Lexical Economy and Optimality

Martin Krämer
University of Ulster
m.kraemer@ulster.ac.uk

Start

This paper compares the predictions of Optimality Theory with general principles of lexical economy. Furthermore it compares both with potential evidence for underlying representations

Results:

- An economic lexicon should contain as few structure as possible, few lexical entries, maximal underspecification in individual lexical entries, minimal number of phonemic contrasts etc.
- According to Prince & Smolensky (1993), Inkelas (1994) OT and in particular Lexicon Optimization predicts fully specified individual lexical entries,
- My finding: this OT prediction depends on the formulation of constraints, but: OT predicts a huge number of lexical entries,
- The 'reality' is somewhere inbetween OT and maximal economy,
- Lexicon Optimization cannot be the most important mechanism to generate lexical entries

1 Economy

(1) Radical Underspecification (Archangeli 1988)

	i	e	ɑ	o	u
high	+				+
low			+		
back				+	+

(2) Lexical economy (Yip 1996: 766)

a. economy of individual lexical entries
b. economy of phoneme inventory
c. economy of phonotactic combinations
d. economy of paradigms

2 Optimality Theory and Lexicon Optimization

Basic idea of OT: Grammar evaluates all possible surface forms by means of ranked violable constraints on output structure.

(3) Richness of the Base (Smolensky 1996:3)

> The source of all systematic cross-linguistic variation is constraint reranking. In particular, the set of *inputs* to the grammars of all languages is the same. The grammatical inventories of a language are the *outputs* which emerge from the grammar when it is fed the universal set of all possible inputs.

(4) Evaluation

/input/	CONSTRAINT A	>>	B	>>	C	>>	D
a. $output_1$	*!				*		
b. $output_2$			*!		*		
c. $output_3$					*!		
☞ d. $output_4$							*
...							
n. $output_\infty$			*!				

But: For each output there is a potentially indefinite number of inputs available.
Just as there is an indefinite number of possible outputs for any given input.

(5) Lexicon Optimization (Prince & Smolensky 1993)

> Suppose that several inputs I_1, I_2, I_3, ..., I_n when parsed by a grammar G lead to corresponding outputs O_1, O_2, O_3, ..., O_n, all of which are realized as the same phonetic form ϕ – these inputs are all phonetically equivalent with respect to G. Now one of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labelled O_k. Then the learner should choose, as the underlying form for ϕ, the input I_k.

(6) Lexicon Optimization

	CONSTRAINT A	B	C	D
a. $I_1 \sim O_1$	*!		*	
b. $I_2 \sim O_2$		*!	*	
c. $I_3 \sim O_3$			*!	
☞ d. $I_k \sim O_k$				*

P&S: Input and output are the same. This avoids unnecessary constraint violations.

(7) *SPEC (Prince & Smolensky, 1993): 'Underlying material must be absent.'
✍ Problem: One constraint on URs among many surface constraints.

(8) Inkelas' (1994) alternation-sensitive restatement of LO:

Given a set S = {S_1, S_2, ..., S_i} of surface phonetic forms for a morpheme M, suppose that there is a set of inputs I = {I_1, I_2, ..., I_j}, each of whose members has a set of surface realizations equivalent to S. There is some $I_i \in I$ such that the mapping between I_i and members of S is the most harmonic, i.e. incurring the fewest marks in grammar for the highest ranked constraints. The learners should choose that I_i as the underlying representation for M.

(9) Schematic tableau for alternation-sensitive Lexicon Optimization

			Constraint A	B	C	D
a.	I_1 ~	S_1	*!		*	
		S_2		*		*
b.	I_2 ~	S_1		*!	*	
		S_2				*
c.	I_3 ~	S_1			*!	
		S_2				*
☞ d.	I_4 ~	S_1				
		S_2				*

(10) Resulting predictions for underspecification:

	Predictable	Unpredictable
Alternating	underspecified	specified
Nonalternating	specified	specified

3 Economy and Optimality

3.1 Economy of individual lexical entries

(11) Correspondence (McCarthy & Prince 1995)

Given two related strings S_1 and S_2. Correspondence is a relation $\Re$ from the elements of S_1 to those of S_2. An element $\alpha \in S_1$ and any element $\beta \in S_2$ are referred to as correspondents of one another when $\alpha \Re \beta$.

