Unbounded recursion in two dimensions,
where syntax and prosody meet

Edward Stabler (UCLA)
Kristine M. Yu (UMass Amherst)

SCiL 2023

* With help from James McCloskey, Ryan Bennett, Emily Elfner, Dónall Ó Baoill, SCiL referees

2023-06-13

Unbounded recursion in two dimensions, where syntax and prosody meet
Edward Stabler (UCLA)
Kristine M. Yu (UMass Amherst)
SCiL 2023
* With help from James McCloskey, Ryan Bennett, Emily Elfner, Dónall Ó Baoill, SCiL referees

Language seems to lack in-principle bounds on both breadth and depth of analysis.

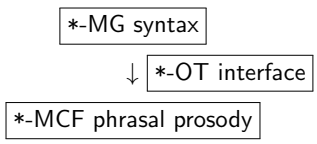
Many computational models of language allow unbounded depth, but not unbounded breadth, unbounded branching. E.g.

– standard grammars in Chomsky hierarchy do not provide it

– the best known treatments of tree automata do not provide it

The easy formal fix, well known in CS, comes from Kleene'56.
The focus this talk is to show how
– with that fix and other minor tweaks in well-known mechanisms –
we can provide a *simple and efficient* implementation
of a *surprising proposal about the syntax-prosody interface in Irish*.

1. **Unbounded branching is empirically motivated:**
   - syntax: flat coordination, adjunction
   - phonology/prosody: flat n-ary branching, e.g., Ft^+
2. **A simple computational architecture:**



implements an idea from Bennett,Elfner&McCloskey 2016

3. **Conclusions, prospects**

2023-06-13

1. Unbounded branching is empirically motivated:
- syntax: flat coordination, adjunction
- phonology/prosody: flat n-ary branching, e.g., Ft^+
2. A simple computational architecture:
*-MG syntax
↓ *-OT interface
*-MCF phrasal prosody
implements an idea from Bennett,Elfner&McCloskey 2016
3. Conclusions, prospects

In addition to arbitrarily deep trees,
linguists have proposed arbitrarily wide trees – unbounded branching –
in both syntax and phonology.

Yu'21:101 observes that linguists use phonological rules like $\omega \rightarrow \sigma^+$. And in OT phonology, it is common that at least some of the candidate structures are n-branching for $n > 2$. In syntax too, there are rules for unbounded coordination and adjunction.

Here we add unbounded branching to 3 components
– syntax, prosody, and a mediating OT interface

The extensions are illustrated with a challenging proposal about Irish, from Bennett, Elfner & McCloskey

We then conclude and mention some future work

Bennett,Elfner&McCloskey'16: Irish coordination

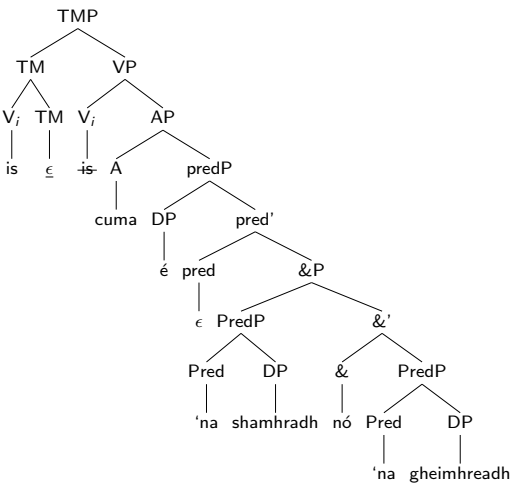
(1) is cuma é 'na shamhradh nó 'na gheimhreadh
COP.PRES no.matter it PRED summer or PRED winter
'It doesn't matter if it's summer or winter'

Bennett,Elfner&McCloskey'16: Irish coordination

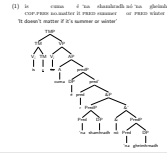
(1) is cuma é 'na shamhradh nó 'na gheimhreadh
COP.PRES no.matter it PRED summer or PRED winter
'It doesn't matter if it's summer or winter'

Bennett&al discuss Irish coordinate structures like this one

(1) is cuma é 'na shamhradh nó 'na gheimhreadh
COP.PRES no.matter it PRED summer or PRED winter
'It doesn't matter if it's summer or winter'

