Towards an Algebraic Morphosyntax

Thomas Graf
mail@thomasgraf.net
http://thomasgraf.net

Stony Brook University

Moscow State University
December 16, 2015

Piece of a Larger Puzzle

- There is a huge number of **morphosyntactic scales**:
  - comparative suppletion (ABC, ABB, *ABA, *AAB)
  - case hierarchy for pronoun suppletion
  - omnivorous number (sg/pl + sg/pl = pl, *sg + sg = sg)
  - resolved gender agreement

    ⋮
- Different syntactic mechanisms seem to be involved
  ⇒ very different syntactic accounts for these phenomena
- **Research Program**
  If we abstract away from the syntactic machinery,
  do we find commonalities among all these scales?

What is the PCC?

Person Case Constraint (PCC)

Whether the direct object (**DO**) and the indirect object (**IO**) of a clause can both be cliticized is contingent on the person specification of **DO** and **IO**.

(1)	Roger	*me/le	leur	a	presésenté.
	Roger	1SG/3SG.ACC	3PL.DAT	has	shown

'Roger has shown me/him to them.'

Questions & Goals

- What are the descriptive properties of PCCs?
  ⇒ algebraic unification in terms of presemilattices
- Can those properties be tied to independently motivated linguistic assumptions? ⇒ connection to feature geometry

What is the PCC?

Person Case Constraint (PCC)

Whether the direct object (**DO**) and the indirect object (**IO**) of a clause can both be cliticized is contingent on the person specification of **DO** and **IO**.

(1) Roger *me/le leur a presésenté.
Roger 1SG/3SG.ACC 3PL.DAT has shown
'Roger has shown me/him to them.'

Questions & Goals

- What are the descriptive properties of PCCs?
  ⇒ algebraic unification in terms of presemilattices
- Can those properties be tied to independently motivated linguistic assumptions? ⇒ connection to feature geometry

Outline

1. Person Case Constraints: An Overview
   - PCC Typology
   - Previous Proposals
2. Characterizing the Class of PCCs
   - The Generalized PCC
   - Algebraic Characterization via Person Locality
3. Connection to Feature Complexity
   - Reducing Person Locality to Feature Complexity
   - Reducing Feature Complexity to Feature Geometries

The PCC: A Closer Look

- attested in a variety of languages, including French, Spanish, Catalan, and Classical Arabic (Kayne 1975; Bonet 1991, 1994)
- specifics of PCC differ between languages, dialects, idiolects

Four Attested PCC Variants

- **Strong PCC** (S-PCC; Bonet 1994)
  DO must be 3.
- **Ultrastrong PCC** (U-PCC; Nevins 2007)
  DO is less local than **IO** (where $3 < 2 < 1$).
- **Weak PCC** (W-PCC; Bonet 1994)
  3**IO** combines only with 3**DO**.
- **Me-first PCC** (M-PCC; Nevins 2007)
  If **IO** is 2 or 3, then **DO** is not 1.

The Four PCC Variants (Walkow 2012)

IO/DO	1	2	3
1	NA	*	✓
2	*	NA	✓
3	*	*	NA

S-PCC

IO/DO	1	2	3
1	NA	✓	✓
2	*	NA	✓
3	*	*	NA

U-PCC

IO/DO	1	2	3
1	NA	✓	✓
2	✓	NA	✓
3	*	*	NA

W-PCC

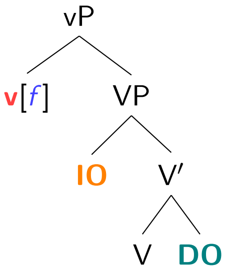
IO/DO	1	2	3
1	NA	✓	✓
2	*	NA	✓
3	*	✓	NA

M-PCC

The PCC in Minimalism

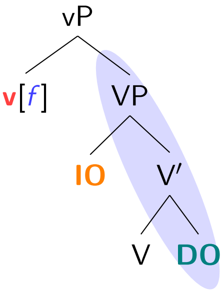
- Variety of proposals, work well empirically:
  - Anagnostopoulou (2005)
  - Nevins (2007)
  - Béjar and Rezac (2009)
  - Walkow (2012)
- **Shared Idea:** PCCs epiphenomenal, arise from more basic **restrictions on the Agree operation**
- **Conceptual Drawbacks**
  - non-standard Agree mechanisms
  - highly specific assumptions about feature system
  - technical, complex
  - hard to determine which assumptions are really needed

