Deriving inflectional paradigms

1 Why does syncretism occur?

(1) A fairly transparent inflectional system: Turkish

	SG	PL	
NOM	ev	ev-ler	ev 'house'
ACC	ev-i	ev-ler-i	
GEN	ev-in	ev-ler-in	
DAT	ev-e	ev-ler-e	
LOC	ev-de	ev-ler-de	
ABL	ev-den	ev-ler-den	

(2) A perfectly transparent inflectional system would be:

(a) completely agglutinative, clear morpheme breaks
(b) semantically completely compositional
(c) no homonymy
(d) no portmanteau-morphemes
(e) no "zero morphemes"
(f) no allomorphy (except maybe purely phonological)
(g) no declensions or conjugations: all words inflect alike

Systems like (2) are rare, maybe nonexistent.

(3) Why should non-transparent inflectional systems exist at all? Answer (Kiparsky 2005): Inflectional paradigms emerge as a by-product of the optimal pairing of morphosyntactic categories and lexical items. The transparent one-meaning-one-form mapping can be suboptimal for at least four reasons:

- Blocking, e.g. *went* / **go-ed*.
- Category hierarchy, e.g. the German genitive plural [–LR, +PL] unifies with both *-es*:[–LR] and *-e*:[+PL] and blocking does not help. Solution: Central categories beat peripheral categories, here number beats case, e.g. *Arm-e* (**Arm-es*).
- Lack of space, e.g. **Arm-e-es* with both [+PL] and [–LR] is ruled out because there is space for only one suffix. Note how this is different from *went* / **go-ed* where both competitors satisfy the category description perfectly.
- No such lexical item, e.g. the English lexicon contains no morphemes for [PRES.1.SG], [1.SG], [SG], etc.

CATEGORY	LEXICAL ITEM	
[BE PRES 1 SG]	*am*:[BE PRES 1 SG]	the perfect match is found
[BE PRES 1 PL]	*are*:[BE PRES]	no such morpheme(s), faithfulness violation

2 Syncretism in OT

2.1 First attempt

(4) The input (universal): A fully instantiated morphosyntactic category constructed out of a universal set of inflectional categories, e.g. number, case, aspect, tense, mood, person. Examples: STEM-PL-POSS-ACC, STEM-PAST-POTENTIAL-PL-1P.

(5) What makes this view attractive: New categories can emerge in languages through grammaticalization.

(6) Labels like ACC can be further decomposed for morphosyntatic purposes. Ideally, the same features that determine the shape of inflectional paradigms also lend themselves to explaining morphosyntactic phenomena (e.g. case assignment). Example:

NOM	[]
GEN	[−LR]
ACC	[−HR]
DAT	[−LR, −HR]

(7) The lexicon: A lexical item has two parts:

- A morphosyntactic category constructed out universal inflectional categories, fully or partially specified, possibly empty (universal).
- A phonological form, possibly empty (language-specific).

(8) A sample of the English lexicon:

tree:'tree'	*fish*:'fish'	∅: 'PRES.SG.1P'
wood:'tree', MASS	∅:'fish, MASS'	*-s*: 'PRES.SG.3P'

(9) The candidate set (universal): All possible lexical items and their combinations.

(10) Constraints (universal) only refer to MARKED values (cf. Kiparsky 1994, Carstairs-McCarthy 1998).

MAX(mN)	Marked input number (e.g. PL) must have a correspondent in the output.
MAX(mC)	Marked input case (e.g. ACC) must have a correspondent in the output.
DEP(mN)	Marked output number must have a correspondent in the input.
DEP(mC)	Marked output case must have a correspondent in the input.
*EXPR	Avoid lexical items (gradient).

(11) This derives underspecification effects: unmarked features may well be present in the input, but they are invisible because no constraint in the grammar ever refers to them.

(12) An input SG is not protected by any faithfulness constraint and thus falls prey to *EXPR. Note that the grammar cannot even tell NOM and SG apart: the difference is purely cosmetic: both are really the unmarked morpheme.