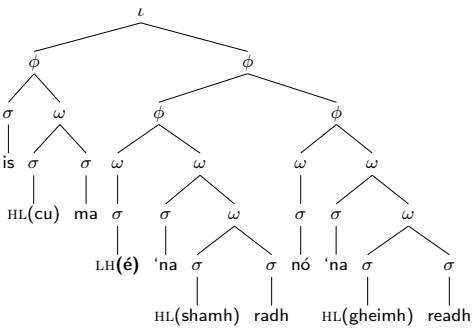


(1) is cuma é 'na shamhradh nó 'na gheimhreadh
COP.PRES no.matter it PRED summer or PRED winter
'It doesn't matter if it's summer or winter'



The syntax of that coordinate structure is roughly like this

(1) is cuma é ‘na shamhradh nó ‘na gheimhreadh
COP.PRES no.matter it PRED summer or PRED winter
‘It doesn’t matter if it’s summer or winter’



2023-06-13

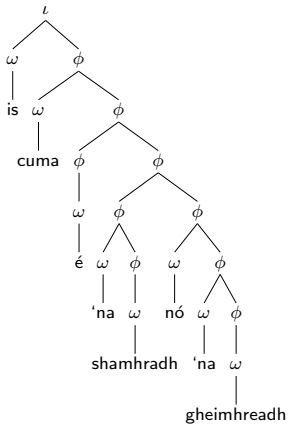
The prosody of the coordinate structure is roughly like this

Here, edge-based pitch accent placement is indicated by LH, HL

This prosodic structure is not the same as the syntactic structure:
– shallower, and the pronoun é is adjoined to ‘na shamhradh.

Why does the prosody differ in this way?

⇒ Bennett&al offer an OT account.



MATCH
Clause → Intonational Phrase
$\text{TMP}(x,y) \to \iota(x,y)$
$\text{TMP}(x,y) \to \phi(x,y)$

Maximal Projection → Phonological Phrase
$\text{XP}(x,y) \to \phi(x,y)$
$\text{XP}(x,y) \to \omega(x,y)$

Head → Prosodic Word
$\text{X}^0(x) \to \omega(x)$
$\text{X}^0(x) \to \sigma(x)$
...

BINARITY * * *
non-binary structures dispreferred
$X(x,y) \to X(x,y)$
$X(x) \to X(x)$
...

EQUALSISTERS * * * * *
sisters of diff cats dispreferred
$X(x,y) \to X(x,y)$ if $\text{root}(x) \neq \text{root}(y)$
...

STRONGSTART
σ as leftmost daughter dispreferred
$\text{X}(x,y) \to X(x,y)$ if $\text{root}(x) = \sigma$
...

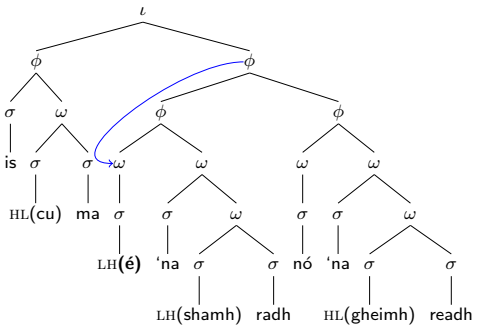
2023-06-13

To implement OT, GEN and constraints are transducers,
climbing the syntax tree to generate the prosodic structures.
Rule left sides are syntax with prosodic structures x, y, z ‘below’ them (these are subtrees, not strings!).
Constraint violations are dispreferred steps, shown in red.

MATCH prefers:
- clauses correspond to intonational phrases ι,
- XPs correspond to phonological phrases ϕ,
- heads correspond to prosodic words ω, and
- (empty structure is ignored.)

Structure on left respects MATCH, but violates other preferences (important: violations assessed for structures above prosodic word level, i.e, ϕ, ι):

(BINARITY) non-binary structures (under ϕ, ι parent) are dispreferred
(EQUALSISTERS) sisters of diff cat (with ϕ, ι parent) dispreferred
(STRONGSTART) leftmost daughter (of ϕ, ι) should not be σ



BINARITY

STRONGSTART *

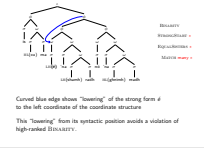
EQUALSISTERS *

MATCH many *

Curved blue edge shows “lowering” of the strong form *é* to the left coordinate of the coordinate structure

This “lowering” from its syntactic position avoids a violation of high-ranked BINARITY.

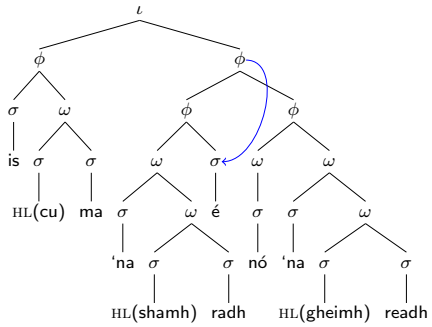
2023-06-13



Bennett&al show that when MATCH is outranked by other preferences, we can get the attested structure – ‘lowering’ the pronoun to 1st coordinate

We will come back to calculating this prosody, but first, let’s make the situation slightly more challenging!