Example: Intuition Behind Nevins (2007)



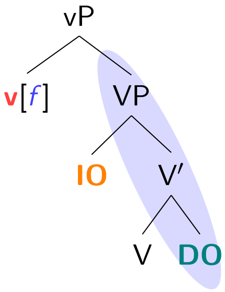
- v needs to agree with a particular feature *f*
- a search domain is established, depending on the type of *f*
- ungrammatical if the domain contains DO but not IO
- v agrees with both DO and IO ⇒ IO and DO must have the same value for *f*

Example: Intuition Behind Nevins (2007)



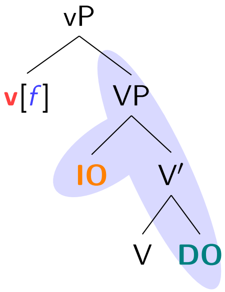
- **v** needs to agree with a particular feature f
- a search domain is established, depending on the type of f
- ungrammatical if the domain contains **DO** but not **IO**
- **v** agrees with both **DO** and **IO** $\Rightarrow$ **IO** and **DO** must have the same value for f

Example: Intuition Behind Nevins (2007)



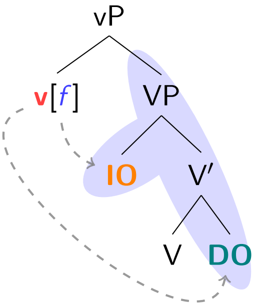
- v needs to agree with a particular feature *f*
- a search domain is established, depending on the type of *f*
- ungrammatical if the domain contains DO but not IO
- v agrees with both DO and IO ⇒ IO and DO must have the same value for *f*

Example: Intuition Behind Nevins (2007)



- **v** needs to agree with a particular feature *f*
- a search domain is established, depending on the type of *f*
- ungrammatical if the domain contains **DO** but not **IO**
- **v** agrees with both **DO** and **IO** ⇒ **IO** and **DO** must have the same value for *f*

Example: Intuition Behind Nevins (2007)



- **v** needs to agree with a particular feature *f*
- a search domain is established, depending on the type of *f*
- ungrammatical if the domain contains **DO** but not **IO**
- **v** agrees with both **DO** and **IO** ⇒ **IO** and **DO** must have the same value for *f*

Example: Assumptions of Nevins (2007)

- **Operations**
  - Agree steps happen concurrently
  - constraints on search domain
  - matching condition on **IO** and **DO**
- **Structure**
  - clitics are PF-realization of Agree
  - **IO** structurally higher than **DO**
- **Features**
  - features are binary valued
  - novel definition of contrastive features
  - feature values can be marked or unmarked
  - specific feature decomposition of person:

Person	Feature Matrix
1	[+author,+participant]
2	[-author,+participant]
3	[-author,-participant]

Evaluation

- Previous accounts work on an empirical level.
- They are complex because they try to do two things at once:
  1. enforce the PCC with Minimalist machinery,
  2. capture the attested typology.
- But that's more ambitious than necessary!

The Secret Power of Merge (Graf 2011; Kobele 2011)

Every syntactic constraint that can be computed with a finite amount of working memory can be enforced purely via Merge.

- The PCCs can be enforced by Merge, we do not need to extend our framework at all.
- The big issue is Point 2: There are $2^6 = 64$ logically possible PCC variants. **Why do we find only 4 PCCs**?

