A Model-Theoretic Framework for Grammaticality Judgements

Denys Duchier Jean-Philippe Prost Thi-Bich-Hanh Dao

LIFO, Université d'Orléans

Formal Grammars, 2009

Foreword

- ungrammatical utterances are an everyday phenomenon
- some utterances are more ungrammatical than others
- JP Prost's PhD thesis [2008]

contributions:

- model-theoretic semantics for property grammars
- loose models for quasi-expressions
- scoring functions for comparative judgements of admissibility

Outline

Sentences of decreasing acceptability

1. Les employés ont rendu un rapport très complet à leur employeur [100%]
   The employees have sent a report very complete to their employer
2. Les employés ont rendu rapport très complet à leur employeur [92.5%]
   The employees have sent report very complete to their employer
3. Les employés ont rendu un rapport très complet à [67.5%]
   The employees have sent a report very complete to
4. Les employés un rapport très complet à leur employeur [32.5%]
   The employees a report very complete to their employer
5. Les employés un rapport très complet à [5%]
   The employees a report very complete to their employer

Gradience

We are interested in two questions: given an expression or a quasi-expression:

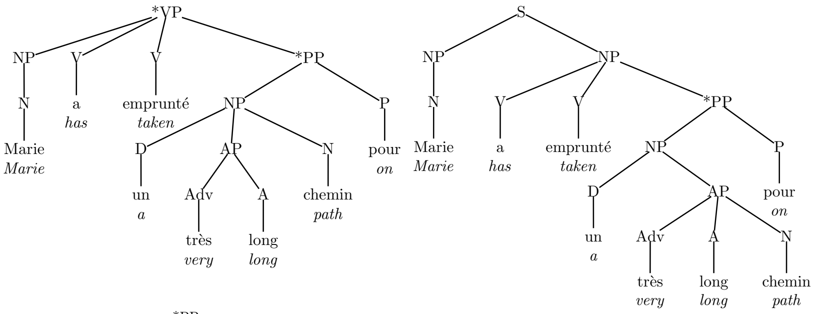
- what is the best (quasi-)analysis for it?
- how grammatical is it?

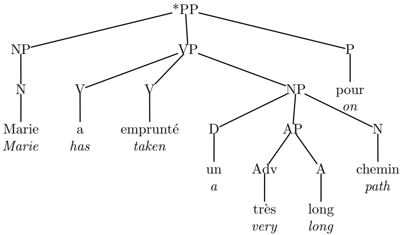
Bas Aarts [2007]:

- intersective gradience (classification)
- subsective gradience (prototypicality)

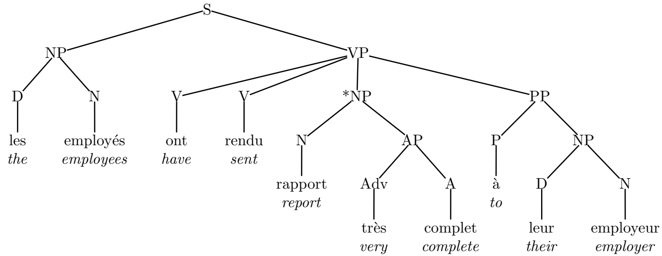
Examples: bat, pinguin

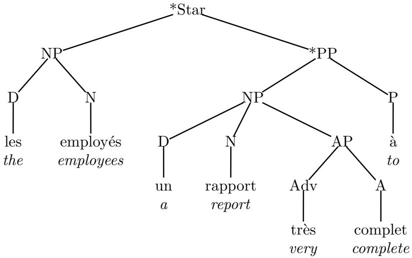
Possible models for a quasi-expression





Models for related quasi-expressions





(some) Formal options

GES/MTS (Pullum&Scholtz [2001], Pullum [2007])

- GES: ill-suited
- MTS:
  - grammar = constraint
  - defined in terms of satisfaction (open to violations)
  - compatible with degrees of ungrammaticality

OT (Prince&Smolensky [1993])