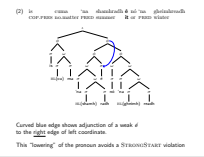
(2) is cuma ‘na shamhradh **é** nó ‘na gheimhreadh
COP.PRES no.matter PRED summer **it** or PRED winter



Curved blue edge shows adjunction of a weak *é* to the right edge of left coordinate.

This “lowering” of the pronoun avoids a STRONGSTART violation

2023-06-13



Bennet&al notice the variant (2), with é further right.

How does that pronoun move into the middle of a coordinate structure?!

This does not look like a syntactic operation!

But if GEN produces this candidate, when é is weak, it can be preferred by preferences already needed in the account of the prev example.

If syntax delivers > 2 coordinates, prosody should handle that too:

(3) is cuma é ['na shamhradh], ['na fhómhar] nó ['na gheimhreadh]
COP.PRES no.matter it PRED summer, PRED autumn, or PRED winter
'It doesn't matter if it's summer, autumn or winter'

(4) is cuma é ['na hearrach], ['na shamhradh], ['na fhómhar] nó ['na gheimhreadh]
COP.PRES no.matter it PRED spring, PRED summer, PRED autumn, or PRED winter
'It doesn't matter if it's spring, summer, autumn or winter'

...

2023-06-13

If syntax delivers > 2 coordinates, prosody should handle that too:

(3) is cuma é ['na shamhradh], ['na fhómhar] nó ['na gheimhreadh]
COP.PRES no.matter it PRED summer, PRED autumn, or PRED winter
'It doesn't matter if it's summer, autumn or winter'

(4) is cuma é ['na hearrach], ['na shamhradh], ['na fhómhar] nó ['na gheimhreadh]
COP.PRES no.matter it PRED spring, PRED summer, PRED autumn, or PRED winter
'It doesn't matter if it's spring, summer, autumn or winter'

...

3 coordinates,

4 coordinates, – Any number!

Let's look at how structures for these could be generated...

...and how – even with unbounded branching and displaced elements – we can nevertheless efficiently calculate optimal prosody.

Chomsky 2018 VS unbounded branching:

> The failure of strong generative capacity of [phrase structure grammar] ... is a failure of principle, as shown by unstructured coordination: e.g.,
>
> "the man was old, tired, tall,..., but friendly".
>
> Even unrestricted rewriting systems fail to provide such structures, which would require an infinite number of rules.

Chomsky 2021 FOR unbounded branching:

Merge($X_1, \ldots, X_n$,WS) = $\{\{X_1, \ldots, X_n\}$, W, Y$\}$

Marcolli,Berwick&Chomsky 2023 VS 3-ary merge:

> [3-ary merge] suffer[s] from both undergeneration and overgeneration with respect to binary Merge.

Kleene 1956: Expressions with *** and +** are regular.
(Chomsky hierarchy *unaffected* by adding *, + to rule right sides)

Here: *-MG, *-OT prosody. (not 3-ary, but unbounded!)

2023-06-13

Chomsky 2018 VS unbounded branching:
The failure of strong generative capacity of [phrase structure grammar] ... is a failure of principle, as shown by unstructured coordination: e.g.,
"the man was old, tired, tall,..., but friendly".
Even unrestricted rewriting systems fail to provide such structures, which would require an infinite number of rules.

Chomsky 2021 FOR unbounded branching:
Merge($X_1, \ldots, X_n$,WS) = $\{\{X_1, \ldots, X_n\}$, W, Y$\}$

Marcolli,Berwick&Chomsky 2023 VS 3-ary merge:
[3-ary merge] suffer[s] from both undergeneration and overgeneration with respect to binary Merge.

Kleene 1956: Expressions with *** and +** are regular.
(Chomsky hierarchy *unaffected* by adding *, + to rule right sides)

Here: *-MG, *-OT prosody. (not 3-ary, but unbounded!)

Chomsky has flip-flopped on unbounded branching several times, beginning in the 1960s.

But both computer scientists and linguists frequently use *-extended grammars and equivalents

We will now use them here, without disrupting a standard and appealing kind of account of the syntax/prosody interface...

