Word and Verb Nets for Semantic Processing

Understanding Verbs and the Verbal Complex

Miriam Butt

DAAD Summer School in Advanced Language Engineering,
Kathmandu University, Nepal

September 4, 2012

Outline Week 2

Day 7 (Today):

- More on Auxiliaries, Modals, etc.
  (Discussion of simplified Austen text)
- Intransitives: Unergatives vs. Unaccusatives
- Verbs of "Reduced" Transitivity

Intransitives: Unergatives vs. Unaccusatives

Crosslinguisticslly intransitives can in fact be divided into two categories (Perlmutter and Postal 1984:98–99).

Unergatives	
Willed or Volitional Acts	*speak, laugh, walk, cry* . . .
Manner of Speaking	*whisper, mumble, bellow* . . .
Animal Sounds	*bark, neigh, roar* . . .
Involuntary Bodily Processes	*cough, sneeze, belch* . . .

Unaccusatives	
Affected Argument	*burn, fall, dry* . . .
Inchoatives	*melt, die, grow* . . .
Existing and Happening	*exist, happen, arise* . . .
Involuntary Emission of Stimuli	*shine, clink, stink* . . .

Examples — Case Marking Difference

Languages like Choctaw make a difference in terms of **case marking** of the subject: the subject of unaccusatives is accusative (yes, this is confusing terminology).

(1) illi-li-tok   kiyo
    die-1st.Nom-Pst not
    'I did not die.'

(2) sa-laksha
    1st.Acc-sweat
    'I sweated.'

Examples — Case Marking Difference

Languages like Urdu make a difference in terms of **case marking** of the subject: the subject is ergative in unergatives (yes, there is method to the madness).

(3) a. nadya   gɑ-yi
    Nadya.F.Sg.Nom go-Perf.F.Sg
    'Nadya went.'

   b. nadya=ne   bol-a
    Nadya.F.Sg=Erg speak-Perf.M.Sg
    'Nadya spoke.'

Reason behind the Confusing Terminology

- Examples like the Choctaw show that although one has an *accusative* argument, this ends up being the subject. Since accusatives are usually objects, one thought about these as *unaccusative*.
- Similarly, if one thinks of ergative structures as being ones in which there is an agent **and** a patient (normal transitive, agentive verbs), then a sentence in which there is **only** an agent argument can be thought of as *unergative*.

The original inventor of these terms has since apologized (Geoffrey Pullum, *The Great Eskimo Vocabulary Hoax*).

Examples — Auxiliary Difference

Languages like Italian and German make a difference as to what auxiliary is used: 'be' is used with unaccusatives, 'have' with unergatives.

(4) a. Ich   bin   gegangen.
       I.Nom be.Pres.1.Sg   go.Perf
       'I went.'

   b. Ich   habe   gesprochen.
      I.Nom have.Pres.1.Sg speak.Perf
      'I spoke.'

Similarly in Urdu, the choice of *light verb* is sensitive to the the difference.

(5) a. nadya   mɑr gɑ-yi/*di-ya
  Nadya.F.Sg.Nom die go-Perf.F.Sg/give-Perf.M.Sg
  'Nadya died.'

 b. nadya=ne bol   *gɑ-ya/di-ya
  Nadya.F.Sg=Erg speak-Perf.M.Sg go-Perf.M.Sg/give-Perf.M.Sg
  'Nadya spoke.'

Different Types of Transitives

- Not all transitives are alike.
- It is well known that so-called *psych-predicates* like *fear, frighten, experience, learn, know* behave differently from agentive transitive verbs.
- These are generally thought of in terms of "reduced" transitivity/agentivity
- But the situation is much more complex and is just beginning to be explored.
- **South Asian languages have a lot to contribute**.

Acquisition of Subcategorization Frames

- One very useful resource is list of verbs along with their properties — these are known as *subcategorization frames*.
- For languages like English and German these have been acquired via a combination of manual and automatic means.
- Raza (2011) presents a first attempt at the semi-automatic acquisition of subcategorization frames for Urdu.
- Raza identifies many challenges for such a task.
- Among them is the fact that nouns and adjectives also take arguments and that it is difficult to distinguish these arguments from verbal ones, particularly because Urdu allows for word order scrambling.

Acquisition of Subcategorization Frames

The following frames are taken from work by Ghulam Raza.

