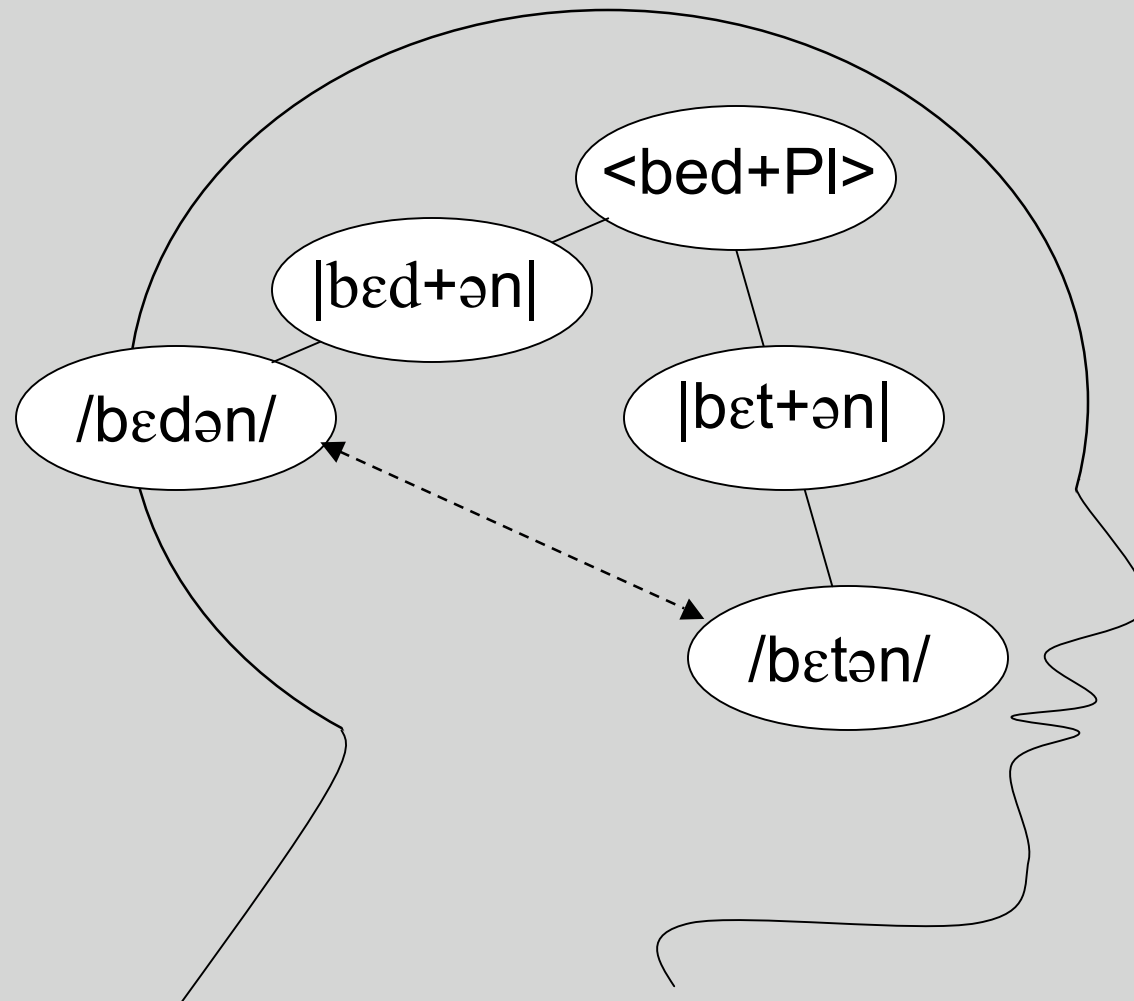
<bed+SG>	*<bed>  bɛt	*<bed>  bɛd	NoSFV	ID	NoVOI	IVV
bɛd+∅  /.bɛd./		*	*!		*	
☞  bɛd+∅  /.bɛt./		*		*		
bɛt+∅  /.bɛd./	*!		*	*	*	
bɛt+∅  /.bɛt./	*!					