*-MG and *-OT do not replace binary operations with 3-ary
but rather allow all operations to be unbounded, like Chomsky 2021.
So of course undergeneration wrt binary is not an issue

*-MG

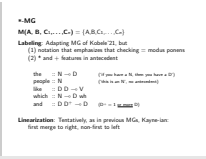
M(A, B, C_1,...,C_n) = {A,B,C_1,...,C_n}

Labeling: Adapting MG of Kobele'21, but
(1) notation that emphasizes that checking = modus ponens
(2) * and + features in antecedent

the	:: N $\multimap$ D	('if you have a N, then you have a D')
people	:: N	('this is an N', no antecedent)
like	:: D D $\multimap$ V	
which	:: N $\multimap$ D wh	
and	:: D D^+ $\multimap$ D	(D^+ = 1 or more D)

Linearization: Tentatively, as in previous MGs, Kayne-ian:
first merge to right, non-first to left

2023-06-13



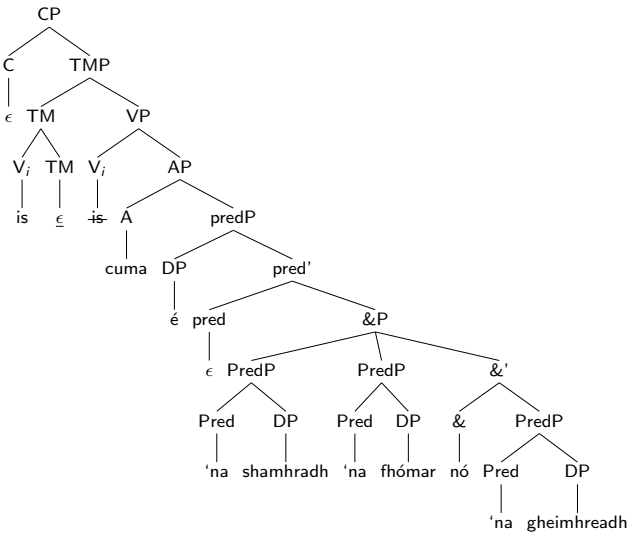
The $\multimap$ notation from linear logic partitions features into pos and neg occurrences: neg on left, pos on right.
This notation highlights the analogy between feature checking and *modus ponens*: neg and pos occurrences check each other.

Labeling is like in previous MG's, but now with Kleene + and *.
(Staying close to previous MG here, we leave recent proposals about labeling theory for future work.)

D^+ checks 1 or more Ds "simultaneously", as explained below.

(This merge and feature checking is simple enough to allow a complete implementation to be given in Appendix B of the paper and on a single slide below. A slightly unpacked version is linked on github. To enforce Kayne-ian linearization requires more bookkeeping – a ∗-MG parser enforcing linear order is also on the github site.)

3-coordinates: X-bar structure



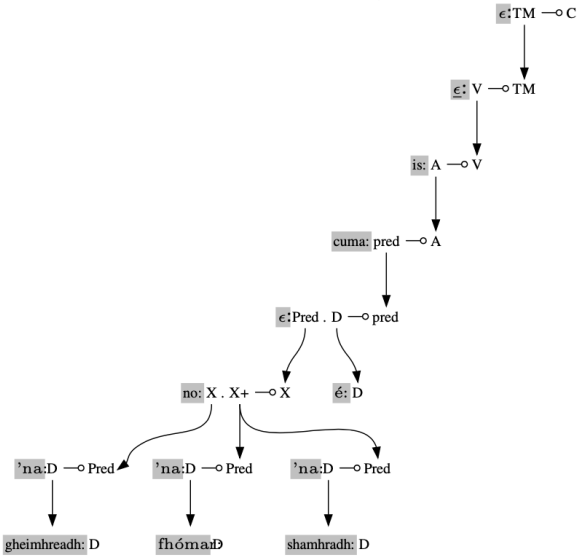
2023-06-13

The minimalist set notation can be drawn as an X-bar structure

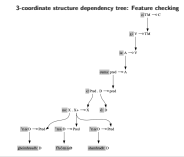
Neeleman et al have a recent paper in Syntax (2023) defending essentially this structure for flat coordination of 3 elements