Argument-taking adjectives in Urdu (Raza)

Nr.	Type of Argument	Example of Adjective Phrase
(i)	Dative Marked	sadr=kO hAsil president=Dat possessed 'possessed by the president'
(ii)	Ablative Marked	adliyah=sE xAif courts=Abl afraid 'afraid of courts'
(iii)	Locative Marked	buxAr=mEN mubtalA fever=Loc.in suffered 'suffered with fever'
(iv)	Adpositional	sihat=kE liyE muzir health=Gen for harmful 'harmful for health'

Simple examples of argument-taking nouns (Raza)

(6) a. istisnA
    'immunity'

  b. muqaddamAt=sE istisnA
    court-case.Pl=Abl immunity
    'immunity from court-cases'

  c. muqaddamAt=sE AInI istisnA
    court-case.Pl=Abl constitutional immunity
    'constitutional immunity from court-cases'

Simple examples of argument-taking nouns

(7) a. barIfiNg
      'briefing'

   b. salAmtI=par barIfiNg
      security=Loc briefing
      'briefing on security'

   c. salAmtI=par tafsIlI   barIfiNg
      security=Loc detailed briefing
      'detailed briefing on security'

Simple examples of argument-taking nouns (Raza)

(8) a. mutAlbA
     'demand'

   b. ArmI-cIf=sE   mutAlbA
       army-chief=Abl demand
       'demand to the army-chief'

   c. ArmI-cIf=sE   qAnUnI mutAlbA
       army-chief=Abl legal   demand
       'legal demand to the army-chief'

Examples of discontinuous NPs (Raza)

(9)a1. sadr=kO$_1$ hAsil$_1$ muqaddamAt=sE$_2$
     president=Dat   possessed   court-cases=Abl

     AInI   istisnA$_2$
     constitutional   immunity

     'Constitutional Immunity from court-cases possessed
     by the president'

    a2. [$_{NP}$[$_{AP}$[$_{KP}$ sadr=kO] hAsil][$_{KP}$ muqaddamAt=sE] AInI istisnA]

      b. muqaddamAt=sE$_2$ sadr=kO$_1$ hAsil$_1$ AInI istisnA$_2$

      c. sadr=kO$_1$ muqaddamAt=sE$_2$ hAsil$_1$ AInI istisnA$_2$

      d. *hAsil$_1$ muqaddamAt=sE$_2$ sadr=kO$_1$ AInI istisnA$_2$

Verbs of Asking

The following is taken from Khan (2009).

Table 2.3: Case markers on addressee in different South Asian languages

Language	Say to		Ask (a question)	
	Marker	**Ref.**	**Marker**	**Ref.**
Punjabi	DAT/ACC	2.3.3.2	DAT/ACC, ABL	2.3.3.2
Saraiki	DAT/ACC	2.3.4.2	ABL	2.3.4.13
Nepali	DAT/ACC	2.3.6.2	COM	2.3.6.5
Manipuri	LOC/DAT	2.3.9.3	LOC/DAT	2.3.9.3
Pashto	ALL	2.3.7.1	ABL	2.3.7.5
Balochi	DAT/ACC	2.3.8.1	ABL	2.3.8.3
Sindhi	DAT/ACC	2.3.5.2	ABL	2.3.5.10
Malayalam	COM	2.3.10.6	COM	2.3.10.6
Urdu/Hindi	ABL/INST/COM	2.3.1.5	ABL/INST/COM	2.3.1.5

Other Verbs of Reduced Transitivity

The following is taken from Khan (2009).

Table 2.4: Classes of NCSA verbs in South Asian languages

Class	Subject Marking	2nd Arg. Marking	Examples
I	NOM/ERG, DAT	ABL	fear
II	NOM/ERG	ABL	ask, beg
III	NOM/ERG	LOC-on/DAT	bless, govern, attack, blame
IV	NOM/ERG, DAT	LOC-on/DAT	trust, doubt, suspect, believe
V	NOM/ERG	COM/DAT	talk, meet, marry
VI	NOM/ERG, DAT	COM	love, hate

Psych Verbs

- Psych verbs tend to be realized via N-V complex predicates in South Asian languages (mostly, but not always).
- This creates its own challenges.

(10) a.
    larki=ne   kahani   yad   k-i
    girl.F.Sg=Erg story.F.Sg.Nom memory.F.Sg.Nom do-Perf.F.Sg
    'The girl remembered a/the story.' (lit.: 'The girl did memory of the story.')

  b. larki=ko   kahani   yad   hε
    girl.F.Sg=Dat story.F.Sg.Nom memory.F.Sg.Nom be.Pres.3P.Sg
    'The girl remembers/knows a/the story.' (lit.: 'Memory of the story is at the girl.')

Background Literature

The information in these slides was taken from the following literature

- Butt, Miriam. 2006. *Theories of Case.* Cambridge University Press.

- Ahmed, Tafseer. 2010. The Unaccusativity/Unergativity Distinction in Urdu. *Journal of South Asian Linguistics* 3. jsal-journal.org

- Khan, Tafseer Ahmed. 2009. *Spatial Expressions and Case in South Asian Languages*, Phd thesis, University of Konstanz. (see the link under "Dissertations" in jsal-journal.org)

- Raza, Ghulam. 2011. *Subcategorization Acquisition and Classes of Predication in Urdu.* Phd thesis, University of Konstanz. (see the link under "Dissertations" in jsal-journal.org)