On Hilberts epsilon operator, pair merge, and the source of asymmetry in adjunction

Chenchen Song

School of International Studies
Zhejiang University

Young Scholars' Symposium, Hangzhou
June 11, 2021

Introduction

In his 2019 UCLA lectures Chomsky recharacterized pair merge in terms of Hilbert's ε-operator.

> *"The simplest operation after... set formation is... pair formation, so we need an operation, pair merge, which will... apply to [adjuncts].... [T]here are unboundedly many dimensions... [and you] can add any number of adjuncts.... [Moreover,] it's not just a set of paired things [but] a sequence... [W]hat we have is a situation where... you generate a finite set [and] form from that set a sequence. It could be any sequence.... You pick one... [and] merge [it] into the construction. This operation of picking a particular element out of the set of sequences is David Hilbert's epsilon operator, which picks a single thing out of a set. It was part of his work on the foundations of mathematics, [a] basic operation." (Chomsky 2019, UCLA Lecture 4)*

Introduction

Pair merge Chomsky's way to derive adjunction since 2000

Merge the basic combinatorial operation in the minimalist program

Essentially two types of merge:

Set merge $(\alpha, \beta) \to \{\alpha, \beta\}$ (a plain set, symmetric)

Pair merge $(\alpha, \beta) \to \langle\alpha, \beta\rangle$ (an ordered pair, asymmetric)

For example:

- SetMerge(eat, cookies) $\vdash$ $[_{\text{VP}}\ [_{\text{V}}$ eat $]\ [_{\text{NP}}$ cookies $]]$
- PairMerge(quickly, eat) $\vdash$ $[_{\text{VP}}\ [_{\text{AdvP}}$ quickly$]\ [_{\text{VP}}$ eat $]]$

Introduction

Chomsky (2000 et seq.) resorts to pair merge to produce the inherent asymmetry in adjunction, which is essentially an additional axiom.

There have been quite a few objections to this approach (see Song 2019 for an overview), but that's not the focus of this talk.

My goal today

Explore how the ε-based new approach changes the picture.

Plan

History

Russell's inverted iota for definite descriptions (Whitehead & Russell 1910):

$$\imath x.F(x)$$

"the unique x such that $F(x)$ is true (i.e., the F)"

Example: $\imath x.\text{King-of-France}(x) =$ "the King of France"

Inspired by Russell, Hilbert invented two generic elements: τ (1922, 1923) and ε (1925, 1926). The latter made it into Hilbert & Bernays (1939).

Hilbert's generic symbols (aka ideal objects):

- $\tau x.F(x)$ The τ-term has property F when *every* individual does so.
- $\varepsilon x.F(x)$ Whenever *some* individual has property F, the ε-term does.

History

Hilbert's main goal was to find a consistent and complete axiom set (i.e., a solid basis) for mathematics, first and foremost for arithmetic. His "basis" was the ε-calculus, an extension of predicate calculus.

Hilbert's program failed due to Gödel's (1931) incompleteness theorems:

First there's no consistent axiomatic system for arithmetic

Second no system can prove its own consistency

But Hilbert's endeavor left us a number of valuable results, including the ε operator. For instance, it can replace $\imath$ and help eliminate $\exists$ and $\forall$.

The ε-calculus (syntax)

The ε-calculus is obtained by adding two components to the first-order predicate calculus:

- a new logical constant ε forming a term $\varepsilon x.F(x)$ from a formula F[1]
- a new axiom for each ε-term of the form

$$F(t) \to F(\varepsilon x.F(x)) \qquad \text{(axiom } \varepsilon\text{)}$$

"if any term t has the property F at all, then $\varepsilon x.F(x)$ does"

NB: ε-terms are **nondeterministic** by definition.

[1] As such, the ε-symbol is strictly speaking a "subnector" (Curry 1963). I keep using the misnomer "operator" since it is already prevalent.