We can derive this X-bar notation, but we don't want it in the derivation

3-coordinate structure dependency tree: Feature checking



2023-06-13



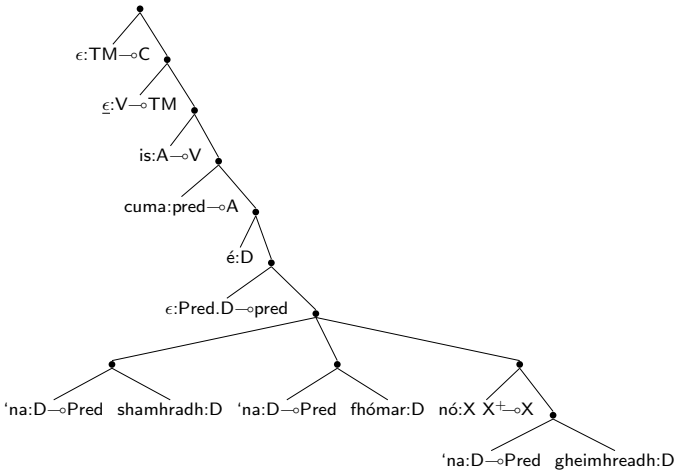
This is MG, so the fact that the X-bar structure is derived is fully determined by the lexical items, which stand in these feature-checking relations

Note how extremely simple the checking relation is in this dependency tree: each antecedent (negative) feature checks a (lower) consequent (positive) feature.

The new thing here is the X^+ feature. The var X can be any feature, and the $+$ indicates that it can check 1 or more matching positive features, 'simultaneously'

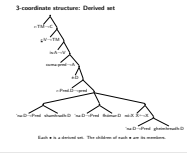
In 'completed' derivations, every feature checked except 1 at root: C.

3-coordinate structure: Derived set



Each • is a derived set. The children of each • are its members.

2023-06-13



To derive a traditional minimalist structure: whenever elements check, put them into a set.

These large nested sets can be displayed as graphs: Each dot is a set and edges indicate is-a-member-of relations.

Again, this is MG, the possible structures are fully determined by the lexical items which are now at the leaves.

(When there is syntactic movement, graphs of the nested sets are not trees, as illustrated by the example in Appendix B of the paper.)

In this derivation, all sets are binary except the coordinate structure with 3 elements.

We can have 2 or 3 or more coordinates there, without bound. The merge rule M and the extended feature-checking rules allow unbounded branching in coordinate structures.

String-based OT ⇒ *-OT on trees: first steps

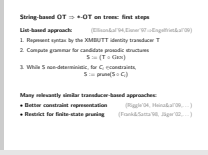
List-based approach: (Ellison&al'94,Eisner'97⇒Engelfriet&al'09)

1. Represent syntax by the XMBUTT identity transducer T
2. Compute grammar for candidate prosodic structures
$$S := (T \circ \textsc{Gen})$$
3. While S non-deterministic, for $C_i \in$constraints,
$$S := \text{prune}(S \circ C_i)$$

Many relevantly similar transducer-based approaches:

- **Better constraint representation** (Riggle'04, Heinz&al'09,...)
- **Restrict for finite-state pruning** (Frank&Satta'98, Jäger'02,...)

2023-06-13


String-based OT ⇒ *-OT on trees: first steps
List-based approach: (Ellison&al'94,Eisner'97⇒Engelfriet&al'09)
1. Represent syntax by the XMBUTT identity transducer T
2. Compute grammar for candidate prosodic structures
S := (T ∘ Gen)
3. While S non-deterministic, for C_i ∈constraints,
S := prune(S ∘ C_i)
Many relevantly similar transducer-based approaches:
• Better constraint representation (Riggle'04, Heinz&al'09,...)
• Restrict for finite-state pruning (Frank&Satta'98, Jäger'02,...)

Adapting finite-state string transducer approaches from Ellison'94, Eisner'97, Albro'97,'05, to the linear finite-state 'extended multi bottom up tree transducers' (XMBUTTs) of Engelfriet&al'09...

We begin with the list-based approach – the most transparent formalization of the Bennett&al proposal – but other relevantly similar OT alternatives should be extended to trees in future work.

The Engelfriet&al'09 definitions do no include *-rules, but they do not presume unique ranking either, so adding *, + rules is not difficult.

Many OT proposals not finite state! (Riggle'04, Koser&Jardine'20, Gerdemann&Hulden'12, Heinz&Lai'13, Lamont'21,'22,'23)

Lamont'23: OT accommodating all this "requires more complex computation than expected [in] a model of phonology"

Bennett&al: Irish prosodic pronoun displacement

(2) is cuma 'na shamhradh **é** nó 'na gheimhreadh
COP.PRES no.matter PRED summer **it** or PRED winter

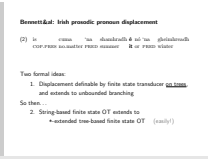
Two formal ideas:

1. Displacement definable by finite state transducer on trees, and extends to unbounded branching

So then...