- grammaticality = optimality
- cannot distinguish between expressions and quasi-expressions

Property Grammars

Property Grammars are the transposition of phrase structure grammars from the GES perspective into the MTS perspective

Production rules as constraints

$$\texttt{NP} \rightarrow \texttt{D N}$$

- GES: rewrite rule
- MTS: constraint
  - satisfied in a tree iff satisfied at every node
  - satisfied at a node iff: **either** the node is not labeled with NP, **or** it has exactly 2 children, the 1st labeled with D, the 2nd labeled with N

Model-theoretic semantics for CFG

A CFG is a set of production rules (1 per non-terminal; use alternation where necessary)

- class of models: trees labeled with categories
- a tree is a model of the grammar iff every rule is satisfied at every node
- $\alpha \to \beta_1 \ldots \beta_n$ is satisfied at a node iff: either the node does not have category α, or it has a sequence of exactly n children labeled respectively β_1 through β_n

Coarse-grained constraints

$$\texttt{NP} \longrightarrow \texttt{D N}$$

For a NP there must be:

(1) a D child
(2) only one
(3) a N child
(4) only one
(5) nothing else
(6) the D child must precede the N child

Properties

obligation	$A : \triangle B$	at least one B child
uniqueness	$A : B!$	at most one B child
linearity	$A : B \prec C$	a B child precedes a C child
requirement	$A : B \Rightarrow C$	if there is a B child, then also a C child
exclusion	$A : B \not\Leftrightarrow C$	B and C children are mutually exclusive
constituency	$A : S?$	the category of any child must be one in S

Fine-grained constraints

$$\texttt{NP} \rightarrow \texttt{D N}$$

becomes:

(1) $\texttt{NP} : \triangle\texttt{D}$ (a D child)

(2) $\texttt{NP} : \texttt{D}!$ (only one)

(3) $\texttt{NP} : \triangle\texttt{N}$ (a N child)

(4) $\texttt{NP} : \texttt{N}!$ (only one)

(5) $\texttt{NP} : \{\texttt{D}, \texttt{N}\}?$ (nothing else)

(6) $\texttt{NP} : \texttt{D} \prec \texttt{N}$ (the D child must precede the N child)

these can be independently violated

Property Grammar for French

S (Utterance)

obligation	: $\triangle$VP
uniqueness	: NP!
	: VP!
linearity	: NP $\prec$ VP
dependency	: NP $\rightsquigarrow$ VP

AP (Adjective Phrase)

obligation	: $\triangle$(A $\vee$ $V_{[\text{past part}]}$)
uniqueness	: A!
	: $V_{[\text{past part}]}$!
	: Adv!
linearity	: A $\prec$ PP
	: Adv $\prec$ A
exclusion	: A $\not\Leftrightarrow$ $V_{[\text{past part}]}$

PP (Propositional Phrase)

obligation	: $\triangle$P
uniqueness	: P!
	: NP!
linearity	: P $\prec$ NP
	: P $\prec$ VP
requirement	: P $\Rightarrow$ NP
dependency	: P $\rightsquigarrow$ NP

NP (Noun Phrase)

obligation	: $\triangle$(N $\vee$ Pro)
uniqueness	: D!
	: N!
	: PP!
	: Pro!
linearity	: D $\prec$ N
	: D $\prec$ Pro
	: D $\prec$ AP
	: N $\prec$ PP
requirement	: N $\Rightarrow$ D
	: AP $\Rightarrow$ N
exclusion	: N $\not\Leftrightarrow$ Pro
dependency	: $N_{\begin{bmatrix}\text{gend} & \boxed{1} \\ \text{num} & \boxed{2}\end{bmatrix}} \rightsquigarrow D_{\begin{bmatrix}\text{gend} & \boxed{1} \\ \text{num} & \boxed{2}\end{bmatrix}}$