(12) Correspondence constraints (McCarthy & Prince 1995: 264)

a. MAX-IO: Every segment in S_1 has a correspondent in S_2. ('No deletion!')

b. DEP-IO: Every segment in S_2 has a correspondent in S_1. ('No insertion!')

c. IDENT(F): Let α be a segment in S_1 and β be any correspondent of α in S_2. If α is [γF] then β is [γF].
('Correspondent segments are identical in feature F.')

(13) Evaluation with DEP(F) constraints

	DEP[high]	DEP[back]	DEP[round]
a. /V/ ~ [e]	*!	*	*
☞ b. /e/ ~ [e]			

(14) Only Identity is requested for features

	IDENT[high]	IDENT[back]	IDENT[round]
☞ a. /V/ ~ [e]			
☞ b. /e/ ~ [e]			

(15) Comparing identity and dependency

	IDENT(F)	DEP(F)
a. /V/ ~ [e]		*
b. /e/ ~ [e]		
c. /i/ ~ [e]	*	*

(16) Turkish backness and roundness harmony

	nom.sg.	nom.pl.	gen.sg.	gen.pl.
a. ‘rope’	ip	ip-l^{j}er	ip-in	ip-l^{j}er-in
b. ‘girl’	kɨz	kɨz-lar	kɨz-ɨn	kɨz-lar-ɨn
c. ‘face’	yüz	yüz-l^{j}er	yüz-ün	yüz-l^{j}er-in
d. ‘stamp’	pul	pul-lar	pul-un	pul-lar-ɨn
e. ‘hand’	el^{j}	el^{j}-l^{j}er	el^{j}-in	el^{j}-l^{j}er-in
f. ‘stalk’	sap	sap-lar	sap-ɨn	sap-lar-ɨn
g. ‘village’	k^{j}öy	k^{j}öy-l^{j}er	k^{j}öy-ün	k^{j}öy-l^{j}er-in
h. ‘end’	son	son-lar	son-un	son-lar-ɨn

(Clements & Sezer 1982: 216)

(17) Optimal UR for Turkish alternating affixes

	HARMONY	IDENT
a. / kɨz-lar/ ~ kɨzlar / ip-lar/ ~ ipler		 *!
b. / kɨz-ler/ ~ kɨzlar / ip-ler/ ~ ipler		*!
☞ c. / kɨz-lAr/ ~ kɨzlar / ip-lAr/ ~ ipler		

3.2 Economy of phoneme inventory

(18) *æ*-tensing in Belfast English (Harris 1990: 91)

Tense *æ* (i.e.[a:])

a. pass, path, laugh, man, Sam

Lax *æ*

b. tap, bat, match, back

c. panel, ladder, wagon

(19) Constraints Kager (1999):

a. *æ*-TENSING : *æC]$_\sigma$.
('No *æ* in syllables closed by fricatives or voiced consonants.')

b. *TENSE-low: 'Low vowels are lax.'

c. IO-IDENT(tense): Let α be a segment in the input, and β be a correspondent of α in the output. If α is [γtense] then β is [γtense].

(20) Example tableau for the tense vowel

	æ-TENSING	*TENSE-low	IO-IDENT
a. pæs	*!		
☞ b. pɑːs		*	

(21) Example tableau for the lax vowel

	æ-TENSING	*TENSE-low	IO-IDENT
☞ a. æsɪd			
b. ɑːsɪd		*!	