	'STEM-SG-ACC'	MAX(mN)	DEP(mN)	MAX(mC)	DEP(mC)	*EXPR
a.	['STEM-SG-NOM']			*!		***
b.	['STEM-SG-ACC']					***!
c.	['STEM-PL-NOM']		*!	*		***
d.	['STEM-PL-ACC']		*!			***
e.	['STEM-SG']			*!		**
f.	['STEM-NOM']			*!		**
g.	['STEM-PL']		*!	*		**
h. ⇒	['STEM-ACC']					**
i.	['STEM']			*!		*

(13) The only possible outputs for /STEM-SG-ACC/ are ['STEM-ACC'] and ['STEM']; the rest are harmonically bounded. The following typology is generated (using OTSoft, Hayes et al. 2003):

	output #1	output #2	output #3	output #4
/STEM-SG-NOM/:	stem	stem	stem	stem
/STEM-SG-ACC/:	stem	stem	stem-acc	stem-acc
/STEM-PL-NOM/:	stem	stem-pl	stem	stem-pl
/STEM-PL-ACC/:	stem	stem-pl	stem-acc	stem-pl-acc

(14) Problems:

- In output #3, number neutralizes instead of case. This goes against the generalization that peripheral categories neutralize more readily than central categories.
- The typology does not reflect the generalization that neutralization takes place preferentially in marked categories: case is either neutralized across the board (#1, #2) or not neutralized at all (#3, #4).
- How can unmarked morphemes *ever* surface? Surely there are languages with morphemes for SG, NOM, etc.

2.2 Second attempt

(15) Why does number neutralize before case in #3? Because MAX(mC) is able to dominate MAX(mN).

	'STEM-PL-ACC'	MAX(mC)	*EXPR	MAX(mN)
a.	['STEM-SG-NOM']	*!	***	*
b.	['STEM-SG-ACC']		***!	*
c.	['STEM-PL-NOM']	*!	***	
d.	['STEM-PL-ACC']		***!	
e.	['STEM-SG']	*!	**	*
f.	['STEM-NOM']	*!	**	*
g.	['STEM-PL']	*!	**	
h. ⇒	['STEM-ACC']		**	*
i.	['STEM']	*!	*	*

(16) We need something to capture the notion of central vs. peripheral categories. Proposal: two universal rankings

	RANKING	INTUITIVE EFFECT
(a)	MAX(mN) ≫ MAX(mC)	It is cheaper to suppress 'ACC' than 'PL'.
(b)	DEP(mN) ≫ DEP(mC)	It is cheaper to insert ['ACC'] than ['PL'].

(17) Result: Pattern #3 is correctly excluded.

	output #1	output #2	output #3
/STEM-SG-NOM/:	stem	stem	stem
/STEM-SG-ACC/:	stem	stem	stem-acc
/STEM-PL-NOM/:	stem	stem-pl	stem-pl
/STEM-PL-ACC/:	stem	stem-pl	stem-pl-acc

(18) How about the generalization that neutralization takes place preferentially in marked categories? In (17), case still neutralizes either across the board or not at all.

(19) Proposal:

- There are a fixed number of slots available for inflection in any given language (functional projections, paradigm cells).
- Central categories are more resilient than peripheral categories because of MAX(mN) ≫ MAX(mC) (see above).

(20) Consequence 1: In a language with exactly one available slot, case is expressed in the singular (the unmarked number), but not in the plural (the marked number) because plural takes precedence by MAX(mN) ≫ MAX(mC), i.e case is neutralized only in the marked number.

/STEM-SG-NOM/:	stem-nom	no marked number, case must emerge
/STEM-SG-ACC/:	stem-acc	no marked number, case must emerge
/STEM-PL-NOM/:	stem-pl	marked number ousts case
/STEM-PL-ACC/:	stem-pl	marked number ousts case

(21) Constraints (preliminary version):

INFL-1	Exactly one slot available.
INFL-2	Exactly two slots available.

(22) This is how case emerges in the singular, but number in the plural:

	'STEM-SG-ACC'	INFL-1	MAX(mN)	DEP(mN)	INFL-2	MAX(mC)	*EXPR
a.	['STEM-SG-NOM']	*!				*	***
b.	['STEM-SG-ACC']	*!					***
c.	['STEM-PL-NOM']	*!		*		*	***
d.	['STEM-PL-ACC']	*!		*			***
e.	['STEM-SG']				*	*!	**
f.	['STEM-NOM']				*	*!	**
g.	['STEM-PL']			*!	*	*	**
h. ⇒	['STEM-ACC']				*		**
i.	['STEM']	*!			*	*	*
	'STEM-PL-ACC'	INFL-1	MAX(mN)	DEP(mN)	INFL-2	MAX(mC)	*EXPR
a.	['STEM-SG-NOM']	*!	*			*	***
b.	['STEM-SG-ACC']	*!	*				***
c.	['STEM-PL-NOM']	*!				*	***
d.	['STEM-PL-ACC']	*!					***
e.	['STEM-SG']		*!		*	*	**
f.	['STEM-NOM']		*!		*	*	**
g. ⇒	['STEM-PL']				*	*	**
h.	['STEM-ACC']		*!		*		**
i.	['STEM']	*!	*		*	*	*