VP (Verb Phrase)

obligation	: $\triangle$V
uniqueness	: $V_{[\text{main past part}]}$!
	: NP!
	: PP!
linearity	: V $\prec$ NP
	: V $\prec$ Adv
	: V $\prec$ PP
requirement	: $V_{[\text{past part}]} \Rightarrow V_{[\text{aux}]}$
exclusion	: $\text{Pro}_{[\text{acc}]} \not\Leftrightarrow$ NP
	: $\text{Pro}_{[\text{dat}]} \not\Leftrightarrow \text{Pro}_{[\text{acc}]}$
dependency	: $V_{\begin{bmatrix}\text{pers} & \boxed{1} \\ \text{num} & \boxed{2}\end{bmatrix}} \rightsquigarrow \text{Pro}_{\begin{bmatrix}\text{type} & \text{pers} \\ \text{case} & \text{nom} \\ \text{pers} & \boxed{1} \\ \text{num} & \boxed{2}\end{bmatrix}}$

Formal definition of property grammars

$\mathcal{L}$ a finite set of labels, $\mathcal{S}$ a finite set of strings

$$\begin{aligned}\mathcal{P}_{\mathcal{L}} = \quad & \{c_0 : c_1 \prec c_2, \\ & c_0 : \triangle c_1, \\ & c_0 : c_1 !, \\ & c_0 : c_1 \Rightarrow c_2, \\ & c_0 : c_1 \not\Leftrightarrow c_2, \\ & c_0 : s_1 ? \mid \forall c_0, c_1, c_2 \in \mathcal{L},\ \forall s_1 \subseteq \mathcal{L}\}\end{aligned}$$

Property grammar

$$G = (P_G, L_G) \qquad P_G \subseteq \mathcal{P}_{\mathcal{L}} \qquad L_G \subseteq \mathcal{L} \times \mathcal{S}$$

syntax tree $\tau = (D_\tau, L_\tau, R_\tau)$

- tree domain D_τ
- labeling function $L_\tau : D_\tau \to \mathcal{L}$
- realization function $R_\tau : D_\tau \to \mathcal{S}^*$

tree domain

a finite subset of $\mathbb{N}_0^*$ closed for prefixes and for left-siblings, where $\mathbb{N}_0 = \mathbb{N} \setminus \{0\}$

arity

$$A_\tau(\pi) = \max \{0\} \cup \{i \in \mathbb{N}_0 \mid \pi i \in D_\tau\}$$

Instances of Properties

Every property in P_G must be checked at every node in D_τ and for all possible choices among its children.

$$\mathcal{I}_\tau[\![c_0 : c_1 \prec c_2]\!] = \{(c_0 : c_1 \prec c_2)@\langle \pi, \pi i, \pi j\rangle \mid \forall \pi, \pi i, \pi j \in D_\tau,\ \ i \neq j\}$$

Instances of Properties

Every property in P_G must be checked at every node in D_τ and for all possible choices among its children.

$$\begin{aligned}
\mathcal{I}_\tau[\![G]\!] &= \cup\{\mathcal{I}_\tau[\![p]\!] \mid \forall p \in P_G\} \\
\mathcal{I}_\tau[\![c_0 : c_1 \prec c_2]\!] &= \{(c_0 : c_1 \prec c_2)@\langle\pi, \pi i, \pi j\rangle \mid \forall \pi, \pi i, \pi j \in D_\tau,\ \ i \neq j\} \\
\mathcal{I}_\tau[\![c_0 : \triangle c_1]\!] &= \{(c_0 : \triangle c_1)@\langle\pi\rangle \mid \forall \pi \in D_\tau\} \\
\mathcal{I}_\tau[\![c_0 : c_1!]\!] &= \{(c_0 : c_1!)@\langle\pi, \pi i, \pi j\rangle \mid \forall \pi, \pi i, \pi j \in D_\tau,\ \ i \neq j\} \\
\mathcal{I}_\tau[\![c_0 : c_1 \Rightarrow s_2]\!] &= \{(c_0 : c_1 \Rightarrow s_2)@\langle\pi, \pi i, \pi j\rangle \mid \forall \pi, \pi i, \pi j \in D_\tau,\ \ i \neq j\} \\
\mathcal{I}_\tau[\![c_0 : c_1 \not\Leftrightarrow c_2]\!] &= \{(c_0 : c_1 \not\Leftrightarrow c_2)@\langle\pi, \pi i, \pi j\rangle \mid \forall \pi, \pi i, \pi j \in D_\tau,\ \ i \neq j\} \\
\mathcal{I}_\tau[\![c_0 : s_1?]\!] &= \{(c_0 : s_1?)@\langle\pi, \pi i\rangle \mid \forall \pi, \pi i \in D_\tau\}
\end{aligned}$$

