Outline – Part 1

Parataxis
- Very small bound on degree of paratactic embedding (PE°)
  - Langendoen 1998
- Right- and left-embedding hypotaxis (R/LE)
  - Readjustment to pseudo-paratactic (ΨP) structure
    - Langendoen 1975, updated with traces
- Center-embedding hypotaxis (CE)
  - Motivating concept of zigzag embedding (ZE)
  - Finite-state modeling of bounded degree of ZE
Polysyndeton (more than one coordinator)

1. He is not quite journalist or carnival barker or orator or interlocutor or master of ceremonies or trained seal.
   • American Publishing House for the Blind (APHB) corpus

2. It's great to see millions on this earth who had nothing but a record of sadness and poverty and misery and hunger and disease have the chance to go up.
   • APHB corpus
Monosyndeton (exactly one coordinator)

1. It provides him with a ... method of getting a hearing ... in any federal income, gift or estate tax dispute ....
   • (APHB corpus)

2. Courts often need ... to settle manslaughter charges, inheritance claims, insurance proceeds, tax problems, and the disposition of jointly held money and property.
   • (APHB corpus)
Asyndeton (no coordinator)

1. I felt exposed, unprotected, somehow afraid of what might happen.
   • SUSANNE corpus

2. The thistle, the nettle, the burdock, the belladonna / Have a future....
Nesting of paratactic structure -- not discussed further

• ... the purified soybean meal is \(_{\text{CAP0}}\) colored, flavored, pressed, shaped and cut into bits that \(_{\text{CVP1}}\) \(_{\text{CV2}}\) look and taste] like \(_{\text{CNP2}}\) bacon \(_{\text{CNP3}}\) chips or strips], pork sausage, ground beef, sliced ham or chicken] and are \(_{\text{CAP2}}\) cheaper and just as nourishing as the real thing].

– APHB corpus
Monosyndetonstons have ‘flat’ structure

income, gift or estate

income, gift or estate
So do asyndetons

- the thistle
- the nettle
- the burdock
- the belladonna
And so *can* polysyndetons

- sadness
- and poverty
- and misery
- and hunger
- and disease
But sometimes polysyndetons show paratactic embedding (PE)

- peel the plums
  - and halve them
    - or slice them
- and halve them
  - or slice them
Factors resulting in PE

A. Members introduced by different connectives.
B. Members introduced by different junctures -- not considered further here.
C. A member introduced by a connective followed by one that is not.
D. Lexical, pragmatic or semantic considerations.
Other examples of Factor A

1. Caroline was going into the possibility of the Pope or his priests and the nuns.
   • APHB corpus
2. The bush babies cluster together and groom each other, or run through the trees in gangs.
   • APHB corpus
Examples of Factor C

1. ... an employer has a right to refuse to hire a man if he doesn’t like the color of his tie, or his diction, his shifty eyes, or his having taken the Fifth Amendment.
   • APHB corpus

2. ... *The Travels of Marco Polo* became an indispensable book for explorers and geographers, historians and mapmakers and the delight of all who travel or dream of doing so.
   • APHB corpus
One PE structure resulting from Factor C

explorers and geographers, historians and mapmakers

explorers and geographers

historians

and mapmakers

explorers

and geographers
Examples of Factor D

1. That’s our job and that’s the job of Brezhnev and Kosygin and Mao Tse-tung and Chou En-lai.
   • APHB corpus

2. Then Dr. White and his faculty and students could assemble and throw rocks at each other and play with matches and burn things down.
   • APHB corpus
One PE structure resulting from Factor D

Brezhnev and Kosygin

Brezhnev

and Kosygin

and Mao Tse-Tung

and Chou En-lai

and Mao Tse-tung

and Chou En-lai
Combination of Factors

1. a company of persons gathered for deliberation and legislation, worship or entertainment
   • Webster’s Collegiate Dictionary 7th ed., definition of assembly

2. The girls are brought in frightened or defiant, cursing and fighting or sullen and withdrawn.
   • APHB corpus

3. Combine grapefruit with bananas, strawberries and bananas, bananas and melon balls, raspberries or strawberries and melon balls, seedless white grapes and melon balls, or pineapple cubes and orange slices.
   • The James Beard Cookbook, recipe for grapefruit salad
One PE structure resulting from Factors A and C

- frightened or defiant, cursing and fighting or sullen and withdrawn
  - frightened
  - or defiant, cursing and fighting
  - or sullen and withdrawn
    - or sullen
    - and withdrawn
Another PE structure resulting from Factors A and C

- Frightened or defiant, cursing and fighting or sullen and withdrawn
  - Frightened or defiant
  - Cursing and fighting
  - Or sullen and withdrawn
    - Or sullen
    - And withdrawn
PE° > 2 is hardly ever attested

frightened or defiant, cursing and fighting or sullen and withdrawn

frightened

or defiant, cursing and fighting or sullen and withdrawn

or defiant

cursing

and fighting or sullen and withdrawn

and fighting

or sullen and withdrawn

or sullen

and withdrawn
Outline – Part 2

• Parataxis
  – Very small bound on degree of paratactic embedding (PE°)
    • Langendoen 1998

