WILEY | Hindawi

Research Article

Calculating Zero Pronominals in Situ: A Type Logical Approach

Haixia Man ⓘ,[1] **Chongli Zou** ⓘ,[2] **and Kesheng Li**[3]

[1]*School of Foreign Studies, University of Science and Technology Beijing, Beijing 100083, China*
[2]*Biquan Academy, Xiangtan University, Xiangtan 411105, China*
[3]*School of Foreign Languages, Hefei Normal University, Hefei 230601, China*

Correspondence should be addressed to Chongli Zou; chlizou@263.net

Zero pronominals challenge Type Logical Grammar in two ways. One, TLG displays a linear resource management regime for semantic composition, meaning that pronominals call for special treatment if they want to do resource multiplication. Two, as a grammar of lexicalism, TLG applies to phonologically realized lexical entries only, illegitimating the phonetically null items during syntactic derivation. Jägor extends the inventory of category-forming connectives of TLG by a third kind of implication that creates categories of anaphoric items and solves the first problem above. This article goes a step further to tackle the second one. In order to formalize the constructions with zero pronominals, we design a ternary category $[A\langle B\rangle C]$ and include the latter into Jägor's system. The proposed system is proof-theoretically well-behaved. It is complete, sound, and decidable. More importantly, zero pronominals of various forms can be derived in the system.

1. Introduction

Natural languages are economical, arguably on all tangible levels, to convey more information with less effort [1, 2]. On the sentence level, various resource reuse strategies are utilized to shun repetitiveness. The multiplicated resources sometimes resort to pronouns, reflexives, and auxiliaries, as in (1)–(3), and sometimes are covert, as the subject in the second coordinate of (4) and the PROs in control constructions of (5) and (6). Interestingly, however, the story goes to the contrary on the semantic side. Each of the repetitive resources, even if it is in zero form, multiplies the meaning of its antecedent and reappears in semantics. For instance, the two PROs in (5) and (6) pick up the meaning representations of their controllers in logical interpretations. They are logically represented as the subject JOE′ and object JEAN′ (in **bold**), respectively, of the upstairs verbs, as shown in (7) and (8).

(1) Joe claims that *he* will win.   pronoun

(2) Joe likes *himself.*   reflexive

(3) Joe walks and Jean *does* too.   ellipsis

(4) Joe *walks and talks.*   coordination

(5) Joe$_i$ promises Jean [PRO$_i$ to stay]   subject control

(6) Joe asks Jean$_i$ [PRO$_i$ to stay].   object control

(7) PROMISE′ (JOE′, JEAN′, STAY′(**JOE′**))

(8) ASK′ (JOE′, JEAN′, STAY′(**JEAN′**))

The issue of how to calculate these resource reuse mechanisms in TLG is an apparent challenge to the grammar because TLG assumes a monostratal model for natural languages, which means that they can only combine neighboring items or constructions. However, all above-mentioned reused linguistic items in (1)–(6), including anaphora, non-constituent coordination, and control, are discontinuous constructions, in which the anaphor/PRO locates several words away from its antecedent/controller. Thus, the grammar has to be capable of coping with discontinuity in order to be a competent natural language grammar. There are mainly two ways to deal with anaphoric constructions in type logical setting. One is to treat pronominals as being triggered by certain lexical items whose semantic representations contain a λ-operator that binds more than one variable occurrence

[3–7]. However, the strategy of this kind forces highly complex lexical entries and coerces formidable mechanisms, like secondary wrapping [5], into syntax. Another way is to introduce into syntax a specifically designed operation to secure the pronominals a semantic interpretation as simple as an identity function. Hepple's permutation operator Δ [8], Jacobson's variable free semantics [9, 10], and Jägor's Lambek calculus with Limited Contraction (LLC) [11] are all attempts of the sort, among which the last attracts the most attention and engenders a series of logical extensions for its theoretical simplicity [12–15].

Jägor [11] includes a limited version of the structural rule of contraction as is shown in Rule 1. Rule 1 does nothing more than allowing the antecedent formulae to be multiplied. The resulted system LLC highlights its vertical slash | that creates categories of anaphoric items. A sign has category $A|B$ iff it needs an antecedent of category B and, if it finds one, behaves like an item of category A. Thus, both the pronoun in (1) and reflexive in (2) are identity functions $\lambda x.x$ of category $np|np$. Making use of | Elimination Rule 2, the simple reflexive sentence (2) is derived as in Figure 1. We use natural deductions for rules or syntactic derivation hereinafter due to the reason stated in [11] as there is a tight correspondence between the structure of proofs and the syntactic structure of the Curry–Howard terms.

