Lack of morpheme segmentability

- if *morpheme* = “the smallest meaningful part of a linguistic expression that can be identified by segmentation”
- “some morphologists have worked with the requirement that the segmentation of words into morphemes must be exhaustive and all meanings must be assigned to a morpheme” (HS:64)
  - Hockett 1947:332 attributes ‘the principle of Total Accountability’ to Harris 1942: ‘Every morph, and every bit of phonemic material, must be determined by (i.e. predictable from) the morphemes and the tagmemes (if any) of which the utterance is composed.’

- Problems for *phonological* segmentation
  - "Base modification" cases showed that not all morphology is concatenative
- Problems for *semantic* segmentation
  - Zero expression
  - Empty morphs
  - "Cumulative expression"/fusion/portmanteaux

“Zero affixes”/”zero expressions”

Meaning but no form:

(4.5) Coptic

čō-i	‘my head’
čō-k	‘your (M) head’
čō	‘your (F) head’
čō-f	‘his head’
čō-s	‘her head’

- ‘cough’ Fort Ware Tsek’ene

1sS duskwus
2sS dinkwus
3sS dukwus
1pS ts’idukwus
2pS dahkwus
3pS ghidukwus

subject prefixes?

Unmarked members of paradigms

- If all meanings must be assigned to a morpheme, then ‘your (f)’ must belong to a phonetically null morpheme
  - -0 ‘your (f)’
- HS 45: “zero morphemes are ad hoc devices that are posited for no purpose other than to save the principle of a concatenation-only model.”

Another approach to zero

- Nida (1965:54): "Principle 4...An overt formal difference in a structural series constitutes a morpheme if in any member of such a series, the overt formal difference and a zero structural difference are the only significant features for distinguishing a minimal unit of phonetic-semantic distinctiveness."
  - "The contrast between the singular sheep /šiyp/ and the plural sheep /šiyp/ consists of a zero and is covert."
- But "we cannot posit a zero unless it contrasts with some non-zero variant. In Japanese, where *sakana* means both 'fish (sg.) and 'fish (pl.)', we cannot posit a zero plural *(sakana-0)* because nowhere in the language does -0_{PL} contrast with a non-zero allomorph." (Aronoff and Fudeman 2011: 17)

Empty morphs

- Form but no meaning
  - ‘the non-absolutive cases share an element’, but ‘the suffixes –re, -di, and –a have no meaning’:

(4.7)

ABSOLUTIVE	*sew*	*fil*	*Rahim*
GENITIVE	*sew-re-n*	*fil-di-n*	*Rahim-a-n*
DATIVE	*sew-re-z*	*fil-di-z*	*Rahim-a-z*
SUBESSIVE	*sew-re-k*	*fil-di-k*	*Rahim-a-k*
	‘bear’	‘elephant’	(male name)

(Haspelmath 1993: 74–5)

Romance verb stem formatives

- or “conjugation markers”; e.g. Spanish

	‘talk’	‘eat’	‘live’
– infinitive	habl-a-r	com-e-r	viv-i-r
– 1pS impf	habl-a-mos	com-e-mos	viv-i-mos

- (stem formatives inherited from Latin; see Aronoff 1994 on Latin)
- Hockett 1947:337: “The conjugation vowels have no meaning.”

Cumulative expression/fusion

- analytic/isolating .....synthetic
- fusional languages are towards synthetic end of continuum
  - identifiable affixes but
  - fused semantic features
    - single phonological element ('formative') expresses two or more semantic elements; a.k.a. "portmanteau" morphemes." (HS 64)

Latin

- Latin as a “fusional” language

		‘lord’	‘song’
sg	nominative	dominus	cantus
	accusative	dominum	cantum
pl	nominative	domini	canti
	accusative	dominos	cantos

- Analysis of Latin
  - -um *acc sg*
  - -i *nom pl*
  - -os *acc pl*
  - -us *nom sg*
- Why are case and number expected to be separately marked?