➢ Right- and left-embedding hypotaxis (R/LE)
  – Readjustment to pseudo-paratactic (ΨP) structure
    • Langendoen 1975 updated with traces

• Center-embedding hypotaxis (CE)
  – Motivating concept of zigzag embedding (ZE)
  – Finite-state modeling of bounded degree of ZE
A right-embedding (RE) hypotactic structure

we recalled that somebody said that everybody knew that we were engaged

that somebody said that everybody knew that we were engaged

that everybody knew that we were engaged
Equivalent pseudo-paratactic ($\Psi P$) structure

we recalled that somebody said that everybody knew that we were engaged

S1 = that somebody said $t_{s_2}$
S2 = that everybody knew $t_{s_3}$
S3 = that we were engaged
A left-embedding (LE) hypotactic structure

my younger brother’s closest friend’s college algebra teacher’s hidden agenda

hidden agenda

my younger brother’s closest friend’s college algebra teacher’s

my younger brother’s closest friend’s

college algebra teacher’s

my younger brother’s closest friend’s

closest friend’s
Equivalent $\Psi PS$

- $N_1 = \text{my younger brother's}$
- $N_2 = t_{N_1} \text{ closest friend's}$
- $N_3 = t_{N_2} \text{ college algebra teacher's hidden agenda}$
- $t_{N_3} \text{ hidden agenda}$
Outline – Part 3

• Parataxis
  – Very small bound on degree of paratactic embedding (PE°)
    • Langendoen 1998

• Right- and left-embedding hypotaxis (R/LE)
  – Readjustment to pseudo-paratactic (ΨP) structure
    • Langendoen 1975 updated with traces

➢ Center-embedding hypotaxis (CE)
  – Motivating concept of zigzag embedding (ZE)
  – Finite-state modeling of bounded degree of ZE
Adverbiaisal attachment to an RE structure results in CE

• The sentence Somebody said that everybody knew that we got engaged last December is structurally ambiguous, depending on which clause last December attaches to:
  – High: somebody said that ...
  – Low: we got engaged
  – Middle: everybody knew that ...

• Each structure exhibits $\text{CE}^\circ = 1$. 
High attachment (L then R)

somebody said that everybody knew that we got engaged last December

somebody said that everybody knew that we got engaged

last December

everybody knew that we got engaged

everybody knew that we got engaged
Low attachment (R then L)

somebody said that everybody knew that we got engaged last December

everybody knew that we got engaged last December

we got engaged last December
Middle attachment (R then L then R)

somebody said that everybody knew that we got engaged last December

dsomebody said that everybody knew that

everybody knew that we got engaged last December

everybody knew that we got engaged

everybody knew that

last December

we got engaged
Preference for high and low to middle attachment

• High and low attachments are preferred to middle attachment.

• Degree of zigzag embedding ($ZE^\circ$) correlates with these preferences.
  – High and low attachment structures have $ZE^\circ = 1$.
  – Middle attachment structures have $ZE^\circ = 2$. 
ΨP structure for high attachment

somebody said that everybody knew that we got engaged last December

S1 = everybody knew that S2 = we got engaged

somebody said that tS1

S1 = everybody knew that tS2

last December

May 11, 2007 Maryland Mayfest
somebody said that everybody knew that we got engaged last December

\( \Psi P \) structure for low attachment

somebody said that \( t_{S_1} \)

\( S_1 = \) everybody knew that \( t_{S_2} \)

\( S_2 = \) we got engaged last December

we got engaged

last December
somebody said that everybody knew that we got engaged last December

somebody said that $t_{S1}$

S1 = everybody knew that we got engaged last December

everybody knew that we got engaged

S2 = we got engaged

everybody knew that $t_{S2}$
Significance of ZE

• ZE is a more sensitive measure of structural complexity than CE.  
  \[ CE^\circ \leq ZE^\circ \leq 2*CE^\circ \]

• ZE° is preserved under mapping to \( \Psi P \) structure.
How regular natural languages might grow onto- or phylo-genetically

1. Start with nonembedding structures, including paratactic ones.
2. Next, recognize or build \( \Psi P \) structures rather than R/LE structures.
3. Next, recognize or build \( \Psi P \) structures with \( ZE^\circ = 1 \).
4. Continue with \( ZE^\circ = 2 \), etc. as needed.
Conclusion

• The procedure approximates but never achieves the coverage of a CE phrase-structure grammar.

• The upper bound on $\text{ZE}^\circ$ (and hence also on $\text{CE}^\circ$) arises organically from the construction. It is not an arbitrarily imposed restriction.