2. String-based finite state OT extends to
*-extended tree-based finite state OT (easily!)

2023-06-13


Bennett&al: Irish prosodic pronoun displacement
(2) is cuma 'na shamhradh é nó 'na gheimhreadh
COP.PRES no.matter PRED summer it or PRED winter
Two formal ideas:
1. Displacement definable by finite state transducer on trees, and extends to unbounded branching
So then...
2. String-based finite state OT extends to *-extended tree-based finite state OT (easily!)

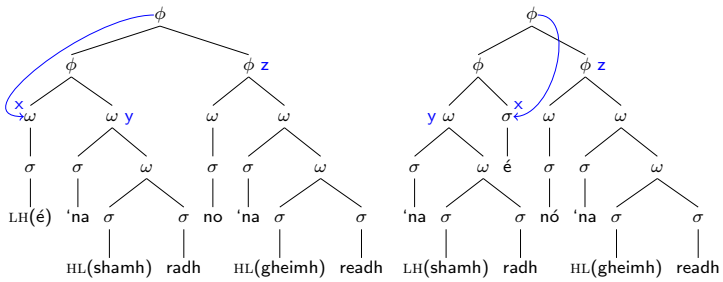
Antecedents:

Engelfriet&al'11, Maletti'11: XMBUTT

Graf'12: MBUTT constraints (w/o OT pruning, *-branching)

Yu'22: MBUTT/MCFG for multiple dependencies in phonology

Much prev work on tree automata does not assume unique ranking: Martens&Niehren'05, Engelfriet&al'09, Bahr'12,...

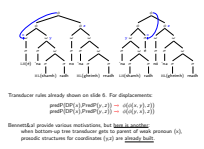


Transducer rules already shown on slide 6. For displacements:

$$\text{predP}(\text{DP}(x),\text{PredP}(y,z)) \rightarrow \phi(\phi(x,y),z))$$
$$\text{predP}(\text{DP}(x),\text{PredP}(y,z)) \rightarrow \phi(\phi(y,x),z))$$

Bennett&al provide various motivations, but here is another:
when bottom-up tree transducer gets to parent of weak pronoun (x),
prosodic structures for coordinates (y,z) are already built.

2023-06-13



Bennett&al:198 say "we assume adjunction here for several reasons".

> *"Elfner (2012:224) presents evidence based on the distribution of pitch accents that prosodic adjunction is the right interpretation Second, we take from recent work by Ito and Mester (e.g., 2006, 2009b, 2012) the idea that prosodic adjunction has a particularly central role"*

Adjunction to either side of left coordinate equally easy –
making room for optimization of the prosodic constraints.

Yu'21:92 "... bracketed strings make edge effects an accident... Without trees, there is nothing for the brackets to be edges of"

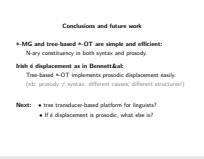
Conclusions and future work

***-MG and tree-based *-OT are simple and efficient:**
N-ary constituency in both syntax and prosody.

Irish é displacement as in Bennett&al:
Tree-based *-OT implements prosodic displacement easily.
(nb: prosody ≠ syntax: different causes, different structures!)

Next:
- tree transducer-based platform for linguists?
- If é displacement is prosodic, what else is?

2023-06-13



Note that in this unified model of prosody+syntax, as in Bennett&al, prosody ≠ syntax!

A. Expressive power

While it is true that standard regular rewrite grammars "need infinitely many rules" for unbounded branching, as Chomsky'18 says. . .

Kleene'56 showed that simply adding * and + to them does not change their expressive power,

so adding * and + to MCFGs, XMBUTTs, and MGs also leaves expressive power unchanged, and requires only minor changes in processing algorithms.

A. Expressive power

While it is true that standard regular rewrite grammars "need infinitely many rules" for unbounded branching, as Chomsky'18 says. . .

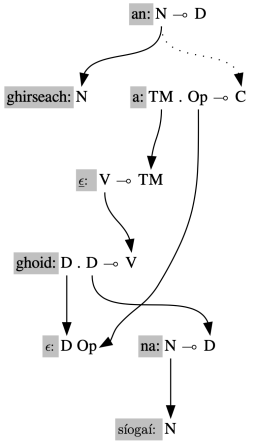
Kleene'56 showed that simply adding * and + to them does not change their expressive power,

so adding * and + to MCFGs, XMBUTTs, and MGs also leaves expressive power unchanged, and requires only minor changes in processing algorithms.

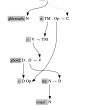
.

