Distributed Lexical Functional Grammar
MOTH 2016 UT Mississauga
April 15th, 2016

Ash Asudeh & Daniel Siddiqi
Oxford University & Carleton University

Abbreviations:

DM: Distributed Morphology (Halle & Marantz 1993)
LFG: Lexical Functional Grammar (Kaplan & Bresnan 1982)
MPP: Minimalist Principles & Parameters (Chomsky 1995)
PFM: Paradigm-Function Morphology (Stump 2001, Stump 2016)
MSI: Morphology-Syntax Interface
LIH: Lexical Integrity Hypothesis
OT: Optimality Theory (Prince & Smolensky 2004)

0. Purpose of the talk

➢ Argue for the benefits of a DM interface with LFG.
➢ Propose an architecture for such an interface.
➢ Discuss the strengths of such a model.

1. Why would LFG want an interface with DM?

➢ LFG typically assumes an *incremental* approach to morphology (Bresnan et al 2016).

   o Words are generated in the lexicon via lexical rules which in essence add morphology to add the featural content of the word.

   o These words then form the atoms of the c-structure and supply the f-structure with its crucial elements.

   o Morphology is *information-increasing* (Stump 2001)

➢ At the same time, most models of morphology have abandoned the incremental approach to morphology over the course of the last two decades.

   o Contemporary morphology is typically *realizational* (Beard 1995)
      a. Morphology *expresses* syntactic information.

      b. Morphology is *associative* (Stump 2001).

      c. Some prevalent realizational models: Anderson (1982, 1992), Halle & Marantz (1993, 1994), Stump (2001, 2016), Starke (2005, 2009), Wunderlich 1996, Ackema & Neelemn 2004.

   o Key reasons to prefer realizational approaches (see Stump 2001):

      a. Multiple exponence.

         • A given grammatical feature has more than one morphological reflex.

      b. Underdetermination.

• The grammatical features expressed by a word-form can be fewer than those associated with the context
   o i.e. morphology can be inherently *underspecified* for the environments it appears in.

o Realizational models come in two varieties:

   ▪ Word & Paradigm

      • Realizational listemes are entire word-forms, typically generated in a paradigm.

      • PFM (Stump 2001) is the dominant model

   ▪ Morpheme-based

      • Listemes are morphemes: minimal form to feature correspondences.

      • DM is the dominant model.

o The work so far in attempting to get LFG to be realizational has been about interfacing it with PFM.

   ▪ Spencer 2003, 2006
   ▪ Luis & Sadler 2003
   ▪ Sadler & Spencer 2004
   ▪ Sadler & Norlinger 2004, 2006
   ▪ Luis & Otoguro 2004
   ▪ Spencer & Sadler 2001
   ▪ Dalrymple 2015

o Nearly all these models make shared assumptions.

   ▪ Realization is a mapping from f-structure to m-structure to PFM.

      • Avoiding c-structure mappings preserves the Lexical Integrity Hypothesis.

      • Mapping to intermediate level preserves the morpheme hypothesis (Aronoff 1994)

▪ Why would LFG want DM, though?

   o The FIRST big question of the day.

   o Formal linguistic theory is overwhelmingly sorted into "camps" of shared assumptions and hypotheses.

      ▪ Realizational morphologists who work with LFG tend to be those that subscribe to the word-based approach.

         • This makes intuitive sense because LFG takes as one of its core assumptions the Lexical Integrity Hypothesis (Aronoff 1976, DiScullo & Williams 1985, LaPointe 1980, Bresnan et al 2016).

      ▪ HOWEVER, in the context of a realizational MSI, You can give up the strictest version of the LIH without giving up much of the motivation for the LIH
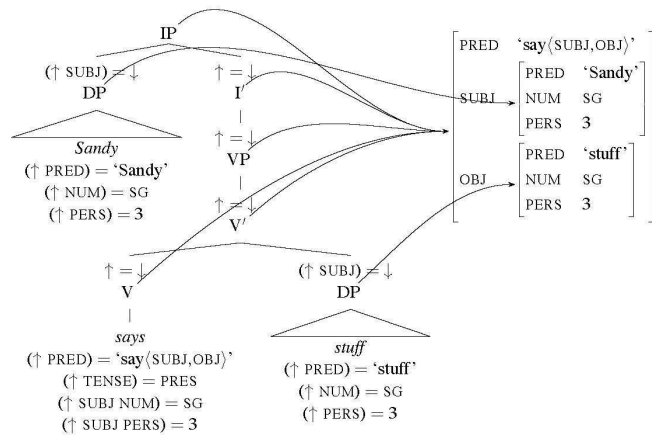
- What does LFG gain by countenancing two alternative theories of the MSI?