(23) An idea based on Jaker (2005): Use ALIGN to get the "only one affix" effect. This constraint requires that all affixes must be at the right/left edge. [Explore the predictions.] This also solves another problem: recall that in a pure blocking analysis, nothing blocks **wents* for 3rd person past, unless we either (i) make /-z/ [-PAST], which is wrong since present is the unmarked tense and should not be visible at all, or (ii) require that there can be only one suffix. The ALIGN solution is a way to implement the second strategy.

(24) Consequence 2: This also explains why unmarked morphemes can surface even though there are no faithfulness constraints protecting them:

- One inflectional slot must be filled.
- If there is no meaningful material available in the input, unmarked material steps in because it is not regulated by any faithfulness constraints (Emergence of the Unmarked).

(25) This is how the unmarked case/number emerges in the singular. Note that the grammar cannot distinguish between ['STEM-NOM'] and ['STEM-SG'], i.e. the distinction is irrelevant.

	'STEM-SG-NOM'	INFL-1	DEP(mN)	INFL-2	DEP(mC)	*EXPR
a.	['STEM-SG-NOM']	*!				***
b.	['STEM-SG-ACC']	*!			*	***
c.	['STEM-PL-NOM']	*!	*			***
d.	['STEM-PL-ACC']	*!	*		*	***
e. ⇒	['STEM-SG']			*		**
f. ⇒	['STEM-NOM']			*		**
g.	['STEM-PL']		*!	*		**
h.	['STEM-ACC']			*	*!	**
i.	['STEM']	*!		*		*

(26) The full typology:

	output #1	output #2	output #3	output #4
/STEM-SG-NOM/:	stem	stem	stem	stem-sg-nom
/STEM-SG-ACC/:	stem	stem	stem-acc	stem-sg-acc
/STEM-PL-NOM/:	stem	stem-pl	stem-pl	stem-pl-nom
/STEM-PL-ACC/:	stem	stem-pl	stem-pl-acc	stem-pl-acc

	output #5	output #6
/STEM-SG-NOM/:	stem-sg/nom	stem-sg/nom
/STEM-SG-ACC/:	stem-acc	stem-acc
/STEM-PL-NOM/:	stem-pl	stem-pl
/STEM-PL-ACC/:	stem-pl-acc	stem-pl

(27) Predictions:

- Peripheral categories neutralize morphological distinctions more readily than central categories. Here: Case neutralizes before number (#2).
- Neutralization takes place preferentially in marked categories. Here: Case neutralizes in the plural first (#6).
- If a language expresses case, it will express number.
- If a language expresses NOM (the unmarked case), it will express ACC (the marked case).
- If a language expresses SG (the unmarked number), it will express PL (the marked number).
- Neutralization always happens towards the unmarked (see also Noyer 1998, Bobaljik 2002).

(28) The theory rules out the following systems:

(a) Pseudo-Latin which neutralizes number instead of case:

	SG	PL
NOM	**rēx**	**rēx**
ACC	rēgem	rēgēs

(b) Pseudo-Latin which neutralizes case in the singular, rather than in the plural:

	SG	PL
NOM	**rēx**	rēgēs
ACC	**rēx**	rēgēm

3 Further test cases

(29) German noun inflection (fragment, 2 numbers, 4 cases):

	SG	PL
NOM	Arm	Arm+e
ACC	Arm	Arm+e
GEN	Arm+es	Arm+e
DAT	Arm+(e)	Arm+en

Partial neutralization of case in both singular and plural. No neutralization of number. [Left as an exercise.]

(30) Hua verb inflection (Haiman 1980, cited in Russell 1995:23), eastern highlands of Papua New Guinea (3 numbers, 3 persons). The forms are imperative singulars. The dual allomorph is a glottal stop followed by the default morpheme.