Evaluation

- Previous accounts work on an empirical level.
- They are complex because they try to do two things at once:
  1. enforce the PCC with Minimalist machinery,
  2. capture the attested typology.
- But that's more ambitious than necessary!

The Secret Power of Merge (Graf 2011; Kobele 2011)

Every syntactic constraint that can be computed with a finite amount of working memory can be enforced purely via Merge.

- The PCCs can be enforced by Merge, we do not need to extend our framework at all.
- The big issue is Point 2: There are $2^6 = 64$ logically possible PCC variants. **Why do we find only 4 PCCs**?

Outline

1. Person Case Constraints: An Overview
   - PCC Typology
   - Previous Proposals
2. Characterizing the Class of PCCs
   - The Generalized PCC
   - Algebraic Characterization via Person Locality
3. Connection to Feature Complexity
   - Reducing Person Locality to Feature Complexity
   - Reducing Feature Complexity to Feature Geometries

The Generalized PCC

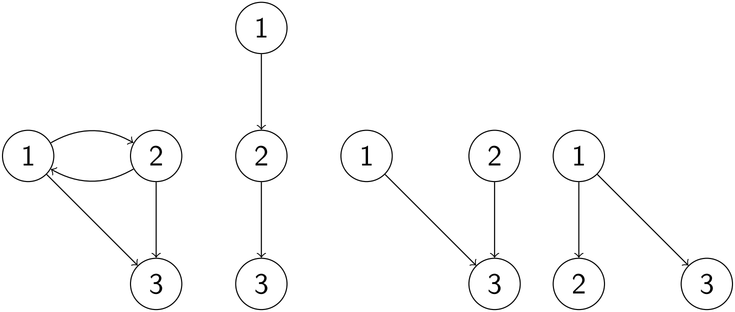
The U-PCC was defined in terms of person locality.
This system can be extended to all four PCC-types.

Generalized PCC (G-PCC)

IO is not less local than **DO** (**IO** $\nless$ **DO**), where

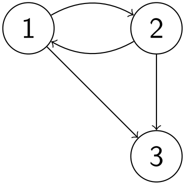
S-PCC:	$1 > 2$	$1 > 3$	$2 > 1$	$2 > 3$
U-PCC:	$1 > 2$	$1 > 3$		$2 > 3$
W-PCC:	$1 > 3$			$2 > 3$
M-PCC:	$1 > 2$	$1 > 3$		

Person Locality Hierarchies



S-PCC U-PCC W-PCC M-PCC

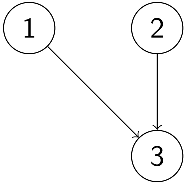
Example 1: S-PCC



$1 > 2$
$1 > 3$
$2 > 1$
$2 > 3$

IO/DO	1	2	3
1	NA	*	✓
2	*	NA	✓
3	*	*	NA

Example 2: W-PCC



$1 > 3$

$2 > 3$

IO/DO	1	2	3
1	NA	✓	✓
2	✓	NA	✓
3	*	*	NA

Presemilattices

The G-PCC gives a unified description of the four PCCs, but we could have drawn any kind of graph.
What makes the previous four structures so special?

First, they are all **presemilattices** (Plummer and Pollard 2012).

Definition (Presemilattices for Linguists)

A structure S is a **presemilattice** iff for all nodes **u** and **v** of S, there is some node **t** such that

- **t** "reflexively dominates" **u** and **v**, or
- **u** and **v** "reflexively dominate" **t**.

Presemilattices

The G-PCC gives a unified description of the four PCCs,
but we could have drawn any kind of graph.
What makes the previous four structures so special?

First, they are all **presemilattices** (Plummer and Pollard 2012).

Definition (Presemilattices for Linguists)

A structure S is a **presemilattice** iff for all nodes **u** and **v** of S, there is some node **t** such that

- **t** "reflexively dominates" **u** and **v**, or
- **u** and **v** "reflexively dominate" **t**.