<bed+PL>	*<bed>  bɛt	*<bed>  bɛd	NoSFV	ID	NoVOI	IVV
☞  bɛd+ən  /.bɛ.dən./		*			*	
bɛd+ən  /.bɛ.tən./		*		*!		*
bɛt+ən  /.bɛ.dən./	*!			*	*	
bɛt+ən  /.bɛ.tən./	*!					*

# The model



# The learning loop



# Ingredients to modelling

- Constraint set
- Data set
- Learning algorithm GLA, Stochastic OT (Boersma 1997, 1998)

# Constraints I

- NoFVoi: word-final obstruents are not voiced.
- NoVoi: obstruents are not voiced.
- IVV: obstruents between vowels are voiced.
- IdVoi: corresponding obstruents in underlying form and surface form should have the same value for voicing.

# Constraints II

- <bed>|bɛd|: connect the morpheme <bed> to the underlying form |bɛd|.
- <bed>|bɛt|: connect the morpheme <bed> to the underlying form |bɛt|.
- <cap>|pɛt|: connect the morpheme <cap> to the underlying form |pɛt|.

# Training data

/tat/<sub>1</sub><morph1-Sg>

/tate/<sub>1</sub><morph1-Pl>

/tat/<sub>2</sub><morph2-Sg>

/tade/<sub>2</sub><morph2-Pl>

/dat/<sub>3</sub><morph3-Sg>

/date/<sub>3</sub><morph3-Pl>

/dat/<sub>4</sub><morph4-Sg>

/dade/<sub>4</sub><morph4-Pl>

# A possible learning path to final devoicing

Comprehension with a random, initial ranking

/.bɛ.dən./<bed+PL>	ID	*<bed>  bɛd	NoVoi	IVV	*<bed>  bɛt	NoSFV
bɛt+ə  /.bɛ.dən./	*!		*		*	
👂  bɛd+ə  /.bɛ.dən./		*	*			

Production with the same ranking

<bed+PL>	ID	*<bed>  bɛd	NoVoi	IVV	*<bed>  bɛt	NoSFV
👄  bɛt+ə  /.bɛ.tən./				←*	←*	
bɛt+ə  /.bɛ.dən./	*!		*		*	
👂  bɛd+ə  /.bɛ.dən./		*!→	*→			
bɛd+ə  /.bɛ.tən./	*!	*		*		

# A possible learning path to final devoicing

New comprehension of / .bɛ.dən./ <bed+PL>

/ .bɛ.dən./ <bed+PL>	ID	IVV	*<bed>  bɛt	*<bed>  bɛd	NoVoi	NoSFV
bɛt+ə  / .bɛ.dən./	*!		*		*	
👂  bɛd+ə  / .bɛ.dən./				*	*	

New production of <bed+PL>

<bed+PL>	ID	IVV	*<bed>  bɛt	*<bed>  bɛd	NoVoi	NoSFV
bɛt+ə  / .bɛ.tən./		*!	*			
bɛt+ə  / .bɛ.dən./	*!		*		*	
👂 🗨️  bɛd+ə  / .bɛ.dən./				*	*	
bɛd+ə  / .bɛ.tən./	*!	*		*		

# A possible learning path to final devoicing

Encountering / .bɛt./ <bed+SG>

/ .bɛt./ <bed+SG>	ID	IVV	*<bed>  bɛt	*<bed>  bɛd	NoVoi	NoSFV
👂  bɛt+∅  / .bɛt./			*		**	*
bɛd+∅  / .bɛt./	*!			*	**	*

Producing <bed+SG>

<bed+SG>	ID	IVV	*<bed>  bɛt	*<bed>  bɛd	NoVoi	NoSFV
👂  bɛt+∅  / .bɛt./			*!→		**→	
bɛt+∅  / .bɛd./	*!		*		*	*
👄  bɛd+∅  / .bɛd./				←*	*	←*
bɛd+∅  / .bɛt./	*!			*	**	

# 3 Simulations

1. All constraints equal, all plasticities equal
2. Faith low, slow plasticity (cf. Jesney & Tessier 2007)
3. Lex equal, fast plasticity