The ε-calculus (semantics)

Hilbert merely used ε as a syntactic tool. Asser (1957) proposed the first model-theoretic semantics for it (see Leisenring 1969), and von Heusinger (1995) further developed it in the context of linguistics.

Asser's idea was to interpret the ε-symbol by a choice function. More generally, there's a clear analogy between axiom ε and the axiom of choice.

$$\forall X.(\varnothing \notin X \Rightarrow \exists f{:}X \to \bigcup X.(\forall A \in X.f(A) \in A)) \qquad \text{(axiom of choice)}$$

"For any set X of nonempty sets, there exists a function f defined on X, such that for every set A in X, $f(A)$ is an element of A."

The ε-calculus (semantics)

Asser extended the usual model of first-order predicate calculus with a choice function Φ to interpret the ε symbol.

$$\mathcal{M} \coloneqq \langle \mathcal{J}, \mathcal{I}, \Phi \rangle$$

"The model $\mathcal{M}$ for the ε-calculus is defined by a mathematical structure $\mathcal{J}$, an interpretation function $\mathcal{I}$, and a choice function Φ."

> Axiom ε essentially says that for any nonempty subset of the universe of discourse, we can choose a particular element from it.

Hence, the ε-operator is sometimes also called the choice operator.

The ε-calculus (semantics) ✂

Asser also considered the empty set. Below is his total-function suggestion:

The ε-operator is interpreted in $\mathcal{M}$ by a function

$$\Phi\colon \mathbf{P}(\mathcal{J}) \to \mathcal{J}$$

from the power set of the universe of discourse $\mathcal{J}$ to $\mathcal{J}$ itself such that for each subset A of $\mathcal{J}$

$$\Phi(A) = \begin{cases} a \in A, & A \neq \varnothing \\ j \in \mathcal{J}, & A = \varnothing \end{cases}$$

"the value of $\Phi(A)$ is a particular element a of A if A is nonempty and an arbitrary element j of $\mathcal{J}$ if A is empty"

The ε-calculus (semantics) ✄

For example, if A is

- Apple, then $\Phi(A)$ is just a particular apple (i.e., a representative)
- Round-Square, then $\Phi(A)$ is an arbitrary element of $\mathcal{J}$, such as a square that looks a bit round or some artwork named "Round Square"

Slater (2017): if there's no such x that satisfies $F(x)$, then the denotation of $\varepsilon x.F(x)$ "is a fiction, which means it is simply a pragmatically chosen individual in the whole world at large."

The ε-calculus (semantics) ✄

In particular, this total-function interpretation of ε helps us understand the eliminability of $\exists$ and $\forall$ in ε-calculus better.

$$\begin{aligned} \exists x.F(x) &\equiv F(\varepsilon x.F(x)) \\ \forall x.F(x) &\equiv F(\varepsilon x.\neg F(x)) \end{aligned}$$

$\exists$ can be directly reformulated by ε, while $\forall$ can be reformulated by ε and $\neg$. If F is true for everything, then $\neg F$ is true for nothing, and $\varepsilon x.\neg F(x)$ picks out an arbitrary element from the whole world, for which F is true.

Linguistics (von Heusinger's application)

Hilbert's ε-operator was first applied in linguistics in the area of semantics. A representative line of research is that of von Heusinger and colleagues since the 1990s on (in)definite NPs and intersentential anaphora (see, i.a., Egli & von Heusinger 1995; von Heusinger 1997, 2013; von Heusinger & Egli 2000; and von Heusinger & Kempson 2004).

- (in)definite NPs: *the man*, *a man*, etc.
- intersentential anaphora: *A man comes. He whistles.*

Dissatisfied with previous approaches (e.g., $\imath$, generalized quantifier, E-type pronoun), they proposed ε-based analyses for the above phenomena.

NB: von Heusinger's application of ε is **semideterministic**.

