omer.lingsite.org/BCGL10.pdf

Assessing the morphosemantic program for φ−features: the prospects for a cross−modularly stable representation

Omer Preminger
Dept. of Linguistics / Language Science Center
University of Maryland

KU Leuven
Brussels
December, 2017

Outline of the talk

- Background
  - Modularity within the grammar: some preliminaries
  - Non-isomorphic mappings are the norm
- The case against isomorphic mapping for φ−features
  - a diagnostic for feature inertness in syntax
  - irreconcilable cross-modular mismatches in inertness
- Is this all just about “markedness”?
  Is there a Calabresean/Nevinsian alternative to all this?
  (No.)
- Discussion

Background

Modularity within the grammar: some preliminaries

- We can distinguish different notions of 'modul(ar(ity))' that are deployed – and sometimes run together – in discussions of the *morphology* ⇔ *syntax* ⇔ *semantics* landscape

  ➤ <u>representational differentiation</u> (e.g. Jackendoff 1997)

  - division into modules based on the different informational primitives each module traffics in

  ➤ <u>information encapsulation</u> (e.g. Fodor 1983)

  - division into modules based on limitations on the flow of information

  ➤ <u>computational differentiation</u> (e.g. Chomsky 1995 (?))

  - division into modules based on different nature of computations that each module carries out

- It's important to note, though, that these are not mutually-exclusive notions;
  - they might not even be mutually *distinguishable* in all cases.

- E.g.: suppose we find that PF and LF "can't do c−command"
  - this could be because PF and LF have <u>representations</u> over which c−command is not (directly) statable
    - say, prosodic representations and predicate-logical ones, respectively
  - or it could be that PF and LF had the representational means to do c−command in principle —
    - but the finer syntactic structure required to perform this computation was <u>encapsulated</u> away from them
      - say, it was "flattened away" during phasal spellout
  - or it could mean that testing for c−command, *qua* <u>computational process</u>, was simply not the kind of computation PF and LF perform

Non-isomorphic mappings are the norm

- Some (representative?) examples:

  1. *open-class predicate of events* ⇔

     verb ⇔

     host of inflectional morphology

- <u>counter-exemplified by:</u>
  - light-verb constructions (where the open-class predicate is, e.g., a noun)
  - complex tense constructions (w/auxiliary verbs)
  - infinitives
  - etc. etc.

2. *recipient of Agent theta-role* ⇔
   bearer of nominative case ⇔
   noun phrase without marked case morphology

   - <u>counter-exemplified by:</u>
     - passive
     - ECM
     - quirky-case languages (e.g. Icelandic)
     - marked-nominative languages (e.g. Oromo)

2b. *subject of predication* ⇔

bearer of nominative case ⇔

noun phrase without marked case morphology

- <u>counter-exemplified by:</u>
  - ECM
  - quirky-case languages (e.g. Icelandic)
  - marked-nominative languages (e.g. Oromo)
  - HIGH-ABS ergative languages (e.g. Georgian, K'ichean)

3. *change-of-state predicate* ⇔

   unaccusative verb ⇔

   morphologically {simpler / more complex} than causative counterpart

   - <u>counter-exemplified by:</u>
     - stative unaccusatives (e.g. Reinhart 2000)
     - unaccusatives that are morphologically derived from causatives <u>and</u> vice versa, in one and the same language (e.g. Hebrew)

➤ This is not surprising:

- it is, in fact, a central part of why we think morphology, syntax, and semantics are distinct modules in the first place

- There is, to be sure, an overarching tendency towards correspondence among the different modules
  - so, for example:
    - open-class predicates of events are *often* verbs
    - verbs are *often* bearers of inflectional morphology
    - and so forth
  - this is what one would expect of a system that has to, at the end of the day, be *learnable*
- But there is no grounds for an <u>expectation</u> that some empirical domain X will lend itself to (combinatorial) cross-modular correspondence
  - in fact it would be quite unexpected

- Think of it this way:
  - if the mappings *phonetics* ⇔ *phonology* ⇔ *morphology* ⇔ *syntax* ⇔ *semantics* ⇔ *pragmatics* were all isomorphic —
    - there wouldn't be much for linguists to do
    - you could just read everything you need off of **sound** and **usage**

➤ That we have stuff to do is a testament to the non-systematicity of the mappings in question.