# 1. All equal

- Default Praat (Boersma & Weenink 2007/2008) settings
- GLA, Stochastic OT
- All constraints on a par initially
- 400 000 pieces of training data, randomized
- Decreasing plasticity, evaluation noise
- 10 virtual learners

# Results I: Variation

	Singular:			Plural:		
morph1	tat	/tat/	100.0	tat+e	/ta.te/	98.6
morph2	tad	/tat/	25.9	tad+e	/ta.de/	98.3
	tat	/tat/	73.6			
morph3	dat	/dat/	100.0	dat+e	/da.te/	98.7
morph4	dad	/dat/	31.0	dad+e	/da.de/	98.8
	dat	/dat/	68.3			

# Final constraint ranking

NoSFV

>>

{ Id, \*|tad, dad|<morph1,3> }

>>

{ \*|tat, dat|<morph2,4> }

>>

{ IVV, NoVoi }

>>

{ \*|tat, dat|<morph1,3>, \*|tad, dad|<morph2,4> }

# The final 'lexicon'

\*|tad|<morph1> >> \*|tat|<morph1> = |tat|

\*|tat|<morph2> >> \*|tad|<morph2> = |tad|

\*|dad|<morph3> >> \*|dat|<morph3> = |dat|

\*|dat|<morph4> >> \*|dad|<morph4> = |dad|

# Interaction leads to variation

IdV >> \*|tat|<morph2>, \*|dat|<morph4>

IDVOICE >> lexical constraints

<morph2+SG>	NSFV	IDV	* tat <morph2>	IVV	NV	* tad <morph2>
☞ 'morph2'  tat  /tat2/			*		**	
'morph2'  tat  /tad2/	*!	*	*		*	
'morph2'  tad  /tad2/	*!				*	*
☹ 'morph2'  tad  /tat2/		*!			**	*





# Lex >> Faith

'morph2'	NSFV	* tat  'morph2'	IDV	IVV	NV	* tad  'morph2'
'morph2'  tat  /tat2/		*!			**	
'morph2'  tat  /tad2/	*!	*	*		*	
'morph2'  tad  /tad2/	*!				*	*
☺ 'morph2'  tad  /tat2/			*		**	*

# The final grammar

NoSFV >> IDVoi >> IVV, NoVoi

Novel inputs:

lradl	NoSFV	IDVoi	IVV	NoVoi
/rad./	*!			*
 /rat./		*		
lrad+ənl				
 /ra.dən./				*
/ra.tən./		*!	*	
lkatl				
/kad./	*!	*		*
 /kat./				
lkat+ənl				
/ka.dən./		*!		*
 /ka.tən./			*	

## 2. Faith low, slow

- All equal to 1., except for Faith starting out at 0 and reranking with slower plasticity than the other constraints

## 2. Results

	Singular:	%	Plural:	%
morph1	tat  /tat1/	100.00	tat+e  /tate1/	99.95
morph2	tat  /tat2/  tad  /tat2/	41.85 58.13	tad+e  /tade2/	99.95
morph3	dat  /dat3/	100.00	dat+e  /date3/	99.95
morph4	dat  /dat4/  dad  /dat4/	41.85 58.14	dad+e  /dade4/	99.94

## 2. Constraint rankings

NoSFV

>>

{|dVoi, |tat, dat|<morph1,3>, |tad, dad|<morph2,4>}

>>

{ NoVoi, IVV }

>>

{ |tat, dat|<morph2,4> }

## 3. Lex equal, fast

- All equal to 1. except for Lex reranking with higher plasticity.

# 3. Results

	Singular:	%	Plural:	%		
morph1	tat  /tat1/	100.0	tat+e  /tate1/	91.9		
			tat+e  /tade1/	8.1		
morph2	tad  /tat2/	99.9	tad+e  /tade2/	99.9		
morph3	dat  /dat3/	100.0	dat+e  /date3/	91.8		
			dat+e  /dade3/	8.6		
morph4	dat  /dat4/	20.0	dat+e  /date4/	15.9		
			dad  /dat4/	79.9	dat+e  /dade4/	4.1
					dad+e  /dade4/	79.9