Compare Hungarian

- “agglutinating”

		‘house’	‘river’
sg	nominative	ház	folyó
	accusative	házat	folyót
pl	nominative	házak	folyók
	accusative	házakat	folyókat

- Analysis of Hungarian
  - -(a)t *acc*
  - -(a)k *pl*
  - (acc pl is 2 suffixes: -(a)k_{pl}-at_{acc})

Latin case suffixes

- What is the segmentation problem?
  - -um *acc sg*
  - -i *nom pl*
  - -os *acc pl*
  - -us *nom sg*
- A *semantic* segmentation problem
  - *-u-m, *-u-m

    *acc-sg, *sg-acc
- Cross-linguistic expectation of separate marking for
  - person and number
  - case and number

Another portmanteau

- Language-internal expectation for separate marking of person and number
- McLeod Lake Tsek’ene

Imperfective paradigms

	‘cry’	‘roast (object)’
	imperfective	**imperfective**
1sS	‘ustsugh	‘usch’èès
2sS	nutsugh	nahch’èès
3sS	‘utsugh	‘ahch’èès
1dS	sìitsugh	sìich’èès
1pS	ts’utsugh	ts’ahch’èès
2pS	‘ahtsugh	‘ahch’èès
3pS	ghutsugh	ghahch’èès

1sS	s-
2sS	nu-
3sS	
1dS	sìì-
1pS	ts’-
2pS	ah-
3pS	gh-

Optative paradigms

	‘cry’	‘roast (object)’
	optative	**optative**
1sS	wustsugh	wusch’èès
2sS	wǫtsugh	wǫhch’èès
3sS	wutsugh	wahch’èès
1dS	wòòtsugh	wòòch’èès
1pS	ts’ootsugh	ts’oohch’èès
2pS	wahtsugh	wahch’èès
3pS	wootsugh	woohch’èès

1sS	s-
2sS	n-
3sS	
1dS	ìì-
1pS	ts’-
2pS	ah-
3pS	gh-

cf. impf.

1sS	s-
2sS	nu-
3sS	
1dS	sìì-
1pS	ts’-
2pS	ah-
3pS	gh-

Perfective paradigms

	‘cry’	‘shoot (O) once’	‘roast O’
	perfective	**perfective**	**perfective**
1sS	ghiitsègh	siich’ǫ	siihch’egh
2sS	ghjįtsègh	sjįch’ǫ	sjįhch’egh
3sS	ghjįtsègh	such’ǫ	sahch’egh
1dS	sughìitsègh	sìich’ǫ	sìich’egh
1pS	ts’ughjįtsègh	ts’uzch’ǫ	ts’ahch’egh
2pS	ghatsègh	sach’ǫ	sahch’egh
3pS	ghughjįtsègh	ghuzch’ǫ	ghahch’egh

perfective		cf. optative	
1sS	ii-	1sS	s-
2sS	n-	2sS	n-
3sS		3sS	
1dS	sìi-	1dS	ìi-
1pS	ts’-	1pS	ts’-
2pS	a-	2pS	ah-
3pS	gh-	3pS	gh-

gh- perfective
“conjugation markers”

s- perfective

ii- and **a-** are portmanteau morphemes
1sSPf 2pSPf

alternatively, zero morphs?
ii- 1sS **-0** Pf? **a-** 2pS **-0** Pf; or **-0** Pf **ii-** 1sS? **a-** Pf **-0** 2pS?

4. For each of the following languages, determine whether the examples exhibit cumulative expression, empty morphs or zero expression. (Some may exhibit more than one of these features.) Explain your answers.

a. Finnish pronouns (partial paradigm)

	1ST P. PL	2ND P. PL	3RD P. PL
NOM	*me* 'we'	*te* 'you'	*he* 'they'
GEN	*meidän*	*teidän*	*heidän*
PAR	*meitä*	*teitä*	*heitä*
ESS	*meinä*	*teinä*	*heinä*
INESS	*meissä*	*teissä*	*heissä*
ELA	*meistä*	*teistä*	*heistä*

nom. "-0"
gen. -idan
par. -ita
ess. -ina
iness. -issa
ela. -ista

-i- empty morph?