Interim summary

- Non-isomorphic cross-modular mappings are the norm

⇒ Therefore, isomorphic mapping in some domain X (e.g. φ−features) is something that needs to be argued for

- Or, if you prefer the conclusion in a more conservative form:
  - isomorphic mappings might be preferable on simplicity & learnability grounds;
  - but given the robust attestation of non−isomorphic mappings —
    - the question of whether domain X involves isomorphic cross-modular mappings is an empirical one;
    - <u>not</u> one to be adjudicated on *a priori* grounds.

- If this all seems fairly anodyne to you, I'm glad;

- But consider:
  - arguments from the behavior of φ−features in one grammatical module are routinely used to adjudicate their behavior in another module
    - to take but one of many examples:
      Nevins 2007 uses facts from the *morphological realization* of 3rd person to argue that 3rd person is (directly) representable in the *syntax*

➤ this line of reasoning is only as sound as the isomorphic-mappings hypothesis.

Accessibility vs. inertness: a diagnostic for feature structure **in syntax**

The diagnostic: gaps in selective targeting

- To begin, I'd like to introduce the diagnostic I'll be using
  - by demonstrating its workings on what is (hopefully) a rather uncontroversial empirical domain

- Consider *wh*−phrases and *wh*−movement:
  - [*wh*] is a feature of certain phrases and (possibly) of certain complementizers
  - in some languages, phrases bearing this feature have a characteristic form
    - though English is probably not such a language (cf. *[hu:]* vs. *[wʌt]*)

- Now, it's fairly easy to convince oneself that whatever featural representation *wh*−phrases bear (say, [+*wh*]) —
  - the complementary featural representation (say, [−*wh*]) is **syntactically inert**
  - i.e., there are no syntactic operations that apply exclusively to <u>non</u>−*wh* phrases

!! This is not (directly) about "binary features" vs. "privative features" or "feature geometries"

- we could model this by saying that "[*wh*] is privative";
- or by saying "no operation can target [−*wh*] only, to the exclusion of [+*wh*]"
  - due to, e.g., some extrinsic markedness hierarchy

- Either way, though, the conclusion is the same: the featural representation borne exclusively by <u>non</u>−*wh* phrases is **syntactically inert**.

- Note also that, with respect to our larger concerns —
  - this is an instance where (at least) syntax and morphology seem to be harmonious with one another:
    - there are languages where all *wh*−phrases bear some characteristic morphology;
    - but even in those languages, non−*wh* phrases do not bear their own characteristic morphology.
  - i.e., in the domain of *wh*, the following appears to hold: **syntactically inert ⇔ morphologically inert**.

Applying the diagnostic to φ−features

- Let us now ask:
  - *which φ−feature configurations are <u>accessible</u>, and which are <u>inert</u>, **in syntax***
- We will investigate this question just as we investigated the corresponding question for [*wh*]

PERSON

CLAIM:

- There are **syntactic** φ−featural relations that target exclusively 1st/2nd person pronouns
  - i.e., bearers of [participant]
- But there are <u>no</u> **syntactic** φ−featural relations that target only lexical noun phrases and 3rd person pronouns, to the exclusion of 1st/2nd person ones

- Here is a case of a probe that seeks exclusively 1st/2nd person targets — a.k.a., *omnivorous agreement* for [participant]
  - from Kaqchikel, a Mayan language of the K'ichean branch

(1) a. ja rat x-**at/*ϕ**-ax-an rja'
FOC you(sg.) COM-**2sg/*3sg.ABS**-hear-AF him
'It was you(sg.) who heard him.'

b. ja rja' x-**at/*ϕ**-ax-an rat
FOC him COM-**2sg/*3sg.ABS**-hear-AF you(sg.)
'It was him who heard you(sg.).'

- NB: Some of you might think that the data we just saw could be afforded a purely “morphological” explanation (cf. Watanabe 2017)
  - but they cannot —
    - morphological analyses are based on the idea that the φ−features of the subject and object are situated in a morphologically-local domain in the first place;
    - this is usually the case in run-of-the-mill transitives; but not always:

(2) x-**at**-*b'e*-**ru**-tz'ët-a' ri a Pedro
COM-**2sg.ABS**-*DIR*-**3sg.ERG**-see-TV DET CL Pedro
'Pedro went to see you.'