  - LFG-internal reason:

    - Extant realizational LFG models typically assume that f-structure is the syntactic level that interfaces with morphology (via m-structure).

      - HOWEVER, f-structure does not make the right structural distinctions to facilitate realization of forms.

        - e.g. agreement information is contributed to the f-structure from many places in the c-structure.

        - WHERE it comes from in the c-structure is important to the morphology, not just that it is there
          - (which is all the f-structure knows).

    - Does not require m-structure

      - Outside of the morpheme hypothesis, m-structure is unparsimonious.

  - LFG-external reason:

    - The LFG "camp" excludes a bunch of morphologists or morphosyntacticians who might otherwise be interested in LFG:

      - Supporters of the morpheme-based hypothesis.

      - Supporters of the Mirror Principle.

        - This is significant because early standard incremental LFG is really good at the Mirror Principle

      - Practitioners seeking syntactocentric explanation.

- Why would DM want LFG?

  - The OTHER big question of the day!

  - DM doesn't have to be married to MPP.

    - There is no reason to exclude the many syntacticians who reject tenets of MPP but would not necessarily reject the tenets of DM.

  - DM doesn't need to be derivational:

    - Declarative models have their appeal!

      - Almost every other model of syntax other than MPP.

      - A syntactic model that is more surface-true is in some ways easier to interface with a realizational model.

      - DM is in the midst of suffering from a glut of post syntactic operations

        - many of which can be alleviated by a spanning account for morpho-syntax (Ramchand 2008, Svenonius 2009, Merchant 2013, Haugen & Siddiqi 2016)

          - Spanning approaches are inherently *declarative* (which makes them compatible with declarative models such as LFG).

        - Many others can be alleviated by OT accounts of morphophonology. (Bye & Svenonius 2014, Haugen 2013, Haugen & Siddiqi 2016)

          - OT accounts are also fundamentally *declarative!*

      - These alternative models of DM that are inherently declarative largely aim to account for particular classes of phenomena:

        - Long distance morphophonological relationships.
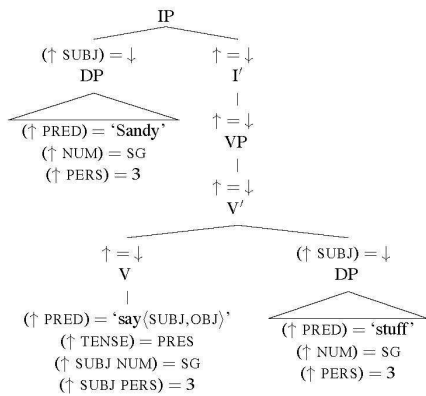   - Stem Allomorphy.
        - Root Suppletion.

2.   Distributed Lexical Functional Grammar (DLFG)

- Step 1: Abandon the hypothesis that C-structure contains phonological strings at all.

- Step 2: Assume that C-structure contains f-descriptions that are NOT linked to words.

  - Not a novel proposal.

  - Asudeh et al 2013, Asudeh & Toivonen 2014

    - F-descriptions for "constructional" meanings are introduced directly into the c-structure (which means constructions do not have to be adopted as theoretical primitives).

- Step 3: **New Proposal**:

  - All there is in the c-structure is these independent f-descriptions.

  - No "words"

  - Depending on your model of morphology, you might also assume:

    - Lexical Identifier
    - Meaning Constructor

  - (↑PRED) = 'Sandy', etc., is minimally sufficient though.