Rule 1. Sequent representation of limited contraction 1

$$\frac{X, x\colon A, y\colon A, Y \Rightarrow M\colon B}{X, x\colon A, Y \Rightarrow M[x/y]\colon B}. \tag{1}$$

Rule 2. | ELIMINATION

$$[M\colon A]_i \quad \cdots \quad \frac{N\colon B|A}{\mathrm{NM}\colon B} |E, i. \tag{2}$$

Despite all these successful treatments over discontinuity, we find that the current categorial machine struggles to legitimately derive constructions concerning zero pronominals, for example, the PROs in (5) and (6). PROs are different from elided constituents in coordination. The latter can be derived by the generalized coordination rule (Rule 3) as shown in Figure 2, whereas the PROs are in embedded clauses rather than a coordination construction, thus not eligible for predicate coordination.

Rule 3. GENERALIZED BOOLEAN COORDINATION SCHEME

$$\frac{\vdots /M\colon np\backslash s \text{ and } \vdots /N\colon np\backslash s}{\lambda x[Mx \wedge Nx]\colon np\backslash s} \text{Conj.} \tag{3}$$

Here is the dilemma for deriving pronominals in TLG. On the one hand, the zero pronominals have to be explicit to carry their meaning for appropriate semantic derivation, while on the other hand, Lambek calculus does not allow zero pronominals to be made explicit as it will induce the structural rule of monotonicity and jeopardizes decidability

FIGURE 1: Type Logical derivation of (2).

FIGURE 2: Type Logical derivation of (4).

of the system by allowing the addition of some formula that is similar to one in the antecedent.

As a result, an ideal system for the present purpose is to extend LLC further to allow covert anaphoric items to be overt without hurting the system's decidability. This is our goal in the present study. It differs from earlier extensions of LLC in such a way that it is a direct logical expansion, rather than a lexical enrichment that includes anaphoric slash in the categories of control verbs [12], and may offer a more meticulous version for the noticing trending works of connecting TLG to computational distributional semantics [16–20]. In Section 2, we will flash out our theoretical assumptions for zero pronominals, then present the axiomatic presentation, Gentzen style sequent formulation, and labeled natural deduction of the new system LLCM in Sections 3–5, respectively, proving that LLCM is complete, sound, and decidable. More linguistic phenomena concerning zero pronominals will also be discussed in Section 5.

2. Anaphora Slot and LLCM

We want to follow Jäger's approach to anaphors [11] and extend LLC in a way that it can accommodate two desirable properties: one, the system should allow covert item to be overt; two, the made-explicit item should be anaphoric. This means that the system should allow both pronominals and zero pronominals. In this way, the PROs in control constructions can take up an actual logical form as well as interpretation during sentence derivation. Thus, (5) and (6) can be derived ideally as in Figures 3 and 4, where *to stay* can take up the right subject identical to its controller, instead of looting direct object of the upstairs predicate—its left neighbor—to be its subject. (To simplify the derivation, we ignore the morphological distinction between finite and infinite VPs and treat *to stay* as a single lexical entry.)

We name the location where the zero pronominal resides "anaphora slot." Thus, an anaphora slot introduction rule in embryo should be like Rule 4.

Rule 4. ANAPHORA SLOT INTRODUCTION

$$\frac{M\colon A \quad N\colon C}{M\colon A \quad W\colon B|A \quad N\colon C}. \tag{4}$$

$$\frac{Joe}{[JOE':np]_i} \quad promises \quad \frac{Jean}{JEAN':np} \quad \frac{to\ stay}{\lambda x.STAY'(x):np\backslash s}$$
$$\frac{\dfrac{\lambda x.x:np|np}{JOE':np}|E,i}{STAY'(JOE'):s}$$

Figure 3: Partial derivation for (5).

$$\frac{Joe}{JOE':np} \quad asks \quad \frac{Jean}{[JEAN':np]_i} \quad \frac{to\ stay}{\lambda x.STAY'(x):np\backslash s}$$
$$\frac{\dfrac{\lambda x.x:np|np}{JEAN':np}|E,i}{STAY'(JEAN'):s}$$

Figure 4: Partial derivation for (6).

Under type logical perspective, introducing an anaphora slot in syntax means that the logic of grammatical composition allows somewhat "redundant" (because they are covert in actual discourse) but not arbitrary category in a valid deduction. This amounts to an assumption that the structural rule of monotonicity is part of the grammar in one way or another. Thus, we name our system LLCM, meaning LLC with limited monotonicity.

We put forward a ternary category $[A\langle B\rangle C]$, in which $[\ldots\langle\ldots\rangle\ldots]$ is an anaphora slot operator (sometimes simplified as $\langle\ \rangle$). It may help to reveal how the category of a zero pronominal is introduced, deleted, or concatenated with categories to its left or right. From our experience with control construction and PROs, there is a preferable concatenative order among the adjacent strings "A, B, C," where B will concatenate with C (if there is a C to B's right) first and the result will further concatenate leftwards with its left neighbor A. Thus, the zero pronominals, when made overt, should obey the limited contraction 2 (Rule 5) in its sequent representation.