(22) Lexicon Optimization in Belfast

i.	*æ*-TENSING	*TENSE-low	IO-IDENT	DEP(F)
a. /pAs/ ~ pɑːs		*		*!
b. /pæs/ ~ pɑːs		*	*!	
☞ c. /pɑːs/ ~ pɑːs		*		

ii.	*æ*-TENSING	*TENSE-low	IO-IDENT	DEP(F)
a. /Asɪd/ ~ æsɪd				*!
b. /ɑːsɪd/ ~ æsɪd			*!	
☞ c. /æsɪd/ ~ æsɪd				

(23) *æ*-tensing without DEP(F)

i.	*æ*-TENSING	*TENSE-low	IO-IDENT
☞ a. /pAs/ ~ pɑːs		*	
b. /pæs/ ~ pɑːs		*	*!
☞ c. /pɑːs/ ~ pɑːs		*	

ii.	*æ*-TENSING	*TENSE-low	IO-IDENT
☞ a. /Asɪd/ ~ æsɪd			
b. /ɑːsɪd/ ~ æsɪd			*!
☞ c. /æsɪd/ ~ æsɪd			

3.3 Economy of phonotactic combinations

Typical example of this type of economy:

(24) Morpheme Structure Constraint on Yucatec Maya roots:

"If both consonants in a root are ejectives (p', t', ts', tʃ', k'), then they must be identical in every other respect."

(Straight 1976: 49)

Not much to say about that yet.

3.4 Economy of paradigms

(25) English plural formation

a.	children	b. dɒgz	dogs
	oxen	kæts	cats
	geese	pɔːz	paws

(26) Blowing up the English lexicon

	ASSIM	$\text{IDENT}_{\text{stem}}$	IDENT
/dɒg -z/ ~ dɒgz			
a. /kæt -z/ ~ kæts			*!
/pɔː -z/ ~ pɔːz			
/dɒg -s/ ~ dɒgz			*
b. /kæt -s/ ~ kæts			
/pɔː -s/ ~ pɔːz			*!
/dɒgz/ ~ dɒgz			
☞ c. /kæts/ ~ kæts			
/pɔːz/ ~ pɔːz			

(27) Yapese[1]

a.	[lúːb]	'breath'
	[lubáːg]	'my breath'

[1] In (27d), the symbol *V* stands for any vowel that might be found in this position, not for an underspecified vowel as elsewhere.

b. [ró:b] 'beard'
[robé:g] 'my beard' (Piggott 1999:64)

c. /root/ + /-ag / -eg/
d. /rootV/ + /-g/ i.e., /luba/, /robe/ + /-g/

(28) Yapese loss of final vowels

/luba/	NoFinalV	DepIO	MaxIO
a. luba	*!		
b. luba□		*!	
☞ c. lub			*

(29) Unexpected vowels

			NoFinalV	DepIO	MaxIO
☞ a.	/lub/	lub			
	+/ag/	lubag			
b.	/luba/	lub			*
	+/g/	lubag			

(30) More Yapese

a. [lú:b] 'breath'
[lubá:g] 'my breath'
[lubá:m] 'your breath'
b. [ró:b] 'beard'
[robé:g] 'my beard'
[robé:m] 'your beard' (Piggott 1999: 64)

(31)Unexpected vowels II

			NoFinalV	DepIO	MaxIO
☞ a.	/lub/	lub			
	+/ag/	lubag			
	/rob/	rob			
	+/eg/	robeg			
☹ b.	/luba/	lub			*
	+/g/	lubag			
	/robe/	rob			*
	+/g/	robeg			

(32)Unexpected vowels III

			DEPIO	MAXIO	IDENT
☞ a.	/lub/	lub			
	+/ag/	lubag			
	/rob/	rob			
	+/eg/	robeg			
☹ b.	/luba/	lub		*	
	+/g/	lubag			
	/robe/	rob		*	
	+/g/	robeg			
c.	/lub/ /rob/	lub			
	+/ag/	lubag			
		rob			
		robeg			*
d.	/luba/ /robe/	lub		*	
	+/g/	lubag			
		rob		*	
		robeg			

✍ Prediction: as many affixes as possible are specified, which amounts to storing the whole form.