Linguistics (von Heusinger's application)

In a nutshell, the proposals of von Heusinger et al. are as follows.

phenomenon	syntax	semantics
definite NP	indexed ε-term (global)	global choice function (Φ)
indefinite NP	indexed ε-term (local, η)	local choice function (f)
anaphora	ε/η coreference	choice function updating

For example (based on von Heusinger 1997, 2013),

- $[\![$*the man*$]\!] \equiv [\![\varepsilon_i x.\mathrm{Man}(x)]\!] = \Phi_i(\mathrm{Man})$
- $[\![$*a man*$]\!] \equiv [\![\eta_1 x.\mathrm{Man}(x)]\!] = f_1(\mathrm{Man})$
- $[\![$*A man comes. The man whistles.*$]\!]$
  $\equiv [\![\mathrm{Come}(\eta_1 x.\mathrm{Man}(x)) \wedge \mathrm{Whistle}(\varepsilon_j x.\mathrm{Man}(x))]\!]$
  $= \mathrm{Come}(f_1(\mathrm{Man})) \wedge \mathrm{Whistle}(\Phi_j(\mathrm{Man}))$
  with $f_1(\mathrm{Man}) = d, \Phi_j = \Phi_{i\langle\!\langle [\![\mathrm{Man}]\!]^{\mathcal{M},g}/d \rangle\!\rangle}$, and $\Phi_j(\mathrm{Man}) = d$

U	the utterance
i	pre-U discourse
j	post-U discourse
1	local context
$\langle\!\langle\ \rangle\!\rangle$	update

Linguistics (von Heusinger's application)

A few issues reflected in von Heusinger et al.'s linguistic application of ε:

1. ε is treated as a syntactic symbol, though not in the "familiar" sense.
2. It probably belongs to the logical form level.
3. It corresponds to determiners (e.g., *the*, *a*) in the familiar sense.
4. Its definition is more relaxed than that in mathematics (e.g., indexing, optional totality).
5. Its foundational status largely matches that in mathematics (i.e., axiom of choice).

In sum, this line of application relates ε to **three** representation levels: grammatical, logical, mathematical.

☝ determiner ☝ ε ☝ choice function

Plan

Chomsky's application: Main thesis

Does Chomsky's application of Hilbert's ε-operator also have multilevel theoretical relevance?

Chomsky's main thesis: The ε-based pair merge theory provides a unified solution for recursive adjunction and coordination.

(1) a. *Tom is young, tall, and happy.*

b. *Tom is a young, tall, and happy guy.* (adapted from Chomsky 2019)

> *"So, what we have is a situation where in order to generate these objects, you generate a finite set, and then you form from that set a sequence. It could be any sequence of elements, and there's in fact infinitely many possible sequences. You pick one out of those, and that sequence, call it S, is the thing that you are then going to merge into the construction. This operation of picking a particular element out of the set of sequences is David Hilbert's epsilon operator, which picks a single thing out of a set." (Chomsky 2019, UCLA Lecture 4)*

Chomsky's application: Key features

Three key features of Chomsky's new theory:

1. It assumes a full-fledged multidimensional structure.

   "[P]air merge [applies] to the simple adjunct case like young man. Young [is] attached to man, but you don't see it in the labeling because it's off in some other dimension. And the unbounded unstructured cases show [that] there are unboundedly many dimensions. . . ."

2. It has a default, built-in recursive configuration.

   "That's the basic object that gives you an unbounded coordination, when you get down to just one case. . . that's just plain adjunction. . . ."

3. It makes use of a "link" component.

   "So we have a sequence of things with a link, and. . . all the links have to be identical, because. . . you are attaching everything to the same point. . . . [E]very member of this coordinated construction. . . is individually predicated of what it links to."