B. Contrast: syntactic movement in Irish

- an ghirseach a ghoid na síogaí
  the girl aL stole the fairies
  'the girl that the fairies stole away'



B. Contrast: syntactic movement in Irish
• an ghirseach a ghoid na síogaí
the girl aL stole the fairies
'the girl that the fairies stole away'



Appendix B of the paper observes:
∗-MG is easily adapted to McInnerney's recent proposal that unboundedly many adjuncts can be merged as sisters of the head they modify

Here, the feature checking dependency structure for a bound empty operator analysis of an Irish relative clause from McCloskey'02 – the dotted line shows the adjunct-check relation.

As described in the paper, any number of additional adjuncts could occur as sister to the noun and relative clause.

C. Implementation: ∗-MG Syntax

In prolog, represent: {A,B,C} with `[A,B,C]`, and

phon : $a_1 \ldots a_i \multimap a_{i+1} \ldots a_{i+j}$ with

`[phon]-[`a_1`,...,`a_i`]-[`a_{i+1}`,...,`a_{i+j}`]`.

```
r([A], A) :- l(A, []-[_]-[]).
r(X0, X) :- select(A, X0, X1), l(A, [F0|AN]-AP-AC0),
    ( nonvar(F0), F0=p(F) -> P=true ; F0=F, P=fail),
    ( select([F|BP]-B, AC0, AC) -> X2 = X1, BC = []
    ; select(B, X1, X2), l(B, []-[F|BP]-BC), AC=AC0
    ), m(B, [F|AN]-AP-AC, []-[F|BP]-BC, ABF),
    '&'(P, F, X2, X3, Cs), mrg(A, B, Cs, ABCs),
    r([ABCs|X3], X).
mrg(A, B, Cs, [A,B|Cs]).
l(_-A-B, A-B-[]).
l([A,B], D) :- l(A, AF), l(B, BF), m(B, AF, BF, D).
m(B, [F0|AN]-AP-AC, BF, AN-AP-ABC) :-
 (nonvar(F0), F0=p(F) -> true ; F0=F),
 (select([F|BP]-B0, AC, AC1) -> B0=B,
   (BP = [] -> AC1 = ABC ;
    BP = [G|BP1], smc([[G|BP1]-B0|AC1], [], ABC, []))
 ; BF = []-[F]-BC, smc(AC, BC, ABC, [])
 ; BF = []-[F,G|BP]-BC, smc([[G|BP]-B|AC], BC, ABC, [])).
'&'(_,_,X,X,[]).
'&'(true,F,X0,X,[C|Cs]) :-
 select(C,X0,X1), l(C,[]-[F]-[]), '&'(true,F,X1,X,Cs).
smc([], D, D, _).
smc([[F|C]-A|L], M, [[F|C]-A|N], Fs) :-
    \+member(F, Fs), smc(L, M, N, [F|Fs]).
```

2 rules for R

1 rule for merge

2 rules for labelling

1 rule for mate

2 rules for &

2 rules for SMC

This, by itself, derives unbounded Irish coordination, along with all the usual MG remnant movement, XX dependencies, etc. See github link.

C. Implementation: XMBUTT Prosody

First steps from §3 of paper, in prolog:

```
head(X) :- member(X, [c,tm,v,a,lpred,pred,b]).
phrase(XP) :- atom_chars(XP, L), last(L, p).
gen(T, Out) :- rule(T, Out).
gen(X/L,T) :-  maplist(gen,L,S), rule(X/S,T).
rule(_/[qw/[X0], qphi/[X1]], qi/[i/[X0, X1]]).
rule(X/[Ph/[]], qw/[w/[Ph/[]]]) :- head(X).
rule(X/[Ph/[]], qphi/[phi/[w/[Ph/[]]]]) :- phrase(X).
rule(_/[], qe).
rule(_/[qw/[X0], qphi/[X1]], qphi/[phi/[X0,X1]]).
rule(_/[qphi/[X0], qphi/[X1]], qphi/[phi/[X0,X1]]).
rule(_/[qe, X0], X0).
rule(_/[X0, qe], X0).
rule(_/[X0], X0).
```

1 rule for head, 1 for phrase, 2 for Gen, and transparent representations of the 9 transducer rules from §3.

This, by itself, derives Figure 1b prosodic structure from X-bar.

Github link adds transducer rules for displaced forms Figure 5a,b,c

A python implementation with composition and pruning is in preparation (also at github link)

Daniel M. Albro. 1997.
Evaluation, implementation, and extension of primitive optimality theory.
Master's thesis, UCLA.

Daniel M. Albro. 2005.
Studies in Computational Optimality Theory, with Special Reference to the Phonological System of Malagasy.
Ph.D. thesis, UCLA.

Patrick Bahr. 2012.
Modular tree automata.
In *Mathematics of Program Construction*, pages 263–299. Springer LNCS 7432.

