

# Lec6: Morphology-II

HUL 242, 2016

## Recap: Strategies for word formation

- bittersweet, homework, sleepwalk
- The second element is usually the head
- The part of speech of compound usually belongs to the second element
- Stress is usually on the first component

## Compound spelling

- May be written together or separately e.g. cook book  
cookbook
- sometimes spelling indicates degree of semantic closeness (but not deterministic) variation black-bird blackbird, black-board  
blackboard
- There may also be more than one element- man about town,  
master of ceremonies

# Productivity

- Enlarging the vocabulary
- Combine elements to form new concepts: German *kabelsalat*
- These elements can be morphological i.e. derivational affixes
- They may also be at the level of lexical items (words)

## More examples

- Move from just compounds to other kinds of multiword expressions
- Proper Names: Jantar Mantar, Pranab Mukherjee
- Predicating expressions yaad karnaa, likh lenaa,
- Collocations (but still idiosyncratic)jal prapat, aam ras
- Idioms: ungli uthaana: accuse
- Reduplications: khat-khat, rang-birangi

## Lexicon

- Define multiwords as : 'a combination of words that must be treated as a unit at some level of linguistic processing' (Calzolari, 2002)
- What percentage of your mental lexicon is simple words - and what percentage is multi-words (such as compounds?)
- 50-70% based on estimates (Jackendoff and others)
- Are these phenomena marginal or central to the way we conceptualize the world?

## Multiwords difficult to learn?

- verb particles: type of multiword expression
- e.g. put on, keep up, get out, fall down, sit down
- Childe corpus: collection of child directed speech in many languages
- children (0-4) and adults produce similar proportion of VPC (verb particle constructions) (Villavicencio et al, 2012)

Sentences Parsed	Children's set 482,137	Adult set 988,101
with VPC	38,326	82,796
with VPC	7.95	8.38

## Why should we care?

- processes that participate in word formation may tell us more about the property of language productivity
- applications with respect to a few areas: lexicography and lexical resources
- applications such as machine translation *Money laundering*= *paise dhonaa* or conversely *gala baithnaa* = *throat sitting* (Try Google Translate)



## Properties of multiwords

- single element to a single meaning
- multiple elements to a single meaning
- Non-compositional : but is it enough?

# Properties of multiwords

## Examples

give a present, take a look, give a whirl, make a mistake, take effect, give a groan, give confidence, take a bow

# Properties of multiwords

## Examples

give a present, take a look, give a whirl, make a mistake, take effect, give a groan, give confidence, take a bow

- Arrange them in order of abstractness (of meaning)?

# Properties of multiwords

## Figurativeness

Classify expressions as Abstract, Metaphorical or Idiomatic?

give a present, give a speech, give advice, give confidence, give orders, give permission, give a groan, give a whirl

- What might be a property common to all of these? Does 'give' mean the same thing?

## Properties of multiwords

### Flexible?

E.g. Add adjective: Sam gave a loud groan  
Passivize? Relativize?

A groan was given by Sam

The groan which Sam gave was very loud

What about *give a whirl* ?

## Properties of multiwords

- Can other nouns occur instead of [uŋgli ut<sup>h</sup>ana]?
- Can these expressions be separated?

## Measure multiwords?

- compare the probability of observing a word  $x$  and  $y$  together (joint probability) with the probability of observing  $x$  and  $y$  by chance (independently)
- if  $x$  and  $y$  are close (multiwords) then joint probability is larger than chance

$$I(x; y) = \log_2 \frac{P(x,y)}{P(x)P(y)}$$