Chomsky's application: Questions

Let us go through this step by step in slightly more formal terms:

1. $\{young,\ tall,\ happy\}$
2. $\Sigma = \{\langle young,\ tall,\ happy\rangle, \langle young,\ happy,\ tall\rangle, \langle tall,\ young,\ happy\rangle, \ldots\}$
3. $S = \varepsilon x.\Sigma(x) = \langle young,\ tall,\ happy\rangle$
4. $S' = \langle \mathrm{CONJ}, \langle young, L\rangle, \langle tall, L\rangle, \langle happy, L\rangle\rangle$

Chomsky's application: Questions

Let us go through this step by step in slightly more formal terms:

1. $\{young, tall, happy\}$
2. $\Sigma = \{\langle young, tall, happy\rangle, \langle young, happy, tall\rangle, \langle tall, young, happy\rangle, \ldots\}$
3. $S = \varepsilon x.\Sigma(x) = \langle young, tall, happy\rangle$
4. $S' = \langle \mathrm{CONJ}, \langle young, L\rangle, \langle tall, L\rangle, \langle happy, L\rangle\rangle$

Questions:

- Where in the grammar are Steps 2–4 encoded?
- What linguistic element encodes ε?
- What is the "link" (L) in Step 4?
- How exactly do we obtain S' from S?
- What happens after Step 4?
- Do we really need pair merge for all this to work?

Tentative answers I

1. $\{$*young, tall, happy*$\}$
2. $\Sigma = \{\langle young,\ tall,\ happy\rangle, \langle young,\ happy,\ tall\rangle, \langle tall,\ young,\ happy\rangle, \ldots\}$
3. $S = \varepsilon x.\Sigma(x) = \langle young,\ tall,\ happy\rangle$
4. $S' = \langle \mathrm{CONJ}, \langle young, L\rangle, \langle tall, L\rangle, \langle happy, L\rangle\rangle$

I speculate Steps 2–3 are at the end of the lexical (sub)array stage.

Tentative answers I

1. $\{young,\ tall,\ happy\}$
2. $\Sigma = \{\langle young,\ tall,\ happy\rangle, \langle young,\ happy,\ tall\rangle, \langle tall,\ young,\ happy\rangle, \ldots\}$
3. $S = \varepsilon x.\Sigma(x) = \langle young,\ tall,\ happy\rangle$
4. $S' = \langle \mathrm{CONJ}, \langle young, L\rangle, \langle tall, L\rangle, \langle happy, L\rangle\rangle$

I speculate Steps 2–3 are at the end of the lexical (sub)array stage.

Rationale:

- They are unlikely to be part of syntactic derivation proper (no derivation rules can generate infinite sets or perform the choice function).
- But they aren't outside syntax either (Step 1 is a lexical subarray and syntactic rules are procedural).
- Step 4 certainly is part of derivation proper (it involves pair merge).

Tentative answers I

1. $\{young,\ tall,\ happy\}$
2. $\Sigma = \{\langle young,\ tall,\ happy\rangle, \langle young,\ happy,\ tall\rangle, \langle tall,\ young,\ happy\rangle, \ldots\}$
3. $S = \varepsilon x.\Sigma(x) = \langle young,\ tall,\ happy\rangle$
4. $S' = \langle \text{CONJ}, \langle young, L\rangle, \langle tall, L\rangle, \langle happy, L\rangle\rangle$

> I speculate Steps 2–3 are at the end of the lexical (sub)array stage.

So, in Chomsky's application ε serves to instruct the derivation by generating some extra (ordering) information on top of lexical subarrays.

Tentative answers II

Suppose Chomsky's application of ε is at the preparatory state of syntactic derivation, **What lexical item encodes it?** It must be lexically encoded due to "inclusiveness" (Chomsky 1995).

$\boxed{\text{I speculate that } \varepsilon \text{ is encoded in CONJ (overt or covert).}}$

Rationale:

- The ordering of the set is imposed by the sequential nature of conjunction.
- Conjunction is signaled by CONJ (which should be present throughout).
- The interconjunct ordering (i.e., the product of ε) doesn't include CONJ.
- This makes CONJ a suitable operator taking the conjunct set as argument.