- Nevertheless, cases like (2) can still feed omnivorous agreement — compare:

(2) x-**at**-*b'e*-**ru**-tz'ët-a' ri a Pedro
COM-**2sg.ABS**-*DIR*-**3sg.ERG**-see-TV DET CL Pedro
'Pedro went to see you.'

(3) a. ja ri a Pedro x-**at**-*b'e*-tzet-ö
FOC DET CL Pedro COM-**2sg.ABS**-*DIR*-see-AF
'It was Pedro that went to see you(sg.).'

b. ja rat x-**at**-*b'e*-tz'et-ö ri a Pedro
FOC you(sg.) COM-**2sg.ABS**-*DIR*-see-AF DET CL Pedro
'It was you(sg.) that went to see Pedro.'

[thanks to Rodrigo Ranero for help collecting these data]

- This is of course just one example of a probe targeting 1st/2nd person targets to the exclusion of all others;
- It certainly doesn't prove the relevant universal negative;
- But to the best of my knowledge, that negative holds:
  - no language has analogous effects involving a probe that omnivorously seeks 3rd person targets (lexical noun phrases *and* 3rd person pronouns)
    - to the exclusion of 1st/2nd person pronouns

➤ *… what about Nevins 2007?*

- As mentioned earlier, the force of Nevins' (2007) argument that 3rd person is (directly) representable is actually confined to morphological representations only
  - see **omer.lingsite.org/ACTL-Nevins2007.pdf** for how to recoup Nevins' full set of result with no direct reference to 3rd person *in the syntax*

NUMBER

- An analogous case can be made about number
  - concentrating here on the opposition between *singular* and *plural*

CLAIM:

- There are **syntactic** φ−featural relations that target exclusively plural DPs

- Kaqchikel again:

(4) a. ja rje' x-**e/*ϕ**-tz'et-ö rja'
FOC them COM-**3pl/*3sg.ABS**-see-AF him
'It was them who saw him.'

b. ja rja' x-**e/*ϕ**-tz'et-ö rje'
FOC him COM-**3pl/*3sg.ABS**-see-AF them
'It was him who saw them.'

- But there are no **syntactic** φ−featural relations that target only singular DPs
  - though the same caveats about proving a negative universal obviously apply

Interim conclusion

- To the extent that there is indeed no convincing case of *omnivorous 3rd person* or of *omnivorous singular* —
  - and to the extent that we find this diagnostic sound (cf. the [*wh*] case discussed earlier) —

➤ We can conclude:

The featural representations of *3rd person* and of *singular* are <u>syntactically</u> inert.

Cross-modular mismatches

Syntax-morphology mismatches

- We've established that "3rd person" and "singular" — *whatever their particular featural representations are* — are representationally inert in the syntax

⇒ To the extent that we can find cases where "3rd person" and "singular" are categorically <u>not</u> inert in the morphology:

➤ we have conclusive evidence a cross-modular mismatch

- Going back to the different notions of modularity discussed at the outset, we might ask:
  - what kind of mismatch are we talking about?

- There is clearly no issue of <u>information encapsulation</u> here
  - syntax determines what features will go where; morphology doesn't (generally) "sprout" new features;
  - whatever features it gets, it gets from the syntax.

⇒ So, suppose we find that, in the morphology, "3rd person" and/or "singular" are decidedly **not** inert —

- is it an issue of <u>representational</u> differences between syntax and morphology?
- or of <u>computational</u> differences between the two?

- I'll return to this when we discuss the prospects for a Calabresean/Nevinsian approach to these issues
  - the short version: representational and computational alternatives can be traded off with each other fairly easily

- For now, our results will simply be phrased relative to < REPR, COMP > — the joint representational and computational properties of φ−features (in a given module).