Two More Restrictions

The number of presemilattices with three nodes is still more than 4. We have to stipulate two more properties:

Top and Bottom

Top For all x, $1 < x$ implies $x < 1$.
'Every person feature is at most as local as 1.'

Bottom There is no $x \neq 3$ such that $x < 3$.
'No person feature is less local than 3.'

Unifying the PCCs

The class of attested PCCs is given by

- the G-PCC IO $\not<$ DO such that
- $<$ defines a presemilattice $\mathcal{P}$ over $\{1, 2, 3\}$, and
- $\mathcal{P}$ respects both Top and Bottom.

Two More Restrictions

The number of presemilattices with three nodes is still more than 4. We have to stipulate two more properties:

Top and Bottom

Top For all x, $1 < x$ implies $x < 1$.
'Every person feature is at most as local as 1.'

Bottom There is no $x \neq 3$ such that $x < 3$.
'No person feature is less local than 3.'

Unifying the PCCs

The class of attested PCCs is given by

- the **G-PCC IO $\not<$ DO** such that
- $<$ defines a presemilattice $\mathcal{P}$ over $\{1, 2, 3\}$, and
- $\mathcal{P}$ respects both Top and Bottom.

Outline

1. Person Case Constraints: An Overview
   - PCC Typology
   - Previous Proposals

2. Characterizing the Class of PCCs
   - The Generalized PCC
   - Algebraic Characterization via Person Locality

3. Connection to Feature Complexity
   - Reducing Person Locality to Feature Complexity
   - Reducing Feature Complexity to Feature Geometries

Top and Bottom Match Feature Complexity

Top and Bottom are stipulations, but express a common intuition: 1 is "maximally complex", 3 "minimally complex".

Example 1: Person Specifications in Nevins (2007)

Person	Specification
1	[+author,+participant]
2	[-author,+participant]
3	[-author,-participant]

Example 2: Alternative Specification a la Nevins (2007)

Person	Specification
1	{participant,author}
2	{participant}
3	{}

Doing Away with Top and Bottom?

Syntactic proposals use feature geometry to derive PCC typology. Can we do the same? Yes, and No.

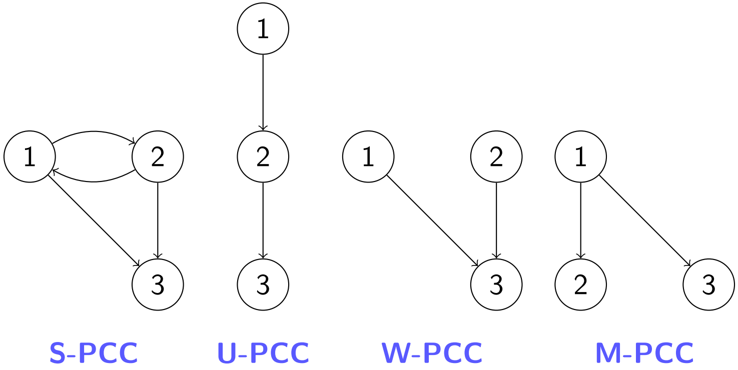
Algebraic Feature Complexity [Idea Sketch]

PCC locality is partially determined by feature complexity:

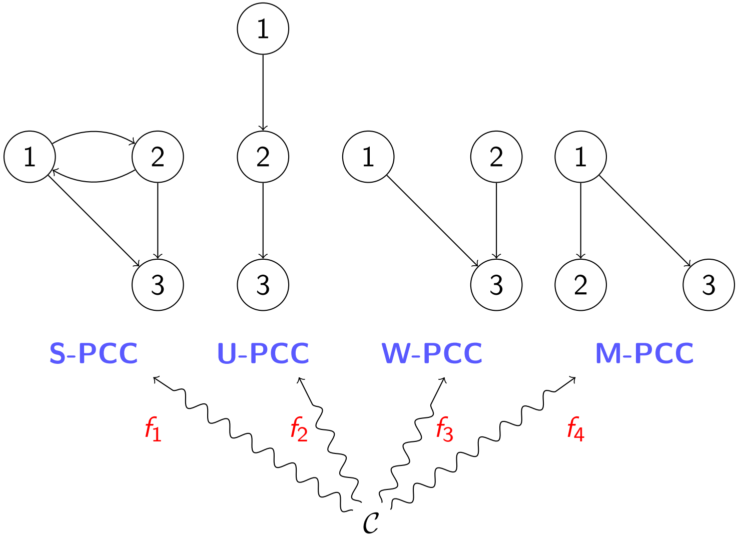
- Person features are ordered by their internal complexity $\Rightarrow$ algebraic structure $\mathcal{C}$
- PCC locality rankings are exactly those structures that
  - can be obtained from $\mathcal{C}$ by a map f such that
  - f preserves certain properties of $\mathcal{F}$

The above is feasible, but more stipulative than one would expect.

Schema of Reduction to Feature Complexity

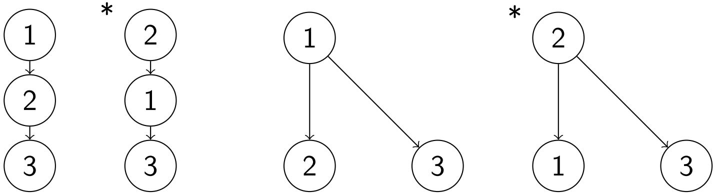


Schema of Reduction to Feature Complexity

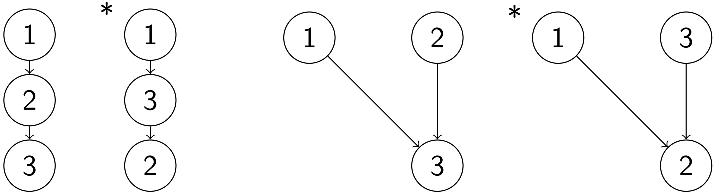


What does $\mathcal{C}$ Look Like?

- $\mathcal{C}$ must assign different complexity to 1 and 2:

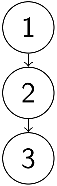


- $\mathcal{C}$ must assign different complexity to 2 and 3:



The Only Viable Shape of $\mathcal{C}$

- The previous observations entail that $\mathcal{C}$ must be

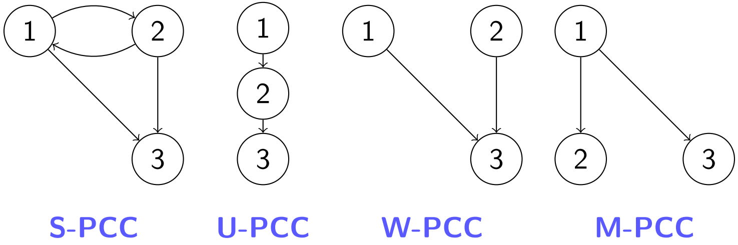


- This is **identical to Zwicky's person hierarchy!**
(Zwicky 1977)