Pronouns: fuse person + number
me 1pS
te 2pS
he 3pS

b. Ndebele imperative verbs

ROOT	IMPERATIVE	GLOSS
lim-	*lima*	'cultivate!'
nambith-	*nambitha*	'taste!'
dl-	*yidla*	'eat!'
m-	*yima*	'stand!'
z-	*yiza*	'come!'
lw-	*yilwa*	'fight!'

-a imperative

yi- empty morph, "augment" to disyllabic---empty morph or phonologically required to satisfy minimal word? does phonological segmentation have to be exhaustive?

Axininca Campa “augment”

Root		__+V…	__+C.:.	__+RED
/na/	*Aug.*		naTA–piroTaanc^{h}i	naTA–naTA–waiTaki
	Nonaug.	na–T–aanc^{h}i	no-na–piroTi	no-na–nona–waiTi
/p/	*Aug.*		pAA–piroTaanc^{h}i	pAA–pAA–waiTaki[38]
	Nonaug.	p–aanc^{h}i		no-wA–nowA–waiTi

c. Serbian present tense verbs: GOVORITI 'to speak, say' and TRESTI 'to shake'

Is infinitive [tresti] a typo for [treseti]?

	SINGULAR	PLURAL
1ST PERSON	*govorim*	*govorimo*
2ND PERSON	*govoriš*	*govorite*
3RD PERSON	*govori*	*govore*

	SINGULAR	PLURAL
1ST PERSON	*tresem*	*tresemo*
2ND PERSON	*treseš*	*tresete*
3RD PERSON	*trese*	*tresu*

Serbian: the answer depends on the segmentation that is assumed. One possibility is:

	SINGULAR	PLURAL		SINGULAR	PLURAL
1ST PERSON	*govor-i-m*	*govor-i-mo*	1ST PERSON	*tres-e-m*	*tres-e-mo*
2ND PERSON	*govor-i-š*	*govor-i-te*	2ND PERSON	*tres-e-š*	*tres-e-te*
3RD PERSON	*govor-i*	*govor-e*	3RD PERSON	*tres-e*	*tres-u*

Under this analysis, the Serbian data exhibit all three phenomena. The morphemes *-m, -mo, -š, -te, and -e/-u* express person and number cumulatively because it is not possible to subdivide them into morphemes meaning 'singular', 'plural', '1st person', etc. The forms *-i* and *-e,* which occurs in five of the six word-forms, are empty morphs because they do not directly correspond to any aspect of meaning. The third person singular has zero expression because there is no morpheme directly corresponding to this grammatical meaning.

Another possible segmentation is:

	SINGULAR	PLURAL
1ST PERSON	*govor-im*	*govor-imo*
2ND PERSON	*govor-iš*	*govor-ite*
3RD PERSON	*govor-i*	*govor-e*

	SINGULAR	PLURAL
1ST PERSON	*tres-em*	*tres-emo*
2ND PERSON	*tres-eš*	*tres-ete*
3RD PERSON	*tres-e*	*tres-u*

This analysis has a disadvantage, in that it does not capture that the suffixes that attach to *govor-* and very similar to the ones that attach to *tres-*. However, under this segmentation, the Serbian data still has cumulative expression, but no empty morphs or zero expression.