English /−z/

- The well known case of φ−feature agreement on English main verbs in the so−called “present simple”
  - /−z/ in *3sg*; ∅ elsewhere

- The single overt cell in the paradigm is arguably the intersection of **multiple** syntactically inert categories

  < 3rd person, singular, nonpast, … >

- Now, given that there is a rule of exponence / vocabulary-insertion rule / … that makes reference to this cell —
  - categories like *3rd person* and *singular*, which are inert in the syntax, cannot be inert in the morphology
  - <u>it follows that:</u>

    < REPR, COMP >$_{syntax}$(φ) ≠ < REPR, COMP >$_{morphology}$(φ)

Number-driven stem suppletion in Hiaki (Harley 2014a,b)

- Harley (2014a,b):
  certain verbs in Hiaki (Uto–Aztecan) supplete based on the number features of their internal argument

(5) a. Aapo/Vempo uka koowi-ta *mea*-k
3sg/3pl the.SG pig-ACC.SG *kill*.SG-PRF
'He/They killed the pig.'

b. Aapo/Vempo ume kowi-m *sua*-k
3sg/3pl the.PL pig-PL *kill*.PL-PRF
'He/They killed the pigs.'

[Harley 2014a:256]

- Harley (2014b:456ff.; see also 2014a:244n26) argues that the “plural form” in a suppletive pair is the default
  - based on the behavior of suppletive verbs in the absence of any number-specified argument

(6) Aman *yahi*-wa/**yevih*-wa
there *arrive*.PL-PASV/**arrive*.SG-PASV
‘Arriving is happening over there.’
or ‘Someone/people/they is/are arriving over there.’

- If this is correct, then the rule triggering the singular form of a suppletive Hiaki verb needs to make reference to <u>singular</u>:

√ARRIVE → *yevih* / [DP_{sg} ___]
→ *yahi* / elsewhere

- The conclusion is the same —
  - *singular*, a syntactically inert category, cannot be morphologically inert
  - thus, once again:

    < REPR, COMP >$_{syntax}$(φ) ≠ < REPR, COMP >$_{morphology}$(φ)

Syntax-semantics mismatches

- It is often assumed that semantically, *plural* is the inert member of the *singular–plural* opposition
  - whereas *singular* means something like *λx.Atomic(x)*
    - Sauerland 2003, *i.m.a.*
      (though see Bale, Gagnon & Khanjian 2011; Martí 2017)

- If this is so, it instantly furnishes another cross-modular mismatch in inertness
  - given that *singular* is the one that is syntactically inert
  - in other words:

$< \text{REPR, COMP} >_{\text{syntax}}(\text{NUM}) \neq < \text{REPR, COMP} >_{\text{semantics}}(\text{NUM})$

Is this just about “markedness”?

(*or:* Is there a Calabresean/Nevinsian alternative?)

Markedness to the rescue?

- Can the idea that, e.g.,
  < REPR, COMP >$_{\text{syntax}}$(φ) = < REPR, COMP >$_{\text{morphology}}$(φ)
  be rescued by an appeal to *markedness*?

- Suppose, for example, that NUMBER in a two-number system (*singular* vs. *plural*) was truly bivalent:
  - plural: [+pl]
  - singular: [−pl]
- Prima facie, there is now a feature (namely [−pl]) which we could then use to construct the (unattested) *omnivorous singular* pattern
- But…

- Suppose we now add a MARKEDNESS HIERARCHY:

  [+pl] ≫ [−pl]

- And we prohibit rules from making exclusive reference to the unmarked member in the hierarchy (in this case, [−pl])
  - cf. Calabrese 1995, 2005 on phonology;
    and Nevins' (2007) adaptation thereof to morphosyntax
    - whereby rules can refer to: **(i)** all values;
      (ii) only contrastive values, or **(iii)** only marked values

- Would this solve our problem?
  - i.e., would it facilitate a model where
    < REPR, COMP >$_{syntax}$(φ) = < REPR, COMP >$_{morphology}$(φ)?

- It seems to me that the answer is no:
  - we saw that, in the morphology, rules of exponence / vocabulary-insertion rules / … need to make exclusive reference to [−pl]
  - we saw that, in the semantics, rules of interpretation (might) need to make exclusive reference to [−pl]

⇒ For this markedness-based approach to work, we'd have to say that:

- rules of exponence / vocabulary-insertion rules / …, and, potentially, rules of interpretation, are not "rules"
- but syntactic agreement is a "rule"

➤ For all I know, this might even be true!