Pertinence

$$P_\tau((c_0 : c_1 \prec c_2)@\langle \pi, \pi i, \pi j \rangle) \quad \equiv \quad L_\tau(\pi) = c_0 \ \wedge \ L_\tau(\pi i) = c_1 \ \wedge \ L_\tau(\pi j) = c_2$$

Pertinence

$$\begin{aligned}
P_\tau((c_0 : c_1 \prec c_2)@\langle \pi, \pi i, \pi j\rangle) &\equiv L_\tau(\pi) = c_0 \ \wedge\ L_\tau(\pi i) = c_1 \ \wedge\ L_\tau(\pi j) = c_2 \\
P_\tau((c_0 : \triangle c_1)@\langle \pi\rangle) &\equiv L_\tau(\pi) = c_0 \\
P_\tau((c_0 : c_1 !)@\langle \pi, \pi i, \pi j\rangle) &\equiv L_\tau(\pi) = c_0 \ \wedge\ L_\tau(\pi i) = c_1 \ \wedge\ L_\tau(\pi j) = c_1 \\
P_\tau((c_0 : c_1 \Rightarrow s_2)@\langle \pi, \pi i, \pi j\rangle) &\equiv L_\tau(\pi) = c_0 \ \wedge\ L_\tau(\pi i) = c_1 \\
P_\tau((c_0 : c_1 \not\Leftrightarrow c_2)@\langle \pi, \pi i, \pi j\rangle) &\equiv L_\tau(\pi) = c_0 \ \wedge\ (L_\tau(\pi i) = c_1 \ \vee\ L_\tau(\pi j) = c_2) \\
P_\tau((c_0 : s_1 ?)@\langle \pi, \pi i\rangle) &\equiv L_\tau(\pi) = c_0
\end{aligned}$$

Satisfaction

$$S_\tau((c_0 : c_1 \prec c_2)@\langle\pi, \pi i, \pi j\rangle) \quad \equiv \quad i < j$$

Satisfaction

$$
\begin{aligned}
S_\tau((c_0 : c_1 \prec c_2)@\langle\pi, \pi i, \pi j\rangle) &\equiv i < j \\
S_\tau((c_0 : \triangle c_1)@\langle\pi\rangle) &\equiv \vee\{L_\tau(\pi i) = c_1 \mid 1 \leq i \leq A_\tau(\pi)\} \\
S_\tau((c_0 : c_1 !)@\langle\pi, \pi i, \pi j\rangle) &\equiv i = j \\
S_\tau((c_0 : c_1 \Rightarrow s_2)@\langle\pi, \pi i, \pi j\rangle) &\equiv L_\tau(\pi j) \in s_2 \\
S_\tau((c_0 : c_1 \not\Leftrightarrow c_2)@\langle\pi, \pi i, \pi j\rangle) &\equiv L_\tau(\pi i) \neq c_1 \;\vee\; L_\tau(\pi j) \neq c_2 \\
S_\tau((c_0 : s_1 ?)@\langle\pi, \pi i\rangle) &\equiv L_\tau(\pi i) \in s_1
\end{aligned}
$$