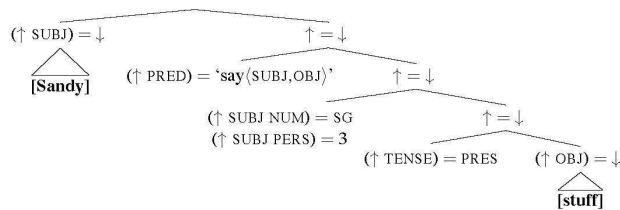1) Traditional LFG c-structure and f-structure for "Sandy says stuff"

$$
\begin{bmatrix}
\text{PRED} & \text{'say}\langle\text{SUBJ,OBJ}\rangle\text{'} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Sandy'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \end{bmatrix} \\
\text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'stuff'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \end{bmatrix}
\end{bmatrix}
$$

c-structure tree for "Sandy says stuff":
- IP
  - DP ($\uparrow$ SUBJ) $=\downarrow$
    - Sandy — ($\uparrow$ PRED) = 'Sandy', ($\uparrow$ NUM) = SG, ($\uparrow$ PERS) = 3
  - I' ($\uparrow = \downarrow$)
    - VP ($\uparrow = \downarrow$)
      - V' ($\uparrow = \downarrow$)
        - V ($\uparrow = \downarrow$)
          - says — ($\uparrow$ PRED) = 'say$\langle$SUBJ,OBJ$\rangle$', ($\uparrow$ TENSE) = PRES, ($\uparrow$ SUBJ NUM) = SG, ($\uparrow$ SUBJ PERS) = 3
        - DP ($\uparrow$ SUBJ) $=\downarrow$
          - stuff — ($\uparrow$ PRED) = 'stuff', ($\uparrow$ NUM) = SG, ($\uparrow$ PERS) = 3

2) Revised c-structure for "Sandy says stuff".

c-structure tree:
- IP
  - DP ($\uparrow$ SUBJ) $=\downarrow$
    - ($\uparrow$ PRED) = 'Sandy', ($\uparrow$ NUM) = SG, ($\uparrow$ PERS) = 3
  - I' ($\uparrow = \downarrow$)
    - VP ($\uparrow = \downarrow$)
      - V' ($\uparrow = \downarrow$)
        - V ($\uparrow = \downarrow$)
          - ($\uparrow$ PRED) = 'say$\langle$SUBJ,OBJ$\rangle$', ($\uparrow$ TENSE) = PRES, ($\uparrow$ SUBJ NUM) = SG, ($\uparrow$ SUBJ PERS) = 3
        - DP ($\uparrow$ SUBJ) $=\downarrow$
          - ($\uparrow$ PRED) = 'stuff', ($\uparrow$ NUM) = SG, ($\uparrow$ PERS) = 3

- This hypothesis enables ANY realizational model to be read directly off the c-structure.
  - Avoids the f-structure mapping problem.
- To get a morpheme-based model, we need to split off the individual f-descriptions into smaller "morpheme" sized nodes.

3) DLFG c-structure "Sandy says stuff"

c-structure tree:
- ($\uparrow$ SUBJ) $=\downarrow$
  - [Sandy]
- ($\uparrow = \downarrow$)
  - ($\uparrow$ PRED) = 'say$\langle$SUBJ,OBJ$\rangle$'
  - ($\uparrow = \downarrow$)
    - ($\uparrow$ SUBJ NUM) = SG, ($\uparrow$ SUBJ PERS) = 3
    - ($\uparrow = \downarrow$)
      - ($\uparrow$ TENSE) = PRES
      - ($\uparrow$ OBJ) $=\downarrow$
        - [stuff]

- Without further architecture this model inherits all DM's problems of dealing with words with stem allomorphy, portmanteau, etc, that require derivational mechanisms.

  - We need to enable a spanning account.

  - Here we adopt a post-linearization spanning account for the reasons given in Haugen & Siddiqi (2016) and Merchant (2013).

- Step 4: **New Proposal**:

  - "Flattening" operates on c-structure and maps c-structure to an ordered list of (sets of) the terminal f-descriptions.

    - Flattening is just the standard tree-theoretic operation of taking the yield of the tree, where the yield is the information in the terminal nodes, preserving their order in the tree.

    - So the yield of a standard vanilla phrase-structure tree would be the string that the tree parses.
    - In our case, as the terminals are sets of f-descriptions, the yield that results from the FLATTEN function is not a string, but an order-preserving list of sets of f-descriptions

  - Stating constraints and operations on the yield of a tree in LFG has been independently motivated prior to this proposal. See Asudeh (2009).

  - Predicts realizational forms can be dependent on linear precedence, not (necessarily) hierarchical relations.