- what it's <u>not</u> is a solution to the problem identified in this talk;
- rather, it's a restatement of the basic claim
  - that, e.g., in the domain of syntax vs. morphology:

$< \text{REPR, COMP} >_{\text{syntax}}(\varphi) \neq < \text{REPR, COMP} >_{\text{morphology}}(\varphi)$

- One possible response is:
  - look for reasons why rules of exponence / vocabulary-insertion rules / … are not "rules" but syntactic agreement is a "rule"
- But I have to say that the prospects here strike me as quite bleak, at the moment
  - for example: there is evidence that in both domains, the relevant "rules" are information-altering *rewrite*-type rules
    - <u>syntax:</u> Preminger 2014, *i.a.*
    - <u>morphology:</u> see Bobaljik 2000, Bonet & Harbour (2012), *i.a.*

- Finally, I'll note here that it is easy to shift the explanatory burden between REPR and COMP in **< REPR, COMP >**
  - e.g.: we *could* assume that syntax (like morphology & semantics) has both [+pl] and [−pl] in its REPRESENTATION
    - but the COMPUTATION in syntax differs from the other two in its inability to reference [−pl]
  - or: we *could* assume (as I have, in other work) that the COMPUTATION in syntax is subject to the same restrictions — e.g. no reference to the *absence* of a feature
    - but that the REPRESENTATION is different in syntax ([pl] vs. [ ]) than it is in the other modules ([+pl] vs. [−pl], or possibly even [−sg] vs. [+sg])

- This sort of analytical leeway should be familiar to you from, e.g., debates on representationalism vs. derivationalism in syntax
  - and I think it's just as inconclusive here as it is there
    - hence the decision to talk about "**< REPR, COMP >**"

Discussion

- We have seen that the way φ−features are treated in syntax (**< REPR, COMP >$_{syntax}$(φ)**) is different from the way they are treated in other modules (e.g. **< REPR, COMP >$_{morphology}$(φ)**)

➤ This is <u>not</u> surprising!

- REMEMBER: Despite what some would have you believe, it is par for the course for the semantics of X to be only a loose fit to the syntax of X, which is only a loose fit to the morphology of X, etc.

- My modest proposal:
  - **Stop treating φ−features as if they should somehow be exempt from this.**

- Concretely, this means:
  - use semantic evidence to adjudicate
    < REPR, COMP >$_{semantics}$(φ)
  - use syntactic evidence to adjudicate
    < REPR, COMP >$_{syntax}$(φ)
  - use morphological evidence to adjudicate
    < REPR, COMP >$_{morphology}$(φ)

    … which is what I have tried to do here.

- And when you see a proposal on “the morphosemantics of X” — check for invalid cross-modular inferences.

A note on language acquisition

- If the **< REPR, COMP >** of φ−features can indeed differ in this fashion, we must ask:
  A. Is **< REPR, COMP >$_X$(φ)** cross-linguistically fixed for every module X of the grammar?
  B. If the answer to (A) is “no”, how is the particular **< REPR, COMP >$_X$(φ)** acquired?

- I see no reason to answer “yes” to (A), especially in light of cases like the English main-verb /−z/
  - since that would involve a massive reduction-to-the-worst-case, where all of < *3rd person, singular, nonpast, …* > would be active & accessible in every language

- Instead, I think it's reasonable to assume that at least in the morphology, there is a default **< REPR, COMP >$_{morph}$(φ)**
  - which the learner only departs from in the face of positive evidence
- It's not hard to imagine what such evidence would be, for a case like English /−z/:
  - a pressure to avoid accidental homophony would lead the learner to conclude that the null exponent must be the 'elsewhere' case
  - and thus, that the *3sg* cell must be active/accessible and not inert
- To the extent that such direct evidence is harder to come by for the learner when it comes to syntax and semantics —
  - **< REPR, COMP >$_{syntax}$(φ)** and **< REPR, COMP >$_{semantics}$(φ)** will be cross-linguistically fixed

➤ <u>BUT:</u>

The demand that these fixed **< REPR, COMP >$_{syntax}$(φ)** and **< REPR, COMP >$_{semantics}$(φ)** be identical to one another is an illegitimate demand

- given that we already know that such strict correspondence is <u>not</u> what we find on the syntax-morphology side.

Thank you!