1 Introduction

(1) *I-language*: a cover term for mind-internal language, as an individual-internal system of “linguistic thought”

(2) *E-language*: a cover term for any form and occurrence of externalized language, used for communicative and other social means

(3) *Externalization*: the mapping from I-language to E-language
   – “not a simple task” Berwick and Chomsky (2011:37)

We know something about E-language, but what does I-language consist of?

(4) *F-selection / the Borer-Chomsky Conjecture* (Baker 2008:156):
   a. Chomsky (2001:10): the Faculty of Language “specifies the features $F$ that are available to fix each particular language $L$ … We adopt the conventional assumption that $L$ makes a one-time selection $[F_L]$ from $F$. These are the features that enter into $L$; others can be disregarded in the use of $L$”
   b. Borer (1984:3): the relevant features belong to “one single component: the inflectional component”.
   c. In Chomsky’s words (2001:2): the variation is “restricted to the lexicon, and … morphological properties, primarily inflectional.”
   d. Chomsky (2007:6): “In addition to Merge …, UG must at least provide atomic elements, lexical items, each a structured array of properties (*features*) to which Merge and other operations apply to form expressions.”

However, it is unclear, from where or how “structured arrays” would enter UG.

(5) UG is *maximally minimal*, not only computationally but also lexically, containing:
   a. A computational faculty, FLN, applying Merge without bounds
   b. The *Universal Lexicon*, UL, of non-composite atomic elements; that is:
      b1 an initial root, $\sqrt{0}$ (Root Zero) = *The Initial Zero Root Thesis*
      b2 an initial functional feature, $F_0$ (Feature Zero) = *The Initial Zero Feature Thesis*

(5b) = The Anti-Lexicalist Hypothesis (or the true lexicalist hypothesis)

We have to think about this “from below” – from the point of view of language evolution (including recent sign languages) and from the point of the newborn (or unborn) individual, who possesses no acquired communication lexicon.
On this view, I- and E-language are non-isomorphically related by ‘(virtual) conceptual necessity.’

2 The growth of language in the individual

(6) *The Copy Theory of Language Growth:* The growth of internal language involves reiterated Copy & Merge of Root Zero and Feature Zero (= $F_0$ / EF) + subsequent specification of their content

(7) *The Generalized Edge Approach:* Both External and Internal Merge are preconditioned by EF ($F_0$)

(8) *Internal language growth (I = Item):*
   a. Stage 1 $\sqrt{1}, \sqrt{2}, ...$ (root carving)
   b. Stage 2 $[I, EF \sqrt{ }] ...$ (I-formation, by free EF merger)
   c. Stage 3 $\{EF, I\}$ (I-extension, by free EF merger)
   d. Stage 4 $\{X, [EF I]\}$ (Item/structure merger)

(9) *Edge Feature Iteration:* Much as a root has to merge with an Edge Feature to build an item, an item has to merge with an additional Edge Feature to build a larger structure (by merging with another item or with a more complex structure).

Symmetric structures cannot be built; that is, the *symmetry problem* (Moro 2000, 2008, Chomsky 2010) vanishes.

Merger of DP and vP: Either one has to serve as a host and merge with $F_0$ – commonly vP:

(10)

Two further instances of Edge Feature Iteration are sketched in (11) and (12).

(11) a. $\overrightarrow{F_0 \text{XP}}$ → b. $\overrightarrow{F_0 \text{XP}}$

(12) a. $\overrightarrow{\text{YP} F_0 \text{XP}}$ → b. $\overrightarrow{\text{YP} F_0 \text{XP}}$

The structures in (11b) and (12b) have the option of either merging with yet another Edge Feature (yielding complex heads) or with a larger structure. In the latter case, the computation either comes to an end (by Stop!) or proceeds by repeated Edge Feature Iteration.

* Adjunct stacking / multiple specifiers
3 I-syntax and The Non-Isomorphy Generalization

Externalization = Materialization + Localization – Materialization being ‘widely undiscussed’

E-language materialization is distinct from I-language materialization, substantiating only an inaccurate and an incomplete representation of I-language rather than ‘copying’ it. – A desirable conclusion, as it suggests that I-languages may be radically more uniform than E-languages, which seems to be essentially correct.

3.1 Abstract Agree vs. overt agreement

(13) There have never been so many trees singing in the rain.

```
P       G
↑______________________________↑
+PL    Agree + PL
```

(14) There should never have been so many trees singing in the rain.

```
P       G
↑______________________________↑
Ø-PL    Agree + PL
```

_The girls is here_ (Henry 1995).