    - English Nominative Case

---

4) English nominative case.

     a. Me and Jack went to the store.
     b. *I and Jack went to the store.
     c. Jack and me went to the store.
     d. Jack and I went to the store.

     e. Us linguists hate pedants.
     f. %We linguists hate pedants.

     g. Me kicking a desk shocked the audience.
     h. *I kicking a desk shocked the audience

- GENERALIZATION: English nominative case is sensitive to linear precedence (in additional structural conditions; adverbs present the usual problems).

  - Realizational morphology can't be sensitive to this type of condition unless it acts on a linearized or flattened structure.

- Step 5: **New Proposal**:

  - REALIZE! Maps flattened structure to phonological forms.

5) "Sandy says stuff".
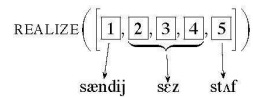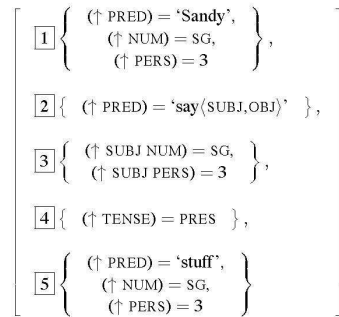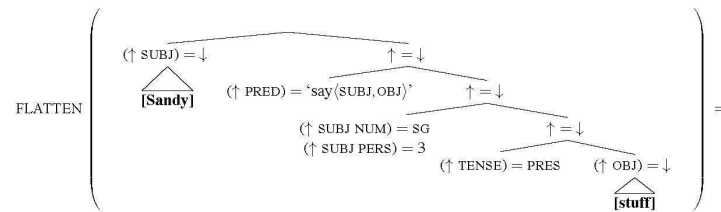
- This mapping is subject to constraints.

  - Will be familiar to practitioners of DM or Nanosyntax

6) Constraints on realization (i.e., the function REALIZE).

  a. Use the fewest listemes you can for the job. (Minimize Exponence: Siddiqi 2009)
  b. Use the listeme that expones the most amount of information in the X^0 it expones. (Subset Principle, Halle & Marantz 1994)
  c. REALIZE may expone multiple adjacent X^0s provided that the f-descriptions exponed by the inserted listeme are as large a subset of the string of adjacent X^0s than that which could otherwise be expressed by separate listemes at the contained X^0s. (Post-linearization Spanning; Haugen & Siddiqi 2016)

3. Stem Allomorphy in DLFG

Compare derivation of "Sandy says stuff" with "Sandy kicks stuff"

7) Sandy kicks stuff.
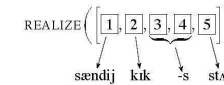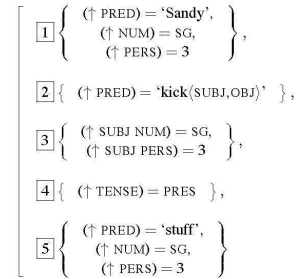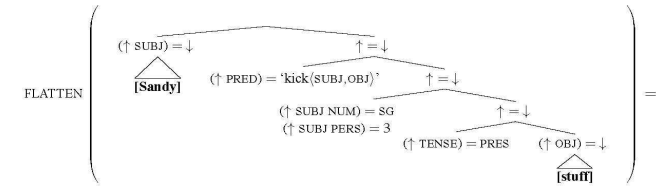
- *Says* expresses 2, 3, and 4 because *says* is a suppletive portmanteau form of [say+s] (despite its transparent spelling)

- *Kick* expresses 2 while *–s* expresses 3 and 4 because while *–s* is a portmanteau, *kick* is the regular expression of 2.

- In this way, DLFG accounts for stem allomorphy with a listing account but accounts for regular morphology with a morpheme-based account.

  - See Haugen & Siddiqi (2016) for details

- Hybrid model.

  - Strengths of both morpheme-based model and word-based models.

4. Conclusions

- Argued that DM need not and ought not exclude LFG.
- Argued that LFG need not and ought not exclude DM.
- Argued that both benefit from bridge-building (i.e. a DM interface with LFG)
- Proposed a possible architecture for such a DM-LFG interface.
- Showed the strengths of this model (hybrid word/morphology declarative model)

References available upon request as a separate attachment.