Tentative answers III

Re the identity of the link, Chomsky has the following remark:

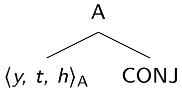
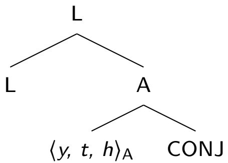
> *"What is L?... Let's take the simplest case: noun phrase and verb phrase coordination, you know, John, Bill, Tom, the young man, etc.... they read the book, walked to the store, and so on.... Whats the linking in those cases? Well, the assumption that comes to mind right away is that L is just n... but that... can't be identified with the [nominalizer]... because this one is just much higher up.... Same with verb: ... the [verbalizer] can't be the same as what we call small v... high up in the derivation. (Chomsky 2019, UCLA Lecture 4)"*

So, we can assume that the link is simply the **syntactic category** the conjunction object merges with, though it still isn't clear what the object in Step 4 really is (e.g., Is it syntactic or logical/semantic?).

The remaining Qs can only be answered with some technical details!

Derivation

Below I work out (a potential version of) the technicalities.

Step	Representation	Rule
1	$\{$CONJ, *young*, *tall*, *happy*$\}$	lexical subarray formation
2	CONJ($\{$*young*, *tall*, *happy*$\}$)	function application (part of CONJ-rule)
2	$\{\langle$*young*, *tall*, *happy*$\rangle, \ldots\}$	sequence set formation (part of CONJ-rule)
3	$\langle$*young*, *tall*, *happy*$\rangle$	ε-rule (part of CONJ-rule)
3	$\{$CONJ, $\langle$*young*, *tall*, *happy*$\rangle\}$	lexical subarray updating (part of CONJ-rule)
4		set merge (derivation proper begins here)
4		set merge

NB: Steps 1–3 are not lexical rules. In minimalist syntax certain rules are required for the lexical array- or numeration-forming stage anyway.

Derivation

A few issues to note:

- $\langle y, t, h \rangle_{\mathrm{A}}$ is labeled A since this example involves adjectival conjuncts.
- CONJ is a nonlabeling category (see Chomsky 2013, Song 2019).
- I assume different derivational details for the attributive and the predicative use of adjectival conjuncts (where the identity of L differs too):
  - Predicative: the conjuncts are not AP's but PredP's (renumeration or similar rule needed, again required in minimalism anyway).
  - Attributive: this is the only adjunction case (more on this later).
- The conjuncts aren't strictly required to enter the derivation as a *tuple*. They may also be in a *set*, with the ordering being part of the environmental info of the derivation. What matters here is merely that the ordering (and its associated asymmetry) is generated by ε rather than merge.

Derivation: Interfaces

The environmental info of the derivation (i.e., the ε-generated ordering) becomes useful at the stage of interface interpretation.

PF $\langle y, t, h \rangle_{\mathrm{A}}$ directly determines the phonological ordering, with a separate rule (perhaps the PF-aspect of the CONJ-rule) inserting the exponent of CONJ before the last conjunct (but this rule may be language-specific).

Derivation: Interfaces

The environmental info of the derivation (i.e., the ε-generated ordering) becomes useful at the stage of interface interpretation.

PF $\langle y,\ t,\ h\rangle_{\mathrm{A}}$ directly determines the phonological ordering, with a separate rule (perhaps the PF-aspect of the CONJ-rule) inserting the exponent of CONJ before the last conjunct (but this rule may be language-specific).

LF $[\![\{\mathrm{CONJ},\ \langle y,\ t,\ h\rangle_{\mathrm{A}}\}_{\mathrm{A}}]\!] = [\![\mathrm{CONJ}]\!]([\![\langle y,\ t,\ h\rangle_{\mathrm{A}}]\!]) = \bigwedge(\langle y', t', h'\rangle) = y' \wedge t' \wedge h'$

The semantics of CONJ (i.e., $\bigwedge$) itself doesn't use the ordering info.

I assume that the ε-generated ordering serves as a discourse-level tool at C-I instead, which can be referenced by pronoun indexing, event sequencing, the interpretation of certain adverbs (e.g., *respectively*), and so on.