(15) a. They would be bought.  
    b. _Sie würden_ gekauft _werden_.
    they would.3PL bought be
    c. _De skulle bli_ köpta.
    they would be bought.PL
    d. _Þeir mundu verða keyptir._
    they.N.M.PL would.3PL be bought.N.M.PL
    ‘They would be bought.’

That is, E-language materialization does not ‘copy’ I-language materialization, instead substantiating only a (variably) incomplete representation of I-language.
Morphological E-syntax categories, such as Person, Tense, Number, Gender, and Case, do **not** have any direct counterparts in I-syntax.

The mismatch between I- and E-language arises as a **single** functional E-category expresses values (1st person, past tense, etc.) that are the outcome of an I-syntactic process (Agree) that computes and values **two or more** I-categories in relation to each other. The reason behind this is that Agree links phases, vP-internal categories being uninterpretable at the interfaces unless they match and get valued in relation to features of the C-system.

**The Full Interpretation Corollary, FIC**

(16) I-syntax categories cannot by fully interpreted phase-internally, hence the full computation is never phase-bounded: a phase-internal category X must be licensed and computed (interpreted/valued) in relation to at least one phase-external category Y.

**3.2 FIC with respect to Person**

The received understanding (Kaplan 1989, etc.) of indexical pronouns (Schlenker 2003:29):

*Fixity Thesis* (a corollary of Direct Reference): The semantic value of an indexical is fixed solely by the context of the actual speech act, and cannot be affected by any logical operators [Kaplanian ‘monsters’]

However, arguments enter syntax as variables (cf. Kratzer 1998), getting their $\phi$-values by means of matching relations. That is, any clause has silent but active Speaker and Hearer features in its C-system, the *logophoric agent* and the *logophoric patient*, $\Lambda_A$ and $\Lambda_P$, matched by event participants ($\text{NP}_{\alpha\phi}$) in the v-system, via $\phi$-variables in the T-system.

(17) $[\text{CP} \ldots \Lambda_A\Lambda_P \ldots [\text{TP} \ldots \text{Pn} \ldots [\text{vP} \ldots \text{NP}_{\alpha\phi} \ldots]]]$

$\uparrow \quad \uparrow \quad \uparrow$

Agree \hspace{1cm} Agree

**NB.** Agree is a **valuing relation**, Control is an **identity relation** (full or partial – established by movement?)
Pn, thus, enters an Agree relation with NP_{obj}, valuing it as NP_{Pn} or NP_{+Pn}, NP_{+Pn} in turn matching Λ_{A} and Λ_{P}, as shown in (18) and (19) (where the arrow reads ‘gets valued as’).

(18) \( NP_{aPn} \rightarrow NP_{Pn} \) or \( NP_{+Pn} \)

(19) a1. +Pn → +Λ_{A}, −Λ_{P} = 1st person by computation
    a2. +Pn → −Λ_{A}, +Λ_{P} = 2nd person by computation
    a3. +Pn → −Λ_{A}, −Λ_{P} = 3rd person by computation
    b. −Pn: = 3rd person by default ('no person')

On this ‘phase-linking approach’ pronouns are exhaustively computed in syntax.

(20) Ali be Sara goft [ke man tora doost daram]. Persian
    Ali to Sara said that I you friend have.1SG
    a. ‘Ali told Sara that he likes her.’ Shifted = (21)
    b. ‘Ali told Sara that I like you.’ Unshifted = (22)

(21) [CP ... {Λ_{A}}_{i} ... [TP ... <Ali>_{k} ... [CP ... {Λ_{A}}_{k} ... [TP ... <man>_{k} ... i ≠ k]
    \[\text{Control} \quad \uparrow \quad \text{Agree} \quad \uparrow \quad \text{Ali} = \{Λ_{A}\}_{k} \quad \text{man} \rightarrow +Λ_{A} \]

(22) [CP ... {Λ_{A}}_{i} ... [TP ... <Ali>_{k} ... [CP ... {Λ_{A}}_{k} ... [TP ... <man>_{i} ... i ≠ k]
    \[\text{Agree} \quad \uparrow \quad \text{man} \rightarrow +Λ_{A} \]

3.3 FIC with respect to Tense

‘Indicative’ past-in-the-past:

(26) Mary had painted the house green.

(27) [CP ... Λ_{A}-Λ_{P} ... T_{S} ... [TP ... Pn ... T_{R} ... [vP ... T_{E} ... ]] \text{TS} = \text{Speech T}
    \text{T_{R}, Reference T}
    \text{T_{E}, Event T} \]
    \[\text{Agree} \quad \text{Agree} \]
'Subjunctive’ non-past-in-the-past:

(29) *María sagði [að Ólafur væri veikur (*í gær)
Mary said that Olaf were.PST.SBJ sick (*in yesterday).
‘Mary said that Olaf was sick (*yesterday).'