3.5 Summary

(33) Predictions of Correspondence OT for URs

	Alternating		Nonalternating	
	predictable	unpredictable	predictable	unpredictable
individual lexical entries	-	+	?	+
phoneme inventory	-/-	+	?	+
phonotactic combinations	-/-	-/-	-/-	-/-
paradigms	+	+	?	+

4 A glimpse at reality?

(34) Turkish novel reduplications (Harrison & Kaun, 2001)

a.	kibrit	kibrit-kabr~~ı~~t	*kibrit-kabrit	'match'
	bütün	bütün-bat~~ı~~n	*bütün-batün	'whole'
b.	mali	mali-muli	*mali-mul~~i~~	'Mali'
	butik	butik-batik	*butik-bat~~i~~k	'boutique'

(35) Pattern Responsive Lexicon Optimization (Harrison & Kaun)

	Predictable	Unpredictable
Alternating	underspecified	specified
Nonalternating	*underspecified*	specified

(36) Harrison & Kaun's analysis

	IDENTI/R	ALIGN[bk]	IDENTB/R
a. /kibrit/ ~ kibrit-kabr**i**t	**!	*	**
✒ b. /kibrit/ ~ kibrit-kabrit	*	**	*
☞ c. /kibrIt/ ~ kibrit-kabr**i**t	*	*	**
d. /kibrIt/ ~ kibrit-kabrit	*	**!	*
e. /butik/ ~ butik-bat**i**k	*!		**
☞ f. /butik/ ~ butik-batik		*	*

☹ Predicts unattested contrast ➔ does not conform to ROTB.

(37) Beckman (1997): harmony = reduction of markedness violations

		*[αF]
a.	X Y | / [αF]	*
b.	X Y | | [αF] [αF]	**

(38) IO-IDENT$_{\sigma 1}$: The vowel in the first syllable of the stem is identical to its underlying specification of feature F.

(39)

/kibrit/	IO-IDENT$_{\sigma 1}$	*[+back]	*[-back]	IO-IDENT
a. [-bk] [-bk] | | k i b r i t			**!	
☞ b. [-bk] \ / k i b r i t			*	

(40) Why is the second vowel not completely underspecified?

	IDENTσ$_1$	IDENT(height)	*BACK	*HEIGHT
a. /k i b r i t/ ~ k i b r i t [-bk] [-bk] [-bk] [-bk] [+hi] [+hi] [+hi] [+hi]			**!	**
💣 b. /k i b r i t/ ~ k i b r i t [-bk] [-bk] [+hi] [+hi]			*	*
💣 c. /k i b r V t/ ~ k i b r i t [-bk] [-bk] [+hi] [+hi]			*	*
☹ d. /k i b r V t/ ~ k i b r i t [-bk] [-bk] [+hi] [+hi] [+hi] [+hi]			*	**!

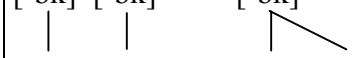
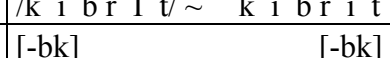
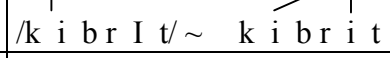
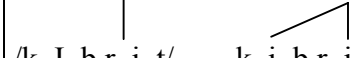
(41) Wrong prediction for reduplication:
/k i b r V t/ → *kibrit-kabrat

We need an 'anti-harmony-constraint'!

(42) *Multiple Correspondence (*MC) (Lamontagne & Rice 1995:218)
Elements of the input and the output stand in a one-to-one correspondence relationship with each other.

(43) *F$_1$ >> *MC >> *F$_2$ (F$_1$ spreads, F$_2$ does not.)

(44) Does *MC do the job?