This makes $\langle \mathrm{CONJ}, \langle \mathit{young}, L\rangle, \langle \mathit{tall}, L\rangle, \langle \mathit{happy}, L\rangle\rangle$ in Step 4 look like a high-level syntacticosemantic representation. It doesn't reflect the actual derivation/composition but just reflects the fact that each conjunct eventually combines with L: $y' \wedge t' \wedge h'(\mathrm{LP}') = y'(\mathrm{LP}') \wedge t'(\mathrm{LP}') \wedge h'(\mathrm{LP}')$.

Derivation: Attributive

The attributive case (typical adjunction):

(2) *Tom is a young, tall, and happy guy.* (=1a)

Two routes:

1. *à la Chomsky, $L = n$*
   $\text{PairMerge}(\{\text{CONJ}, \langle y,\ t,\ h\rangle_{\text{A}}\}, guy_{\text{N}}) = \langle\{\text{CONJ}, \langle y,\ t,\ h\rangle_{\text{A}}\}, guy_{\text{N}}\rangle_{\text{N}}$
2. *à la Song 2019, $L = Cat$*
   $\text{SetMerge}(\text{SetMerge}(\text{Cat}, \{\text{CONJ}, \langle y,\ t,\ h\rangle_{\text{A}}\}), guy_{\text{N}})$
   $= \text{SetMerge}(\{\text{Cat}, \{\text{CONJ}, \langle y,\ t,\ h\rangle_{\text{A}}\}\}_{\text{Cat}}, guy_{\text{N}})$
   $= \{\{\text{Cat}, \{\text{CONJ}, \langle y,\ t,\ h\rangle_{\text{A}}\}\}_{\text{Cat}}, guy_{\text{N}}\}_{\text{N}}$

The Cat-route dispenses with pair merge and derives adjunction via set merge plus agreement on categorial features instead.

Digression: Song's (2019) adjunction theory

Song derives adjunction via set merge plus agreement on categorial features, done by an independently motivated "defective category" (Cat).

Long story short, Cat is a categorially underspecified functional category that gets assimilated into any fully specified category it merges with via agreement. This results in an extension in structure but not in label, and whatever material Cat carries along from a previous cycle[2] forms a two-segment category with its host in the main cycle.

[2] Moreover, this is a root categorization cycle, where the $\sqrt{\mathrm{Comp}}$ part is either a root or a root-like chunk ("atomized" in an even earlier cycle). See Song (2019) for details.

Interim summary

The version of technical details presented here shows that Chomsky's ε-based treatment of recursive coordination/adjunction is **not obligatorily** accompanied by pair merge (i.e., these are separate theories).

The inherent asymmetry of adjunction may also be derived in a pair-merge-free fashion (e.g., Song 2019).

Interim summary

The version of technical details presented here shows that Chomsky's ε-based treatment of recursive coordination/adjunction is **not obligatorily** accompanied by pair merge (i.e., these are separate theories).

The inherent asymmetry of adjunction may also be derived in a pair-merge-free fashion (e.g., Song 2019).

Chomsky's application of ε can be related to the three representation levels in von Heusinger's application too:

- grammatical (lexicosyntactic)
- logical (ε)
- mathematical (choice function)

NB: Chomsky's application of ε is also **semideterministic**.

Plan

A case study

Chinese coordinative compounds provide an interesting case to illustrate (and complement) the foregoing theory.

- They involve root coordination under the level of categorization (so L = categorizer), which Chomsky didn't mention.
- Their distinctions from homophonous phrasal coordination demonstrate the difference between the two types of little x's Chomsky mentioned.
- They provide a scenario where the ε-choice is phonologically informed (vis-à-vis Chomsky's examples, where it is semantically informed).

In the final section of this talk I briefly illustrate the above points.

Facts

Chinese coordinative compounds have the following basic characteristics.

1. They can consist of any number of items.

   e.g., *dà-xiǎo* ‘big-small; size’, *gāo-fù-shuài* ‘tall-rich-handsome; a perfect guy’,

   yī-shí-zhù-xíng ‘clothes-food-residence-travel; everyday activities’,

   chái-mǐ-yóu-yán-jiàng-cù-chá ‘firewood-rice-oil-salt-sauce-vinegar-tea; daily life details’

Facts

Chinese coordinative compounds have the following basic characteristics.