Ryan Bennett, Emily Elfner, and James McCloskey. 2016.
Lightest to the right: An apparently anomalous displacement in Irish.
Linguistic Inquiry, 47(2):169–234.

Noam Chomsky. 2018.
Syntactic structures. some retrospective comments.
In Norbert Hornstein, Howard Lasnik, Pritty Patel-Grosz, and Charles Yang, editors, *Syntactic Structures after 60 Years*. De Gruyter Mouton.

Noam Chomsky. 2021.
Minimalism: Where are we now, and where can we hope to go.
Gengo Kenkyu, 160:1–41.

Noam Chomsky and George A. Miller. 1963.
Introduction to the formal analysis of natural languages.
In R. Duncan Luce, Robert R. Bush, and Eugene Galanter, editors, *Handbook of Mathematical Psychology, Volume II*, pages 269–321. Wiley.

Jason Eisner. 1997.
Efficient generation in primitive optimality theory.
In *Procs. 35th Annual Meeting of the Association for Computational Linguistics.*

Emily Elfner. 2012.
Syntax-Prosody Interactions in Irish.
Ph.D. thesis, University of Massachusetts, Amherst.

Mark T. Ellison. 1994.
Phonological derivation in optimality theory.
In *Procs. 15th Int. Conf. on Computational Linguistics*, pages 1007–1013.

Joost Engelfriet, Eric Lilin, and Andreas Maletti. 2009.
Extended multi bottom–up tree transducers: Composition and decomposition.
Acta Informatica, 46(8):561–590.

Robert Frank and Giorgio Satta. 1998.
Optimality theory and the generative complexity of constraint violability.
Computational Linguistics, 24:307–315.

Dale Gerdemann and Mans Hulden. 2012.
Practical finite state optimality theory.
In *Procs 10th International Workshop on Finite State Methods and Natural Language Processing.*

Thomas Graf. 2012.
Concealed reference-set computation: How syntax escapes the parser's clutches.
In Anna Maria Di Sciullo, editor, *Towards a Biolinguistic Understanding of Grammar. Essays in Interfaces*, pages 339–362. John Benjamins.

Jeffery Heinz and Regine Lai. 2013.
Vowel harmony and subsequentiality.
In *Procs 13th Meeting on the Mathematics of Language.*

Junko Ito and Armin Mester. 2012.
Recursive prosodic phrasing in Japanese.
In *Prosody Matters: Essays in Honor of Elisabeth Selkirk*, pages 280–303. Elsevier.

S. C. Kleene. 1956.
Representation of events in nerve nets and finite automata.
In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 3–42. Princeton University Press.

Gregory M. Kobele. 2021.
Minimalist grammars and decomposition.
Technical report, Universität Leipzig.

Nate Koser and Adam Jardine. 2020.
The complexity of optimizing over strictly local constraints.
In *Procs 43rd Annual Penn Linguistics Conference*, page 125–134.

Andrew Lamont. 2022.
Directional Harmonic Serialism.
Ph.D. thesis, University of Massachusetts, Amherst.

Andreas Maletti. 2011.
Tree transformations and dependencies.
In *Proceedings of the 12th Conference on Mathematics of Language*, LNAI 6878, pages 1–20. Springer.

Wim Martens and Joachim Niehren. 2005.
Minimizing tree automata for unranked trees.
In *Procs 10th Annual Symposium on Database Programming Languages*, pages 232–246. Springer LNCS 3774.

Robert C. Berwick Matilde Marcolli, Noam Chomsky. 2023.
Mathematical structure of syntactic merge.
Computing Research Repository, arXiv:2305.18278.

James McCloskey. 2002.
Resumption, successive cyclicity, and the locality of operations.
In Samuel D. Epstein and T. Daniel Seely, editors, *Derivation and explanation in the Minimalist Program*, pages 184–226. Blackwell.

James McCloskey. 2022.
The syntax of Irish Gaelic.
Technical report, University of Califoria, Santa Cruz.
Forthcoming.

Andrew McInnerney. 2022.
The Argument/Adjunct Distinction and the Structure of Prepositional Phrases.
Ph.D. thesis, University of Michigan.

Jason Riggle. 2004.
Generation, Recognition, and Learning in Finite State Optimality Theory.
Ph.D. thesis, University of California, Los Angeles.

Kristine M. Yu. 2021.
Computational perspectives on phonological constituency and recursion.
Catalan Journal of Linguistics, 77:114.

Kristine M. Yu. 2022.
Representations for multiple dependencies in prosodic structures.
Proceedings of the Society for Computation in Linguistics, 5:15.

22