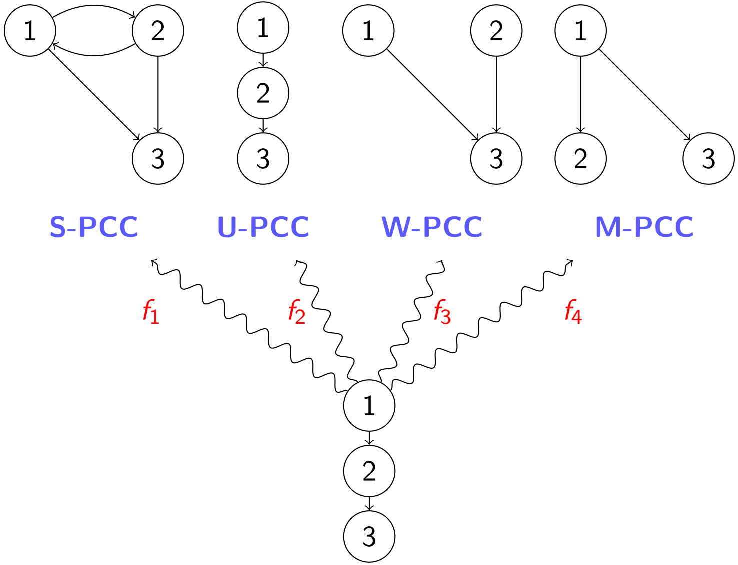
From $\mathcal{C}$ to Person Locality

- The 4 PCCs are generated from $\mathcal{C}$ by those maps that
  - **preserve connectedness** ($\approx$ Presemilattice)
  - **preserve maximality** ($\approx$ Top)
  - **preserve lack of daughter nodes** ($\approx$ Bottom)
- But where does $\mathcal{C}$ come from?
  Can we obtain $\mathcal{C}$ from some feature geometry $\mathcal{G}$?

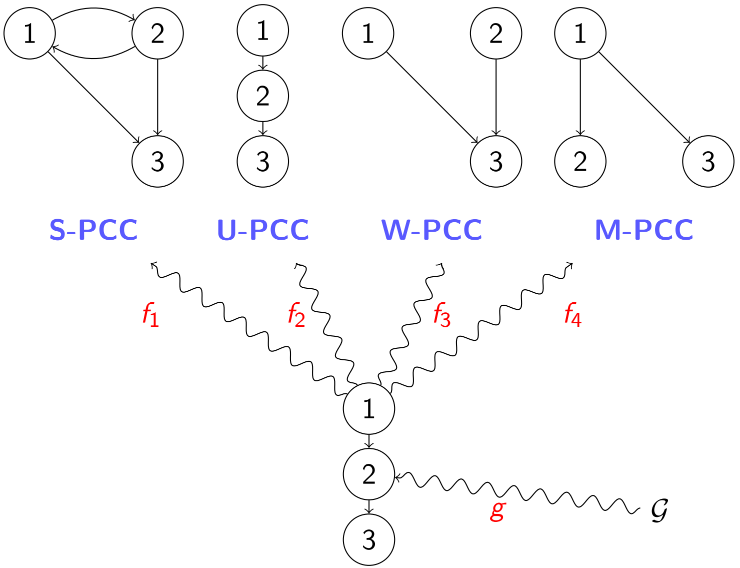
Obtaining $\mathcal{C}$ from Feature Geometries



Obtaining $\mathcal{C}$ from Feature Geometries



Obtaining $\mathcal{C}$ from Feature Geometries



Using Nevin's Geometry

$\mathcal{C}$ is easily obtained from the feature specification in Nevins (2007) if person complexity is determined by the number of features.

Reminder: Set-Theoretic Specification a la Nevins (2007)

Person	Specification
1	{participant,author}
2	{participant}
3	{}

This counting measure also works for unnatural specifications:

Example: Specification with Distinguished Feature for 3

Person	Specification
1	{participant,author,non-addressee}
2	{participant,addressee}
3	{non-participant}

Using Nevin's Geometry

$\mathcal{C}$ is easily obtained from the feature specification in Nevins (2007) if person complexity is determined by the number of features.

Reminder: Set-Theoretic Specification a la Nevins (2007)

Person	Specification
1	{participant,author}
2	{participant}
3	{}

This counting measure also works for unnatural specifications:

Example: Specification with Distinguished Feature for 3

Person	Specification
1	{participant,author,non-addressee}
2	{participant,addressee}
3	{non-participant}

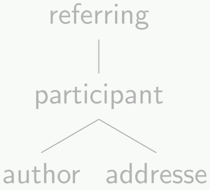
Another Feature Geometry: Harley and Ritter (2002)

- Without restrictions on what counts as a complexity measure, any feature geometry can be the basis for $\mathcal{C}$.
- But **some feature geometries are compatible with more complexity measures** than others.