### 3. Final constraint ranking

{ \*|tat,dat|<morph2,4>, \*|tad,dad|<morph1,3> }

>>

NoSFV

>>

IdVoi

>>

NoVoi, IVV

>>

{ \*|tad,dad|<morph2,4>, \*|tat,dat|<morph1,3> }

### 3. However

- Mistakes produced by the Lex-fast learners are attested in child speech (Kerkhoff 2007):
  - adult /pɛtən/ <caps> becomes \*/pɛdən/
  - adult /bɛdən/ <beds> becomes \*/bɛtən/

# Allomorphy

- English:

*a pen but an oak*

→ listed in the lexicon

→ phonology chooses variant

# An oak

<INDEFART+oak>	DEP	ONSET	*<INDEFART>  ən	*<INDEFART>  ə
ə+ok  /.ə.ok./		*!		*
ə+ok  /.ə.ʔok./	*!			*
  ən+ok  /.ə.nok./			*	

# 4. Simulation

- All constraints equal
- Default Praat settings

# Results

{ \*<IndefArt>|bip|, Dep, Onset }

>>

\*<IndefArt>|an|

>>

\*<IndefArt>|a|

<IndefArt+oak> → |an+oak| → /.a.noak./ ~ 100%

<IndefArt+pen> → |a+pen| → /.a.pen./ ~ 100%

	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
<b>DEP</b>	112.399	112.399	1.000000
<b>ONSET</b>	112.273	112.273	1.000000
*<INDEFART> bip	112.037	112.037	1.000000
*<INDEFART> an	100.075	100.075	1.000000
*<INDEFART> a	87.889	87.889	1.000000

<INDEFART+oak>	DEP	ONSET	*<INDEFART> bip	*<INDEFART> an	*<INDEFART> a
a+oak  /.a.ok./		*!			*
a+oak  /.a.?ok./	*!				*
☞  an+oak  /.a.nok./				*	
bip+oak  /.bi.pok./			*!		

<INDEFART+pen>	DEP	ONSET	*<INDEFART> bip	*<INDEFART> an	*<INDEFART> a
☞  a+pen  /.a.pen./					*
an+pen  /.an.pen./				*!	
bip+pen  /.bip.pen./			*!		

# Implications

- Lexicon as computation/constraint ranking → enables direct interaction with the grammar
- Acquiring a lexicon proceeds as acquiring the grammar → no specific learning mechanism needed (cf. Tesar et al. 2003)
- Proceeds online; lexicon and grammar can be acquired simultaneously
- One grammar and lexicon at a time (cf. Jarosz 2006)

# Outlook

- Convergence issue
- More cases (phonology, allomorphy, morphophonology)
- Harmonic Grammar

# References

- Apoussidou, D. (2007). *The learnability of metrical phonology*. Dissertation, LOT Dissertation Series 148.
- Boersma, P. & D. Weenink (2007/2008). Praat: doing phonetics by computer (Version 4.4-5.0.34). [Computer program]. [www.praat.org](http://www.praat.org).
- Kerkhoff, A. (2007). *Acquisition of Morpho-Phonology*. Dissertation, LOT Dissertation Series 164.
- Jarosz, G. (2006). Rich lexicons and restrictive grammars - Maximum likelihood learning in Optimality Theory. Dissertation, Baltimore, Maryland.
- Jesney, K. & A.M. Tessier (2007). Re-evaluating learning biases in Harmonic Grammar. In *UMOP 36: Papers in Theoretical and Computational Phonology*, ed. Michael Becker, 69-110. Amherst, MA: GLSA.
- Prince, Alan & Paul Smolensky (1993). Optimality Theory: constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder. Technical Report RuCCS-TR-2. Rutgers Center for Cognitive Science.
- Tesar, B., J. Alderete, G. Horwood, N. Merchant, K. Nishitani & A. Prince (2003). Surgery in language learning. In *Proceedings of the 22<sup>nd</sup> West Coast Conference on Formal Linguistics (WCCFL)*, eds. G. Garding and M. Tsujimura, 477-490. Somerville, MA: Cascadilla Press. ROA 619.
- Wolf, M. (2007). Lexical insertion occurs in the phonological component. To appear in Bernard Tranel (ed.), *Understanding allomorphy: Perspectives from Optimality Theory*. London: Equinox.