1. They can consist of any number of items.

   e.g., *dà-xiǎo* 'big-small; size', *gāo-fù-shuài* 'tall-rich-handsome; a perfect guy', *yī-shí-zhù-xíng* 'clothes-food-residence-travel; everyday activities', *chái-mǐ-yóu-yán-jiàng-cù-chá* 'firewood-rice-oil-salt-sauce-vinegar-tea; daily life details'

2. They are mostly idiomatic (e.g., the above examples).

Facts

Chinese coordinative compounds have the following basic characteristics.

1. They can consist of any number of items.

   e.g., *dà-xiǎo* 'big-small; size', *gāo-fù-shuài* 'tall-rich-handsome; a perfect guy', *yī-shí-zhù-xíng* 'clothes-food-residence-travel; everyday activities', *chái-mǐ-yóu-yán-jiàng-cù-chá* 'firewood-rice-oil-salt-sauce-vinegar-tea; daily life details'

2. They are mostly idiomatic (e.g., the above examples).

3. They are generally categorially flexible.

   e.g., (i) *Zhè shì yí-gè chānzá le* *chái-mǐ-yóu-yán-jiàng-cù-chá*$_{\mathrm{N}}$ *de píngfán de àiqíng gùshi.*
   'This is an ordinary love story filled with daily details.' vs.

   (ii) *Yì-qiān-jù qíng-huà, dōu bù-rú hé nǐ yìqǐ* *chái-mǐ-yóu-yán-jiàng-cù-chá*$_{\mathrm{V}}$.
   'A thousand romantic words is not as good as having an ordinary life together with you.'

While 1 is also true for phrasal coordination, 2–3 are compound-specific.

Facts

Chinese coordinative compounds have the following basic characteristics.

4. They often allow flexible internal ordering.

   e.g., *dài-tì*/*tì-dài* 'replace-replace; replace', *xún-zhǎo*/*zhǎo-xún* 'search-search; search', *gāo-fù-shuài*/*gāo-shuài-fù* 'a perfect guy', *fēng-shuāng-yǔ-xuě*/*yǔ-xuě-fēng-shuāng* 'wind-frost-rain-snow/rain-snow-wind-frost; hardships of journey or life'

The ordering is determined by several factors (Chen 2008, Xu 2016, Hsieh 2021), with prosody being a most important one (e.g., level tone $<$ oblique tone). Other factors include semantics, stylistics, and encyclopedic info.

Facts

Chinese coordinative compounds have the following basic characteristics.

4. They often allow flexible internal ordering.

   e.g., *dài-tì/tì-dài* 'replace-replace; replace', *xún-zhǎo/zhǎo-xún* 'search-search; search', *gāo-fù-shuài/gāo-shuài-fù* 'a perfect guy', *fēng-shuāng-yǔ-xuě/yǔ-xuě-fēng-shuāng* 'wind-frost-rain-snow/rain-snow-wind-frost; hardships of journey or life'

The ordering is determined by several factors (Chen 2008, Xu 2016, Hsieh 2021), with prosody being a most important one (e.g., level tone $<$ oblique tone). Other factors include semantics, stylistics, and encyclopedic info.

Synchronically there's usually a standardized ordering for each word, but the flexibility is still there, which distinguishes co-compounds from multisyllabic roots, whose internal ordering is strictly fixed.
e.g., *pánghuáng/*huángpáng* 'hesitate', *zhǎnzhuǎn/*zhuǎnzhǎn* 'toss and turn'

Phrasal coordination also has flexible ordering, but that's usually semantically determined (or totally stochastic) and not subject to standardization or clear acceptability variation (except for idioms).

Analysis

The ε-based analysis for coordinative compounds is straightforward. Take *gāo-fù-shuài* 'tall-rich-handsome; a perfect guy' for example.

