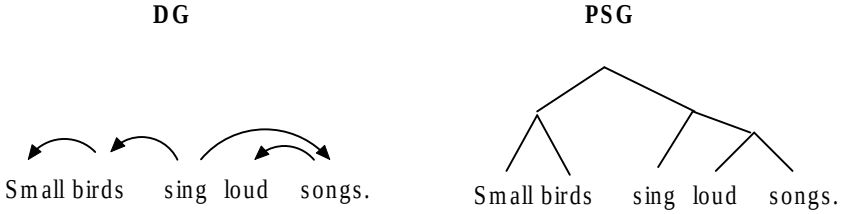
ESSLLI 2000: Dependency Grammar 1 - Overview

1. How DG is different from PSG

Grammars, and theories of grammar, can be classified according to whether the basic unit of sentence structure is:

- The phrase (PSG)
- The dependency between two words (DG).



Each approach implies the other:

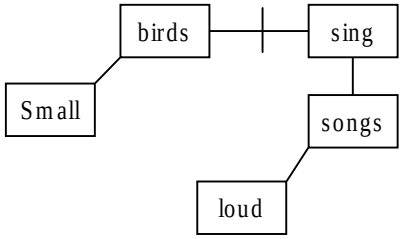
- PSG implies inter-word dependencies (but only if one word is chosen as the phrasal head - this is an optional extra, chosen in X-bar syntax).
- DG implies phrases (a word plus its dependents and their phrases constitutes a phrase).

So one important question is whether the difference is significant - are they just notational variants?

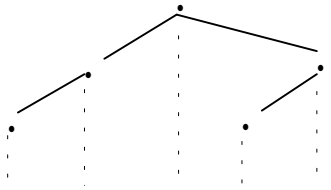
2. The history of DG and PSG

Most of the history of linguistics is the history of DG, with PSG invoked only for subordinate clauses - e.g. the **noun phrase** is a discovery/invention of the 20th century:

- Panini (2600 years ago, India) recognised, distinguished and classified semantic, syntactic and morphological dependencies - see Bharati, *Natural Language Processing*
- The Arabic grammarians (1200 years ago, Iraq) recognised government and syntactic dependency structure, and the grammarians of Basra and Kufa almost declared war on each other over the possibility of mutual dependency! - Owens, *The Foundations of Grammar.*
- The Latin grammarians (800 years ago) recognised 'determination' and dependency structures. - Percival, "Reflections on the History of Dependency Notions"
- School grammars of English in Europe and USA taught sentence-analysis in terms of dependency, and the 'sentence diagramming' which has been popular since the late 19th century (using a system invented in USA) is DG:



- Lucien Tesnie`re (1930s, France) developed a relatively formal and sophisticated theory of DG grammar for use in **schools**. - Tesniére, *Éléments de syntaxe structurale,* first drafted in 1939 but published only in 1959, after his death. His structure diagram is called a 'stemma' and is widely used in DG literature.

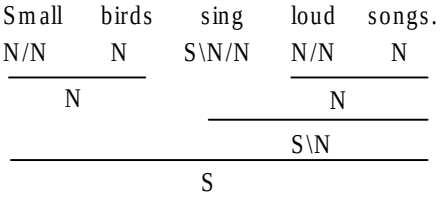


Small birds sing loud songs.

This **bottom-up** approach is still widely used in **Europe** (by linguists in Germany, France, Scandinavia, Czechoslovakia, Russia), and by Russians and slavists in USA (Mel'cuk, Shaumyan, Nichols).

BUT in 1930 Leonard Bloomfield in the **USA** developed a **top-down** approach: Immediate-Constituent Analysis, which turned into PSG. (This was largely inspired by the German psychologist Wundt - Percival, "On the Historical Source of Immediate Constituent Analysis".)

Meanwhile, in Europe, logicians (notably Ajdukiewicz, 1935) were developing **Categorial Grammar**, which is often classified as a kind of DG although the words combine to make phrases just as in PSG.



This was being taught in Harvard by Jehoshua Bar-Hillel, who had a lot of influence on Chomsky - e.g. he wrote the first 'generative grammar' of a natural language (Newmeyer, *Linguistic Theory in America*:36*).* In this sense, Chomsky 'studied DG' with him, and decided it was just a notational variant of PSG. He may never have considered any 'pure' version of DG.