Admissibility

A syntax tree τ is admissible iff it satisfies the *projection property*, i.e. $\forall \pi \in D_\tau$:

$$\begin{aligned} A_\tau(\pi) = 0 \quad &\Rightarrow \quad \langle L_\tau(\pi), R_\tau(\pi) \rangle \in L_G \\ A_\tau(\pi) \neq 0 \quad &\Rightarrow \quad R_\tau(\pi) = \sum_{i=1}^{i=A_\tau(\pi)} R_\tau(\pi i) \end{aligned}$$

$\mathcal{A}_G =$ admissible syntax trees for grammar G

Strong Models

$$I^0_{G,\tau} = \{r \in \mathcal{I}_\tau[\![G]\!] \mid P_\tau(r)\}$$
$$I^+_{G,\tau} = \{r \in I^0_{G,\tau} \mid S_\tau(r)\}$$
$$I^-_{G,\tau} = \{r \in I^0_{G,\tau} \mid \neg S_\tau(r)\}$$

$\tau : \sigma \models G$

a syntax tree τ is a strong model of property grammar G, with realization σ, iff it is admissible and $R_\tau(\varepsilon) = \sigma$ and $I^-_{G,\tau} = \emptyset$

Loose Semantics

admissible trees for utterance σ

$$\mathcal{A}_{G,\sigma} = \{\tau \in \mathcal{A}_G \mid R_\tau(\epsilon) = \sigma\}$$

fitness

$$F_{G,\tau} = I^+_{G,\tau} / I^0_{G,\tau}$$

loose models

$$\tau : \sigma \mathrel{|\!\!\approx} G \quad \text{iff} \quad \tau \in \operatorname*{argmax}_{\tau' \in \mathcal{A}_{G,\sigma}} (F_{G,\tau'})$$

Postulates

- Failure cumulativity
- Success cumulativity
- Constraint weighting
- Constructional complexity
- Propagation

Weighted Property Grammar

weighted property grammar $G = (P_G, L_G, \omega_G)$:

- (P_G, L_G) is a property grammar
- $\omega_G : P_G \rightarrow \mathbb{R}$ assigns a weight to each property

Instance location

We write at(r) for the node where property instance r applies.

$\forall p \in \mathcal{P}_{\mathcal{L}},\ \forall \pi_0, \pi_1, \pi_2 \in \mathbb{N}_0^*$:

$$\begin{aligned} \mathrm{at}(p@\langle \pi_0 \rangle) &= \pi_0 \\ \mathrm{at}(p@\langle \pi_0, \pi_1 \rangle) &= \pi_0 \\ \mathrm{at}(p@\langle \pi_0, \pi_1, \pi_2 \rangle) &= \pi_0 \end{aligned}$$

Sets of instances at node π

If B is a set of instances, then $B|_\pi$ is the subset of B of all instances applying at node π:

$$B|_\pi = \{r \in B \mid \mathrm{at}(r) = \pi\}$$

The sets of instances pertinent, satisfied, and violated at node π:

$$I^0_{G,\tau,\pi} = I^0_{G,\tau}|_\pi \qquad I^+_{G,\tau,\pi} = I^+_{G,\tau}|_\pi \qquad I^-_{G,\tau,\pi} = I^-_{G,\tau}|_\pi$$

Cumulative weights at node π

cumulative weights of pertinent, satisfied, and violated instances at node π:

$$W^0_{G,\tau,\pi} = \sum \{\omega_G(x) \mid \forall x@y \in I^0_{G,\tau,\pi}\}$$
$$W^+_{G,\tau,\pi} = \sum \{\omega_G(x) \mid \forall x@y \in I^+_{G,\tau,\pi}\}$$
$$W^-_{G,\tau,\pi} = \sum \{\omega_G(x) \mid \forall x@y \in I^-_{G,\tau,\pi}\}$$

