

Morphology

- ▶ Study of how words form
- ▶ Derivational morphology: create a new *lexeme* from a base
 - estrangle (v) => estrangement (n)
 - become (v) => unbecoming (adj)
 - ▶ May not be totally regular: enflame => inflammable
- ▶ Inflectional morphology: word is inflected based on its context
 - I become / she becomes
 - ▶ Mostly applies to verbs and nouns

Morphological Inflection

► In Spanish:

		singular			plural		
		1st person	2nd person	3rd person	1st person	2nd person	3rd person
		yo	tú vos	él/ella/ello usted	nosotros nosotras	vosotros vosotras	ellos/ellas ustedes
indicative	present	llego	llegas ^{tú} llegás ^{vos}	llega	llegamos	llegáis	llegan
	imperfect	llegaba	llegabas	llegaba	llegábamos	llegabais	llegaban
	preterite	llegué	llegaste	llegó	llegamos	llegasteis	llegaron
	future	llegaré	llegarás	llegará	llegaremos	llegaréis	llegarán
	conditional	llegaría	llegarías	llegaría	llegaríamos	llegaríais	llegarían

Morphological Inflection

- ▶ Not just verbs either; gender, number, case complicate things

Declension of Kind [hide ▲]					
	singular			plural	
	indef.	def.	noun	def.	noun
nominative	ein	das	Kind	die	Kinder
genitive	eines	des	Kindes, Kinds	der	Kinder
dative	einem	dem	Kind, Kind¹	den	Kindern
accusative	ein	das	Kind	die	Kinder

- ▶ Nominative: I/he/she, accusative: me/him/her, genitive: mine/his/hers
- ▶ Dative: merged with accusative in English, shows recipient of something
 - I taught the children <=> Ich unterrichte die Kinder
 - I give the children a book <=> Ich gebe den Kindern ein Buch

Irregular Inflection

- ▶ Common words are often irregular
 - ▶ I am / you are / she is
 - ▶ Je suis / tu es / elle est
 - ▶ Soy / está / es
- ▶ Less common words typically fall into some regular *paradigm* — these are somewhat predictable

Agglutinating Languages

- ▶ Finnish/Hungarian (Finno-Ugric), also Turkish: what a preposition would do in English is instead part of the verb

		active	passive
1st		halata	
long 1st ²		halatakseen	
2nd	inessive ¹	halatessa	halattaessa
	instructive	halaten	—
3rd	inessive	halaamassa	—
	elative	halaamasta	—
	illative	halaamaan	—
	adessive	halaamalla	—
	abessive	halaamatta	—
	instructive	halaaman	halattaman
	partitive	halaamista	
4th	nominative	halaaminen	
	partitive	halaamista	
5th ²		halaamaisillaan	

inessive
 1st sing. halaten
 2nd sing. halattesi
 3rd sing. halattensa
 1st plur. halattamme
 2nd plur. halattanne
 3rd plur. halattaneen

illative
 1st sing. halattani
 2nd sing. halattasi
 3rd sing. halattansa
 1st plur. halattamiin
 2nd plur. halattamiin
 3rd plur. halattamiin

halata: “hug”

illative: “into”

adessive: “on”

- ▶ Many possible forms — and in newswire data, only a few are observed

Morphologically-Rich Languages

- ▶ Many languages spoken all over the world have much richer morphology than English
 - ▶ CoNLL 2006 / 2007: dependency parsing + morphological analyses for ~15 mostly Indo-European languages
 - ▶ SPMRL shared tasks (2013-2014): Syntactic Parsing of Morphologically-Rich Languages
- ▶ Word piece / byte-pair encoding models for MT are pretty good at handling these if there's enough data