3. Is DG just a notational variant of PSG?

A number of logicians, including Bar-Hillel, proved that DG (including Categorial Grammar) is **weakly** equivalent to a context-free PSG (Gaifman, "Dependency systems and phrase-structure systems"). This result is generally accepted.

But it is **not** a notational variant of PSG because it is not **strongly** equivalent - i.e. it does not allow the same analyses:

- Phrases are implicit, not explicit, so
  - phrases cannot be classified separately from their heads.
- Relations are explicit, not implicit, so
  - relations can be classified and labelled.
- All phrases must be endocentric, so
  - apparently exocentric constructions such as gerunds ('an NP with the internal structure of a clause') are a fundamental challenge.
- No non-terminal nodes are allowed, so DG does not allow:
  - 'unary' branching (e.g. NP consisting of just N),
  - VP (contrasting with V).

4. The main theories in the DG family

DG, like PSG, is just one element in a theory of sentence structure. A wide range of more comprehensive theories include DG rather than PSG, and vary along much the same lines as theories which assume PSG:

- Some reject transformations while others accept them.
- Some recognise a single level of syntax, while others disperse syntactic phenomena over a range of different levels of structure which map onto one another more or less freely.
- Some are sufficiently formalised to be used in computer systems, while others are relatively informal.
- Some insist on 'projectivity' (each word in a stemma 'projects' directly to its node, without crossing the projection line of any other word - i.e. phrases must be continuous), while others don't.

Here is an incomplete list of DG-based theories (in alphabetic order), with the names of their main proponents:

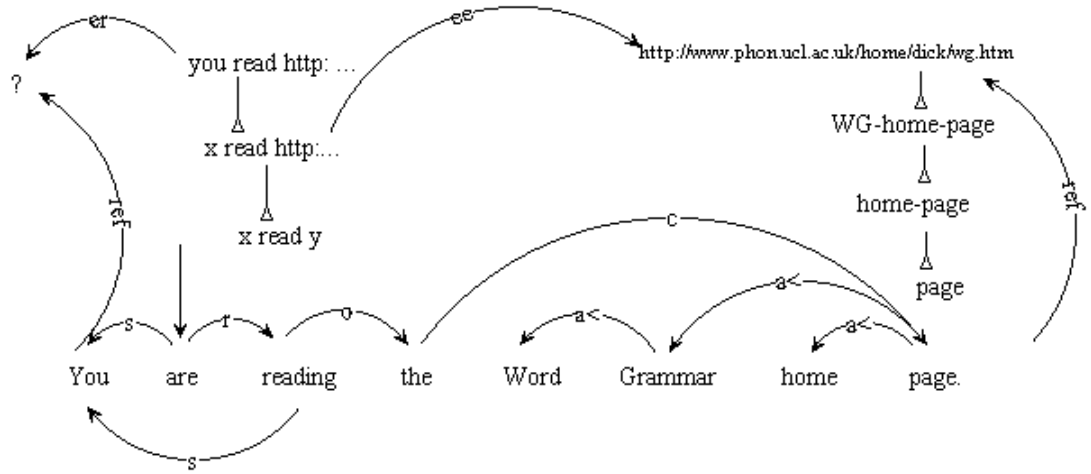
- Case Grammar (Anderson)
- Daughter-Dependency Theory (Hudson)
- Dependency Unification Grammar (Hellwig)
- Functional-Generative Description (Sgall)
- Lexicase (Starosta)
- Meaning-Text Model (Mel'cuk)
- Metataxis (Schubert)
- Unification Dependency Grammar (Maxwell)
- Constraint Dependency Grammar (Maruyama)

5. Word Grammar

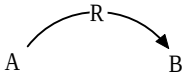
This is a theory which I have been developing since the early 1980s; see separate bibliography. It is firmly based on DG, and since I believe it has the best possible combination of other features this is the theory that I shall use to illustrate the (many) strengths and (very few!) weaknesses of DG.

The other main features of WG:

- It is **monostratal** - there is only one syntactic structure per sentence, paired with a semantic structure and a phonological one.
- It is **enriched** - it allows multiple dependencies - e.g. *You* depends on two words ('has two parents') in *You are reading*.
- It generalises by means of **default inheritance** based on the 'isa' relationship.
- It allows **labelled relationships** (in an isa hierarchy).
- It is **non-modular** and **cognitive**: language is an area of the general **network** of knowledge.



5. Notation: how to read a network



B depends on A in relation R
i.e. B is A's R

B is a A; i.e. B is either
an example or a subclass
of A