Example: Harley and Ritter (2002) Needs a Weighted Measure

1 and 2 are structurally equivalent: same number of features, same structural representation $\Rightarrow$ features must be weighted

Person	Specification
1	{ref,part,auth}
2	{ref,part,addr}
3	{ref}



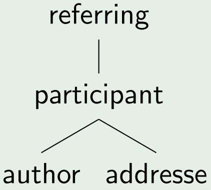
Another Feature Geometry: Harley and Ritter (2002)

- Without restrictions on what counts as a complexity measure, any feature geometry can be the basis for $\mathcal{C}$.
- But **some feature geometries are compatible with more complexity measures** than others.

Example: Harley and Ritter (2002) Needs a Weighted Measure

1 and 2 are structurally equivalent: same number of features, same structural representation $\Rightarrow$ features must be weighted

Person	Specification
1	{ref,part,auth}
2	{ref,part,addr}
3	{ref}



Technical Summary

- **Natural algebraic characterization** of the attested PCCs:
  - a ban against specific person locality configurations (G-PCC),
  - locality structures must be presemilattices,
  - locality structures respect both Top and Bottom.
- Going **one level deeper**:
  - person complexity must be $1 > 2 > 3$,
  - person complexity restricts shape of locality structures (stipulative right now, but algebraically fairly natural).
- Going **even deeper**:
  - person complexity determined by feature geometry
  - no obvious natural link at this point, but some geometries derive person complexity more easily

What's Next

- At this point there's **too many algebraic solutions**.
- We need to look at morphosyntax beyond person:
  1. number
  2. gender
  3. animacy
  4. case
  5. comparatives
- All phenomena should follow from a given feature geometry once all parameters have been fixed
  - mapping from feature geometry to complexity structures
  - mappings from complexity structures to locality structures

References I

Anagnostopoulou, Elena. 2005. Strong and weak person restrictions: A feature checking analysis. In *Clitics and affixation*, ed. Lorie Heggie and Francisco Ordoñez, 199–235. Amsterdam: John Benjamins.

Béjar, Susana, and Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40:35–73.

Bonet, Eulàlia. 1991. *Morphology after syntax: Pronominal clitics in Romance*. Doctoral Dissertation, MIT, Boston, MA.

Bonet, Eulàlia. 1994. The Person-Case Constraint: A morphological approach. In *The Morphology-Syntax Connection*, number 22 in MIT Working Papers in Linguistics, 33–52.

Graf, Thomas. 2011. Closure properties of minimalist derivation tree languages. In *LACL 2011*, ed. Sylvain Pogodalla and Jean-Philippe Prost, volume 6736 of *Lecture Notes in Artificial Intelligence*, 96–111. Heidelberg: Springer.

Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78:482–526.

Kayne, Richard S. 1975. *French syntax: The transformational cycle*. Cambridge, Mass.: MIT Press.

References II

Kobele, Gregory M. 2011. Minimalist tree languages are closed under intersection with recognizable tree languages. In *LACL 2011*, ed. Sylvain Pogodalla and Jean-Philippe Prost, volume 6736 of *Lecture Notes in Artificial Intelligence*, 129–144.

Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory* 25:273–313.

Plummer, Andrew, and Carl Pollard. 2012. Agnostic possible world semantics. In *LACL 2012*, ed. Denis Béchet and Alexander Dikovsky, number 7351 in Lecture Notes in Computer Science, 201–212. Heidelberg: Springer.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic*. Oxford: Oxford University Press.

Walkow, Martin. 2012. *Goals, big and small*. Doctoral Dissertation, University of Massachusetts Amherst.

Zwicky, Arnold. 1977. Hierarchies of person. In *Chicago Linguistic Society*, volume 13, 714–733.

Why IO $\not<$ DO?

Reminder: Unifying the PCCs

The class of attested PCCs is given by

- the **G-PCC IO $\not<$ DO** such that
- $<$ defines a presemilattice $\mathcal{P}$ over $\{1, 2, 3\}$, and
- $\mathcal{P}$ respects both Top and Bottom.