(30) [CP ... T$_{S1}$ ... ... say-T$_{E1}$ ... [CP ... T$_{S2}$ ... ... sick-T$_{E2}$ ...
\[\uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \]
Agree Control Agree
past simultaneous non-past (‘present’)

The features of the speech event (in the sense of Sigurðsson 2004), including $\Lambda_A$-$\Lambda_P$ and T$_S$, are the Kaplanian monsters, present in the C-system of every single clause.

Even languages that largely lack subjunctive morphology, such as English, have the same ‘subjunctive syntax’ as does Icelandic; that is, they also have a monstrous T$_S$, shifted under control in SOT contexts. Evidently, there are ‘subjunctive E-differences’ but no corresponding I-differences between these language types.

Grammatical E-categories are entities that express relations between two or more I-elements (F-atoms). The \textit{Non-isomorphy Generalization}:

(31) The mapping from I-language to E-language is non-isomorphic; that is, there are no one-to-one relations between the minimal building elements of I- and E-language.

5. On externalization

Variation does not arise as a consequence of languages having or not having I-language categories; it arises as languages externalize these categories differently.

5.1 Person computation vs. person agreement

(17) [CP ... $\Lambda_A$-$\Lambda_P$ ... [TP ... Pn ... [vP ... NP$_{\phi \ldots}$]]]
\[\uparrow \uparrow \uparrow \uparrow \]

The syntax operates on atomic elements (NP_{np}, Pn, Λ_A, Λ_P), while overt ϕ-agreement operates with a ‘compact unit’: the value (e.g., 1st person) that results from NP_{np} matching Λ_A, Λ_P.

5.2 Tense computation vs. tense agreement

(29) Mary said that Olaf were.PST.SBJ sick (*in yesterday).

(30) [CP … T_{S1} … say-T_{E1} … [CP … T_{S2} … sick-T_{E2} …
  \[ \uparrow \uparrow \uparrow \uparrow \uparrow \] \[ \downarrow \downarrow \downarrow \downarrow \] ]
  Agree Control Agree
  past simultaneous non-past (‘present’)

Uninterpretable Tense Agreement; even found in some infinitival ECM complements:

(32) a. María segist munu fara. [PRES.IND – PRES.INF]
  Mary says-herself will.INF go
  ‘Mary says that she will go/leave.’

b. María sagðist mundu fara. [PST.IND – PST.INF]
  Mary said-herself would.INF go
  ‘Mary said that she would go/leave.’

Non-SOT languages and split SOT languages do not apply Tense Agreement here, instead using the simple present tense, as illustrated for Russian and Japanese in (33) and (34).

(33) Tanja skazala [čto ona tancuet]. [Russian]
  Tanja said that she dances (Comrie 1986:275)
  ‘Tanja said that she was dancing
  (at the moment of Tanja’s saying so).’

(34) Taroowa [Hanakoga Siatoruni iru] to itta. [Japanese]
  Taro Hanako Seattle-in is that said (Ogihara 1996:5)
  ‘Taro said that Hanako was in Seattle
  (at the moment of Taro’s saying so).’

However, the same tense interpretation: ‘non-past relative to the past saying in the matrix clause’.
Evidently, in Icelandic morphology, +PAST is silently copied onto $T_{S2}$ under control and spelled out on the verb in $T_{E2}$ under Agree with $T_{S2}$ – not so in Russian/Japanese.

Tense Agreement (overt SOT) behaves like a reflex of sorts, utilizing a syntactic control relation between $T_{E1}$ and $T_{S2}$ as a kind of a path or a gateway to pass down the morphological tense value from the matrix verb (established under $T_{S1}$/T$_{E1}$ Agree/valuation).

Strikingly, the value in examples like (29) is shifted (+PAST), while the syntactic control relation between $T_{E1}$ and $T_{S2}$ establishes an unshifted identity relation (as control relations generally do; here, the identity is temporal ‘sameness’, simultaneity).

Tense Agreement is thus distinct from the I-syntactic matching processes that yield tense interpretation.

- Operates with a non-syntactic feature (here +PAST)
- Utilizes the Control/Agree ‘path’ in a directional top-down externalization process.

5.3 Case agreement

PF top-down also seen in meaningless case agreement of PRO

(35) Ólafur fór bara einn*/einan í veisluna.
Olaf.NOM went just alone.NOM*/ACC to party.the
‘Olaf just went alone to the party.’

(36) Hún bað Ólaf [að PRO fara bara einan í veisluna].
she.NOM asked Olaf.ACC to go just alone.ACC to party.the
‘She asked Olaf to just go alone to the party.’