		IO-ID$_{\sigma 1}$	*[+back]	*[-back]	*MC
a.				*	*
☹ b.				*	*
c.				*	*
d.				*	*
💣 e.				*	
f.				*	*

(45) UniqueLink (U-LINK): Every feature is exclusively associated to one segment in the output.

(46) U-LINK and URs of nonalternating harmonic forms in Turkish

	IDENT σ_1	IDENT (height)	*COLOUR	U-LINK	*HEIGHT	IDENT (colour)
a. [-bk] [-bk] [-bk] [-bk] /k i b r i t/ ~ k i b r i t [+hi] [+hi] [+hi] [+hi]			**!		**	
☠ b. [-bk] [-bk] /k i b r i t/ ~ k i b r i t [+hi] [+hi]			*	**!	*	
☠ c. [-bk] [-bk] /k i b r V t/ ~ k i b r i t [+hi] [+hi]			*	**!	*	
☞ d. [-bk] [-bk] /k i b r V t/ ~ k i b r i t [+hi] [+hi] [+hi] [+hi]			*	*	**	
☞ e. [-bk] [-bk] /k i b r V t/ ~ k i b r i t [+hi] [+hi] [+hi] [+hi]			*	*	**	

Formulation of constraints is crucial to generate the right UR.
Choice between (d) and (e) has to be made outside EVAL.

5 Finish

- If we know something about the UR we can bend OT until it gives us the right UR, but a priori OT does not tell us much about URs.
- Economy lies not at the heart of OT.
- Lexicon Optimization is a misnomer – should be Optimization of Evaluation.
- Independent economy principles are probably also not the right guide to an understanding of underlying representations.

References

Archangeli, Diana (1988). Aspects of underspecification theory. *Phonology* 5. 183-207.

— (2000). Evaluation in OT. *The Linguistic Review* 17. 219-230.

Beckman, Jill N. (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* **14**, 1-46.

Borowsky, Toni (2000). Word-faithfulness and the direction of assimilations. *The Linguistic Review* 17. 1-28.

Clements, George N. & Engin Sezer (1982). Vowel and consonant disharmony in Turkish. In: Harry van der Hulst & Norval Smith (eds.). *The Structure of Phonological Representations*. Vol. II. Dordrecht: Foris. 213-255.

Harris, John (1990). Derived phonological contrasts. In Susan Ramsaran (ed.) *Studies in the pronunciation of English. A Commemorative Volume in Honour of A. C. Gimson*. London: Routledge. 87-105.

Harrison, K. David & Abigail Kaun (2000). Pattern-Responsive Lexicon Optimization. *NELS* 30. 327-340.

— (2001). Patterns, pervasive patterns and feature specification. In T. Alan Hall (ed.) *Distinctive Feature Theory*. Berlin: Mouton de Gruyter. 211-236.

Inkelas, Sharon (1994). The Consequences of Optimization for Underspecification. In *NELS* 25. 287-302.

Kager, René (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Lamontagne, Greg & Keren Rice (1995). A Correspondence Account of Coalescence. In Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.). *University of Massachusetts Occasional Papers in Linguistics* 18. University of Massachusetts, Amherst. 211-224.

McCarthy, John & Alan Prince (1995). Faithfulness and Reduplicative Identity. *UMOP* 18. 249-384.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms., Rutgers University and University of Colorado.

Smolensky, Paul (1996). *The initial state and 'Richness of the Base' in Optimality Theory*. Technical Report, Johns Hopkins University, Baltimore.

Steriade, Donca (1995). Underspecification and Markedness. In John Goldsmith (ed.), *The handbook of phonological theory*. Oxford: Basil Blackwell. 114-174.

Straight, Stephen H. (1976). *The Acquisition of Maya Phonology. Variation in Yucatec Child Language*. New York & London: Garland Publishing.

Yip, Moira (1996). Lexicon Optimization in Languages without Alternations. In Jacques Durand & Bernard Laks (eds.) *Current trends in phonology: Models and methods*. Vol. 2. ESRI, University of Salford Publications, Salford. 757-788.