Rule 5. SEQUENT REPRENTATION OF LIMITED CONTRACTION

$$\frac{X, (x: A, (y: A, z: C)), Y \Rightarrow M: D}{X, (X: A, z: C), Y \Rightarrow M[x/y]: D}. \tag{5}$$

In the coming two sections, we will define LLCM of its axiomatic version and Gentzen style sequent formulation, and prove that it is sound, complete, and decidable under such expansion.

3. Model Theory of LLCM

Now we extend the inventory of LLC category-forming connectives by the ternary operator $[\langle\rangle]$. So the set of LLCM categories $\mathscr{F}$ over a collection of atomic categories A is given below.

Definition 1. LLCM CATEGORIES

If F is a well-formed LLCM categories, then F/F, $F\backslash F$, $F|F$, $F\cdot F$ and $[F{<}F{>}F]$ are also well-formed LLCM categories.

All well-formed LLCM categories are recursively defined as in Definition 1. Next, a sound and complete model-theoretic interpretation for LLCM is presented.

Definition 2. MODEL of LLCM (This model is based on LLC in [11] and a preliminary version of this system is given in our earlier work [21].)

A Model $\mathscr{M}$ for LLCM is a tuple $<W, R, S, T, \sim, f, g>$, where W is a non-empty set of linguistic signs, $T \subseteq W^4$ is a quaternary relation on W; $R, S \subseteq W^3$ are ternary relations on W; $\sim\ \subseteq W^2$ is a binary relation on W; f is a function from atomic categories to subsets of W; and g is a function from LLCM-categories to W. The verification relation between points in W and LLCM categories is defined as follows:

$$\|p\|\ \mathscr{M} = f(p) \subseteq W$$
$$\|A\bullet B\|\ \mathscr{M} = \{x | \exists yz[Rxyz\ \&\ y \in \|A\|\ \mathscr{M}\ \&\ z \in \|B\|\ \mathscr{M}]\}$$
$$\|A\backslash B\|\ \mathscr{M} = \{x | \forall yz[Rzyx\ \&\ y \in \|A\|\ \mathscr{M} \Rightarrow z \in \|B\|\ \mathscr{M}]\}$$
$$\|A/B\|\ \mathscr{M} = \{x | \forall yz[Rzxy\ \&\ y \in \|B\|\ \mathscr{M} \Rightarrow z \in \|A\|\ \mathscr{M}]\}$$
$$\|A|B\|\ \mathscr{M} = \{x | \exists y\ [Sxyg(B)\ \&\ y \in \|A\|\ \mathscr{M}]\}$$
$$\|A\langle B\rangle C\|\ \mathscr{M} = \{x | \exists yzu[Txyzu\ \&\ y \in \|A\|\ \mathscr{M}\ \&\ z \in \|B\|\ \mathscr{M}\ \&\ u \in \|C\|\ \mathscr{M}]\}$$

The ternary relation R can be taken as ordinary syntactic concatenation between linguistic signs. $Rxyz$ means that if y and z occur adjacently in that order, the combination of the two gets an x. Relation S is a ternary relation in [11]. It is similar to R but it is responsible for anaphoric resolution. $Sxyz$ means that x is changed into y if there is an element similar to z (noted as $y \sim z$ in meaning postulates) available to be antecedent for anaphora resolution. Relation T models anaphora slot operation and $Txyzu$ means that x is the result of inserting z in between y and u.

The following meaning postulates hold:

MP1 $\forall xyzwu[Rxyz\ \&\ Szwu \Rightarrow \exists\ v\ [Sxvu\ \&\ Rvyw]]$

MP2 $\forall xyzwu[Rxyz\ \&\ Sywu \Rightarrow \exists\ v\ [Sxvu\ \&\ Rvwz]]$

MP3 $\forall xyzwuv[Rxyz\ \&\ Sywu\ \&\ Szvu \Rightarrow \exists r[Sxru\ \&\ Rrwv]]$