Morpheme-based lexicon

- Descriptive (elegance) considerations alone suggest problems for morpheme-based model
  - semantic segmentation problems
  - base modification: need for morphological rules as well as lexical entries

Strict word-form lexicon

- “consists entirely of word forms, both simple and complex”

Advantages of strict word-form lexicon

- Descriptive
  - semantically unpredictable words (e.g. *reader*)
  - words formed from affixes that are no longer productive (e.g. *arrival, *confusal*)
    - ("Productive": "morphological patterns that can be used to create new words" HS 67)
- Psycholinguistic
  - words with high "token frequency"
    - are better remembered (HS 68)
    - accessed faster (HS 73)
  - suggests word storage

Disadvantages of strict word-form lexicon

- # words which must be memorized (in some lgs.)
  - Witsuwit'en inflectional possibilities, regular verbs
    - 4 tense/aspects
    - 7 subjects
    - 2 polarities
    - Are all 56 forms really memorized?
    - + regular derivational affixes...
  - Turkish verbs have "at least 2000" forms (HS)

Evidence for word-internal structure

- Strict word-based lexicon assumes morphological rules apply to whole words. But:
- Morphological phenomena that refer to word-internal structure
  - Dutch past participles
    - *spreken* ‘to speak’, *ge-sproken*
    - *be-spreken* ‘to discuss’, *be-sproken*

Witsuwit’en inceptive formation

- refers to word-internal structure
- -je ‘sg. goes (on foot)’
  - inceptive t- (s): tɛzje ‘he/she left (walking), started to walk’
  - *continuative* derivation
    - nəsəje ‘he/she walked around’
    - inceptive ne#d- (e): nedinje ‘he/she started to walk around’
- w-Git ‘dig’
  - noozGit ‘he/she dug around’
  - newdinGit ‘he/she started to dig around’

Phonological phenomena refer to word-internal structure

- HS Italian s-voicing example
  - [s]/[z] in complementary distribution
  - Intervocalic s-voicing applies
    - within roots: a[z]ola ‘buttonhole’, ca[z]a ‘house’
    - after unproductive prefixes: re[z]istenza ‘resistance’
    - before suffixes: ca[z]e ‘houses’
    - after productive C-final prefixes: di[z]onesto ‘dishonest’
  - Intervocalic s-voicing doesn’t apply
    - after clitic: la[s]irena ‘the siren’
    - root-initially within compound: tocca[s]ana 'cure all‘
    - ➢after productive V-final prefixes: a[s]ociale ‘asocial’

- S-voicing must see morphological structure?
  - a-[s]ociale
  - ca[z]-e
  - di[z]-honesto
- Nespor and Vogel 1987: s-voicing applies PWd internally; PWd construction sensitive to morphological structure
  - $_{PWd}$[a]$_{PWd}$[[s]ociale]
  - $_{PWd}$[ca[z]-e]
  - $_{PWd}$[di[z]-honesto]; *$_{PWd}$[di[s]]$_{PWd}$[honesto] because Italian PWd must end in a vowel

HS: moderate word-form lexicon

- Both words, word-schemata in lexicon

(4.11) word lexical entries (Russian)

a. [/ruka/$_N$ 'hand.NOM.SG']

b. [/ruku/$_N$ 'hand.ACC.SG']

c. [/rʲiba/$_N$ 'fish.NOM.SG']

d. [/rʲibu/$_N$ 'fish.ACC.SG']

e. [/sestra/$_N$ 'sister.NOM.SG']

f. [/sestru/$_N$ 'sister.ACC.SG']

(4.12) word-schema lexical entries (Russian)

a. suffixes

[/Xa/$_N$ 'x.NOM.SG'] [/Xu/$_N$ 'x.ACC.SG']

b. roots

[/rukX/$_N$ 'hand'] [/rʲibX/$_N$ 'fish'] [/sestrX/$_N$ 'sister']

morphological patterns as lexical entries

But which complex words are listed?

- for one thing, “the set of words in a language is never quite fixed” HS 71
- Psycholinguistic literature: factors leading to word-form storage
  - outputs of non-concatenative morphology (*Väter*)
  - phonological changes in base (*divinity*)
  - high token frequency (*insane*) relative to base (*sane*)