Scoring factors

quality index, satisfaction ratio, and violation ratio at node π:

$$W_{G,\tau,\pi} = \frac{W^+_{G,\tau,\pi} - W^-_{G,\tau,\pi}}{W^+_{G,\tau,\pi} + W^-_{G,\tau,\pi}} \quad \rho^+_{G,\tau,\pi} = \frac{|I^+_{G,\tau,\pi}|}{|I^0_{G,\tau,\pi}|} \quad \rho^-_{G,\tau,\pi} = \frac{|I^-_{G,\tau,\pi}|}{|I^0_{G,\tau,\pi}|}$$

Scoring factors

to account for constructional complexity:

$$T_{G,\tau,\pi} = \{c : C \in P_G \mid L_\tau(\pi) = c\}$$

completeness index:

$$C_{G,\tau,\pi} = \frac{|I^0_{G,\tau,\pi}|}{|T_{G,\tau,\pi}|}$$

Index of grammaticality

index of precision:

$$P_{G,\tau,\pi} = kW_{G,\tau,\pi} + l\rho^{+}_{G,\tau,\pi} + mC_{G,\tau,\pi}$$

index of grammaticality:

$$g_{G,\tau,\pi} = \begin{cases} P_{G,\tau,\pi} \cdot \frac{1}{A_\tau(\pi)} \sum_{i=1}^{A_\tau(\pi)} g_{G,\tau,\pi i} & \text{if } A_\tau(\pi) \neq 0 \\ 1 & \text{if } A_\tau(\pi) = 0 \end{cases}$$

$g_{G,\tau,\varepsilon}$ is the score of loose model τ

Index of grammaticality

index of precision:

$$P_{G,\tau,\pi} = kW_{G,\tau,\pi} + l\rho^{+}_{G,\tau,\pi} + mC_{G,\tau,\pi}$$

index of grammaticality:

$$g_{G,\tau,\pi} = \begin{cases} P_{G,\tau,\pi} \cdot \frac{1}{A_\tau(\pi)} \sum_{i=1}^{A_\tau(\pi)} g_{G,\tau,\pi i} & \text{if } A_\tau(\pi) \neq 0 \\ 1 & \text{if } A_\tau(\pi) = 0 \end{cases}$$

$g_{G,\tau,\varepsilon}$ is the score of loose model τ

Pearson's correlation coefficient

$$\rho = 0.4857$$

Index of coherence

index of anti-precision:

$$A_{G,\tau,\pi} = kW_{G,\tau,\pi} - l\rho^{-}_{G,\tau,\pi} + mC_{G,\tau,\pi}$$

index of coherence:

$$\gamma_{G,\tau,\pi} = \begin{cases} A_{G,\tau,\pi} \cdot \frac{1}{A_\tau(\pi)} \sum_{i=1}^{A_\tau(\pi)} \gamma_{G,\tau,\pi i} & \text{if } A_\tau(\pi) \neq 0 \\ 1 & \text{if } A_\tau(\pi) = 0 \end{cases}$$

$\gamma_{G,\tau,\varepsilon}$ is the score of loose model τ

Index of coherence

index of anti-precision:

$$A_{G,\tau,\pi} = kW_{G,\tau,\pi} - l\rho^{-}_{G,\tau,\pi} + mC_{G,\tau,\pi}$$

index of coherence:

$$\gamma_{G,\tau,\pi} = \begin{cases} A_{G,\tau,\pi} \cdot \frac{1}{A_\tau(\pi)} \sum_{i=1}^{A_\tau(\pi)} \gamma_{G,\tau,\pi i} & \text{if } A_\tau(\pi) \neq 0 \\ 1 & \text{if } A_\tau(\pi) = 0 \end{cases}$$

$\gamma_{G,\tau,\varepsilon}$ is the score of loose model τ

Pearson's correlation coefficient

$$\rho = 0.5425$$

Conclusion

Property grammars are well-suited to the task of modeling graded grammaticality.

- model-theoretic *strong* semantics
- analyzing quasi-expressions:
  - loose models
  - fitness score to determine optimal loose models
- comparative admissibility of quasi-expressions:
  - scoring functions
  - Prost [2008] has shown that these functions can be tuned to agree well with human judgements
- contraint solver under construction