Step	Representation	Rule
1	$\{\text{CONJ}, \sqrt{\text{GĀO}}, \sqrt{\text{FÙ}}, \sqrt{\text{SHUÀI}}\}$	lexical subarray formation
2	$\text{CONJ}(\{\sqrt{\text{GĀO}}, \sqrt{\text{FÙ}}, \sqrt{\text{SHUÀI}}\})$	function application (part of CONJ-rule)
2	$\{\langle\sqrt{\text{GĀO}}, \sqrt{\text{FÙ}}, \sqrt{\text{SHUÀI}}\rangle, \ldots\}$	sequence set formation (part of CONJ-rule)
3	$\langle\sqrt{\text{GĀO}}, \sqrt{\text{FÙ}}, \sqrt{\text{SHUÀI}}\rangle$	ε-rule (part of CONJ-rule)
3	$\{\text{CONJ}, \langle\sqrt{\text{GĀO}}, \sqrt{\text{FÙ}}, \sqrt{\text{SHUÀI}}\rangle\}$	lexical subarray updating (part of CONJ-rule)
4	$\langle\sqrt{\text{G}}, \sqrt{\text{F}}, \sqrt{\text{S}}\rangle$ CONJ	set merge (label-less) *Neither CONJ nor $\sqrt{}$ is a labeling category.*
4	n n $\langle\sqrt{\text{G}}, \sqrt{\text{F}}, \sqrt{\text{S}}\rangle$ CONJ	set merge (categorization) *Semantically, not only the entire word but also its components get nominalized, or in Chomsky's notation: $\langle \mathit{CONJ}, \langle\sqrt{\text{G}}, n\rangle, \langle\sqrt{\text{F}}, n\rangle, \langle\sqrt{\text{S}}, n\rangle\rangle$.*

Plan

Summary

Overall, we can reach the following conclusions:

- Unlike in mathematics, the ε-operator is semideterministic in both von Heusinger's and Chomsky's linguistic application. Our case study further shows that the ε-choice is made based on multiple factors.
- Chomsky's tuple-notation is a compact, high-level representation that doesn't reflect the actual derivation (or composition).
- Chomsky's ε-theory doesn't absolutely hinge on pair merge. It's possible to apply ε in a set-merge-only setting.
- There's much going on in the lexical-array-forming stage, which deserves more attention than it currently enjoys in minimalism.

Summary

Chomsky (2000 et seq.) resorts to pair merge to produce the inherent asymmetry in adjunction, which is essentially an additional axiom.

My goal today

Explore how the ε-based new approach changes the picture.

Both ε and pair merge serve to generate asymmetry, but the former type of asymmetry is clearly not that inherent in adjunction:

- ε: asymmetry between multiple coordinated items
- adjunction: (essentially) asymmetry between two merging sisters

So, insofar as adjunction is concerned, the addition of ε to the axiom set of minimalism doesn't change the picture in a fundamental way. But ε clearly is a very useful conceptual tool in its own right.

Thank you!

Selected references I

Chomsky, Noam
The minimalist program
MIT Press, 1995

Chomsky, Noam
Minimalist inquiries: The framework
Step by step: Essays on minimalist syntax, 89–156
MIT Press, 2000

Chomsky, Noam
UCLA lectures
https://linguistics.ucla.edu/noam-chomsky/
Apr 29–May 2, 2019

Selected references II

Hilbert, David & Paul Bernays
Grundlagen der Mathematik, Vol. 2
Springer, 1939

von Heusinger, Klaus
Definite descriptions and choice functions
Logic, language and computation, 61–91
Kluwer, 1997

von Heusinger, Klaus
The salience theory of definiteness
Perspectives on linguistic pragmatics, 349–374
Springer, 2013

Selected references III

Leisenring, Albert
Mathematical logic and Hilbert's epsilon-symbol
Macdonald, 1969

Song, Chenchen
On the formal flexibility of syntactic categories
University of Cambridge dissertation, 2019