Maybe our problem with reducing the PCCs to feature geometries is due to our peculiar choice of **G-PCC**?

Spoiler

It is not.

Why IO $\not<$ DO?

Reminder: Unifying the PCCs

The class of attested PCCs is given by

- the **G-PCC IO $\not<$ DO** such that
- $<$ defines a presemilattice $\mathcal{P}$ over $\{1, 2, 3\}$, and
- $\mathcal{P}$ respects both Top and Bottom.

Maybe our problem with reducing the PCCs to feature geometries is due to our peculiar choice of **G-PCC**?

Spoiler

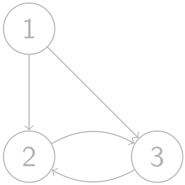
It is not.

Typology with Other Constraints

	P1	P2	P3	P4
IO $\not<$ DO	S	U	W	M
DO $<$ IO	W	U	S	M2

Me-second PCC (M2-PCC): If there is a DO, IO must be 1. [unattested]

- Under IO $\not<$ DO, M2-PCC is given by



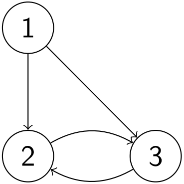
- Weakening Bottom to allow for this structure also brings in

Typology with Other Constraints

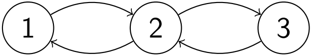
	P1	P2	P3	P4
IO $\not<$ **DO**	S	U	W	M
DO $<$ **IO**	W	U	S	M2

Me-second PCC (M2-PCC): If there is a **DO**, **IO** must be 1. [unattested]

- Under **IO** $\not<$ **DO**, **M2-PCC** is given by



- Weakening Bottom to allow for this structure also brings in



Typology with Additional Structures

	P1	P2	P3	P4	P5	P6
IO $\nless$ DO	S	U	W	M	M2	I
DO $<$ IO	W	U	S	M2	M	N

Indiscriminate PCC (I-PCC): No IO-DO clitic combinations. [Cairene Arabic (Shlonsky 1997:207, Walkow p.c.)]

Null PCC (N-PCC): Any clitic combination.

Typology with Additional Structures

	P1	P2	P3	P4	P5	P6
IO $\not<$ **DO**	S	U	W	M	M2	I
DO $<$ **IO**	W	U	S	M2	M	N

Indiscriminate PCC (I-PCC): No **IO**-**DO** clitic combinations. [Cairene Arabic (Shlonsky 1997:207, Walkow p.c.)]

Null PCC (N-PCC): Any clitic combination.

The Full Extended Typology

	P1	P2	P3	P4	P5	P6
IO $\not<$ DO	S	U	W	M	M2	I
DO $<$ IO	W	U	S	M2	M	N
IO $<$ DO	W	U	S	M2	M	N
DO $\not<$ IO	S	U	W	M	M2	I

Implications

- Choice of G-PCC has minor effect on predicted PCC typology.
- Allowing structures P5 and P6 requires a change to Bottom/Preservation of lack of daughters.
- However, the complexity ranking $\mathcal{C}$ stays the same
  $\Rightarrow$ problem of linking $\mathcal{C}$ to feature geometry unchanged

The Full Extended Typology

	P1	P2	P3	P4	P5	P6
IO $\not<$ **DO**	S	U	W	M	M2	I
DO $<$ **IO**	W	U	S	M2	M	N
IO $<$ **DO**	W	U	S	M2	M	N
DO $\not<$ **IO**	S	U	W	M	M2	I

Implications

- Choice of **G-PCC** has minor effect on predicted PCC typology.
- Allowing structures P5 and P6 requires a change to Bottom/Preservation of lack of daughters.
- However, the complexity ranking $\mathcal{C}$ stays the same $\Rightarrow$ problem of linking $\mathcal{C}$ to feature geometry unchanged