	SG	DU	PL	
1	hu-ve	hu-ʹve	hu-pe	/hu-/ ‘do’
2	ha-pe	ha-ʹve	ha-ve	
3	hi-ve	ha-ʹve	ha-ve	

Partial neutralization of person in dual and plural (only 1 vs. 2/3). “This is a perfectly general fact throughout the morphology of Hua, and indeed, many Papuan languages: all non-singular third persons are treated as though they were the corresponding second persons.” [Left as an exercise.]

(31) A minor pattern: Due to stem Ablaut, 3.SG and 2/3.PL are homophonous.

	SG	DU	PL	
1	mu-ve	mu-ʹve	mu-pe	/mi-/ ‘give’
2	mi-pe	mi-ʹve	mi-ve	
3	mi-ve	mi-ʹve	mi-ve	

Both person and number are partially neutralized in this conjugation. No examples of paradigms with number, but no person neutralization. [How to incorporate this minor pattern into the analysis?]

(32) Baerman (2005) notes several counterexamples to almost any conceivable generalization regarding person and number, but concludes:

> There are some decided tendencies, for example, the prevalence of syncretism of 1/2 and 2/3 person in non-singular numbers (suggesting that these do constitute natural classes at some level), and the frequency with which the third person serves as a default in diachronic change. The challenge for morphological theory is to allow the formal model to be open-ended, while still giving an account of the fact that some patterns are common, and others rare. (Baerman 2005:60)

(33) Question: How to derive the generalization that portmanteau-morphemes are more likely to neutralize than agglutinative morphemes? In other words, how to exclude the following type of Pseudo-Latin-Turkish:

	SG	PL
NOM	rēg-s	**rēg-ler-s**
ACC	rēg-em	**rēg-ler-s**

4 The role of the lexicon

(34) In blocking theories, patterns of neutralization crucially depend on the presence/absence of particular lexical entries as well as their content or absence thereof (underspecification).

(35) Example: Bresnan (2001) gives an OT analysis of the paradigm of the English verb *be*.

(36) Lexical entries:

am:[BE PRES 1 SG]
is:[BE PRES 3 SG]
are:[BE PRES]
art:[BE PRES 2 SG]

(37) Constraints (note that *SG and *3 refer to unmarked values):

FAITH$^{p\&n}$	Preserve input person and number in the output.
*PL / *SG	No plural / singular in the output.
*1 / *2 / *3	No first / second / third person in the output.

(38) Questions:

- Does Bresnan's analysis predict the neutralization asymmetries discussed in this handout? Why/why not?
- If you manipulate the content of the lexical entries (e.g. assume that any combination of tense, person, and number has a corresponding lexical entry), what happens?
- What is the theoretical significance of all this?

Bibliography

Baerman, M. n.d. Syncretism: An annotated bibliography. `http: //www.surrey.ac.uk /LIS/MB/Bibliography.htm`.

Baerman, M. 2005. Typology and the formal modelling of syncretism. In *Yearbook of Morphology 2004*, 41–72. Dordrecht: Springer.

Bobaljik, J. D. 2002. Syncretism without paradigms: Remarks on Williams 1981, 1994. In *Yearbook of Morphology 2001*, 53–86. Dordrecht: Kluwer.

Bresnan, J. 2001. Explaining morphosyntactic competition. In M. Baltin and C. Collins (Eds.), *Handbook of Contemporary Syntactic Theory*. Malden, Mass.: Blackwell.

Carstairs, A. 1987. *Allomorphy in Inflexion*. London: Croom Helm.

Carstairs-McCarthy, A. 1998. Comments on the Paper by Noyer. In S. G. Lapointe, D. K. Brentari, and P. M. Farrell (Eds.), *Morphology and Its Relation to Phonology and Syntax*, 286–301. Stanford, California: CSLI Publications.

Haiman, J. 1980. [A Grammar of Hua].

Hayes, B., B. Tesar, and K. Zuraw. 2003. "OTSoft 2.1", software package. `http: //www.linguistics.ucla.edu /people /hayes /otsoft/`.

Kiparsky, P. 1994. Remarks on markedness. Paper presented at the Second Trilateral Phonology Weekend (TREND), Stanford University/UC Santa Cruz, January 22, 1994.

Noyer, R. 1998. Impoverishment Theory and Morphosyntactic Markedness. In S. G. Lapointe, D. K. Brentari, , and P. M. Farrell (Eds.), *Morphology and Its Relation to Phonology and Syntax*, 264–285. Stanford, California: CSLI Publications.

Russell, K. 1995. Morphemes and candidates in Optimality Theory. ROA-44-0195.