(37) [CP … Olaf.ACC … [CP … PRO … alone.ACC … ]

↑_________________________↑↑___________↑

Control Agree(ment)

Cf. also case concord:

(38) Við hittum [alla þessa fjóra frægu leikara].
we met all these four famous actors
N A.M.PL A.M.PL A.M.PL A.PL.DEF A.M.PL
Participle agreement:

(39) a. *Þeir voru kosnir.*
    the.N.M.PL were.3PL elected.N.M.PL

b. *Þær voru kosnar.*
    they.N.F.PL were.3SG elected.N.F.PL

‘They were elected.’

The syntactic Event/Participant relation gets reflected by morphological non-syntactic elements in Icelandic while it does not in English (and many other languages).

The I-/E-split also evidenced by case assignment to arguments (Maling 2001, Sigurðsson 2012, a.o).

(40) A few Dat taking Icelandic verbs – corresponding verbs in German take Acc

    dreifa ‘spread’, fagna ‘welcome’, fleygja ‘throw away’, fleyta ‘float’, fljúga ‘fly’ (e.g. an aeroplain),
    ríða ‘ride’ (e.g. a horse), róa ‘row’, sigla ‘sail’, snúa ‘turn’, stjórna ‘control, govern, rule’,

(41) A few Gen taking Icelandic verbs – corresponding verbs in German take Acc

    krefjast ‘demand’, leita ‘look for, search for’, minnast ‘(be able to) remember’, neyta ‘consume’,
    njóta ‘enjoy’, óska ‘wish for’, sakna ‘miss’, æskja ‘wish for’, þarfust ‘need’

case is an externalization strategy, relating only indirectly to underlying semantic/syntactic distinctions within individual E-languages–not across languages. Case variation non-parametric?
However: The ‘ingredients’ in case systems – ‘well-behaved chaos’:

(42) a. NPs are event licensed by heads: Voice, Appl, v, p, n, a, …
    b. Event licensers of arguments may be case star augmented (cf. Chomsky 2001)
    c. The augmented case star may be variably marked: *, *+, *++
    d. Nom = syntactically a non-case (regardless of its overt marking)

**Accusativity:**

(43) … [TP .. Voice*/₃₃ .. [ NP₂ .. v*–V NP₁ ..]] → NP₁/ACC in PF

NP₂/Ø (=’Nom’) in PF

**Ergativity:**

(44) .. [TP .. Voice*/₃₃ .. [ NP₂ .. v–V NP₁ ..]] → NP₂/ERG in PF

NP₁/Ø (=’Nom’/’Abs’) in PF

Historically, individual cases presumably ‘begin life’ as category specific: Erg as a Voice case, Acc as a v case, Gen as an n case, Dat as Appl case, etc.

Cases commonly behave like externalization (PF) viruses, invading or contaminating additional category domains, Gen and Dat invading the v-system, etc.

The Icelandic case system (much of it):

(45) |    | 0 | * | *+ | *++ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>v system</td>
<td>Nom</td>
<td>Acc</td>
<td>Dat</td>
</tr>
<tr>
<td>b.</td>
<td>Voice system</td>
<td>Nom</td>
<td>Dat (?)</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>Appl system</td>
<td>–</td>
<td>Dat</td>
<td>Acc</td>
</tr>
<tr>
<td>d.</td>
<td>n system</td>
<td>–</td>
<td>Gen</td>
<td>Dat</td>
</tr>
<tr>
<td>e.</td>
<td>p system</td>
<td>–</td>
<td>Acc</td>
<td>Dat</td>
</tr>
<tr>
<td>f.</td>
<td>a system</td>
<td>Nom</td>
<td>Dat</td>
<td>Gen</td>
</tr>
</tbody>
</table>

Variable marking of event licensors (v, v*, v**, etc.) = E-parameters, unlinked to UG but yielding E-variation. Seemingly chaotic but ‘systematically regulated.’
Deep transfer, from I-syntax to the semantic interface and deep PF or E-syntax (including abstract morphology)

Surface transfer, from fully computed E-syntax to phonology and phonetics.

With Merge and deep transfer in place, the individual makes use of whatever surface transfer modes that are independently available to the organism: audible, visible, or tactile. ‘Lexical PF insertion’ takes place under surface transfer (with I-syntax still in hindsight); hence, it cannot apply to I-syntactic structures, be it by insertion into terminal nodes or by phrasal spell-out.

I-language relations, including Person and Tense computation, pronominal indexical shift, perspective or ‘speech’ time shift in Sequence of Tenses contexts, Event/Participant relations, and so on, do not seem to vary. Rather, what varies is how these relations are expressed by the units of E-languages, including morphological E-categories, such as Person, Number, Gender, Case and Tense.

Some references