MP4 $\forall xyzwu[Rxyz\&Szwu\&y\sim u \Rightarrow Rxyw]$

MP5 $\forall Aw[w \in \|A\|\mathscr{M} \Rightarrow w \sim g(A)]$

MP6 $\forall xyzu[Rxyz \Rightarrow Txyuz]$

MP7 $\forall xyzu[Txyzu \Rightarrow \exists\ v\ [Rxyv\ \&\ Rvzu]]$

MP8 $\forall xyzuvw[Rxyz\ \&\ Tzuvw \Rightarrow \exists t[Txtvw\ \&\ Rtyu]]$

MP9 $\forall xyzuvw[Rxyz\ \&\ Tyuvw \Rightarrow \exists t[Txtvw\ \&\ Rtuz]]$

MP10 $\forall xyzuvw[Rxyz\ \&\ Tzuvw \Rightarrow \exists t[Txuvt\ \&\ Rtyw]]$

MP11 $\forall xyzuvw[Rxyz\ \&\ Tyuvw \Rightarrow \exists t[Txuvt\ \&\ Rtwz]]$

MP12 $\forall xyzuvwst[Rxyz\ \&\ Tyuvw\ \&\ Tzsvt \Rightarrow \exists ab[Txavb\ \&\ Raus\ \&\ Rbwt]]$

MP13 $\forall xy[Rxyy \Rightarrow y \subseteq x]$

MP14 $\forall B[\|B\|\mathcal{M} \neq \varnothing]$

MP1–5 mean postulates about relation S [11]. MP 6–7 are about relation T, exhibiting an important feature of LLCM. They amount to say that if x contains an anaphora slot and x is composed by y and z, then x can also be composed by y and v among which v is the result of conjoining z and the covert u. In other words, a complex sign with an anaphora slot can be represented similarly either with or without the anaphor participating in its syntactic composition. MP8–12 show how categories in the ternary operator $[\langle\rangle]$ compose with its neighboring categories. MP13 says that the extension over a set of linguistic signs is closed under R. The last postulate is a structural postulate complementary to MP6 and MP7.

Definition 3. AXIOMATIC VERSION OF LLCM

The axiomatic version of LLCM is the system that is obtained when the following 12 axioms and 4 rules are added to the axiomatic version of L:

A1 $A \bullet B|C \longrightarrow (A \bullet B)|C$

A2 $A|B \bullet C \longrightarrow (A \bullet C)|B$

A3 $A|C \bullet B|C \longrightarrow (A \bullet B)|C$

A4 $A \bullet B|A \longrightarrow A \bullet B$

A5 $A \bullet C \longrightarrow [A\langle B\rangle C]$

A6 $[A\langle B\rangle C] \longrightarrow A \bullet (B \bullet C)$

A7 $D \bullet [A\langle B\rangle C] \longrightarrow [(D \bullet A)\langle B\rangle C]$

A8 $[A\langle B\rangle C] \bullet D \longrightarrow [(A \bullet D)\langle B\rangle C]$

A9 $D \bullet [A\langle B\rangle C] \longrightarrow [A\langle B\rangle (D \bullet C)]$

A10 $[A\langle B\rangle C] \bullet D \longrightarrow [(A\langle B\rangle (C \bullet D)]$

A11 $[A\langle B\rangle C] \bullet [D\langle B\rangle E] \longrightarrow [(A \bullet D)\langle B\rangle (C \bullet E)]$

A12 $A \bullet A \longrightarrow A$ (monotonicity)

Deductive rules:

D1

$$\frac{A \longrightarrow B}{A|C \longrightarrow B|C}. \tag{6}$$

D2

$$\frac{A \longrightarrow B}{[A\langle D\rangle C] \longrightarrow [B\langle D\rangle C]}. \tag{7}$$

D3

$$\frac{A \longrightarrow B}{[D\langle C\rangle A] \longrightarrow [D\langle C\rangle B]}. \tag{8}$$

D4

$$\frac{A \longrightarrow B}{[D\langle A\rangle C] \longrightarrow [D\langle B\rangle C]}. \tag{9}$$

Then we can prove the soundness and completeness of the axiomatic version in a way that closely follows the proof for L in [22] and LLC in [11]. We will start with the axiom 6 and D2. The rest cases are already proved in [11].

Theorem 1. *SOUNDNESS*

If $LLCM \vdash A \longrightarrow B$, *then for all models M,* $\|A\| \mathcal{M} \subseteq \|B\| \mathcal{M}$.

Proof. $A \bullet C \longrightarrow [A\langle B\rangle C]$

Suppose $x \in \|A \bullet C\| \mathcal{M}$. Then there are $y \in \|A\| \mathcal{M}$ and $z \in \|C\| \mathcal{M}$ such that $Rxyz$. According to MP14 and MP6, there is a $u \in \|B\| \mathcal{M}$ such that $Txyuz$. Hence, for $x \in \|A \bullet C\| \mathcal{M}$, there is $y \in \|A\| \mathcal{M}$, $u \in \|B\| \mathcal{M}$, and $z \in \|C\| \mathcal{M}$ such that $Txyuz$, thus $x \in \|[A\langle B\rangle C]\| \mathcal{M}$.

Now prove $[A\langle B\rangle C] \longrightarrow A \bullet (B \bullet C)$.

Suppose $x \in \|[A\langle B\rangle C]\| \mathcal{M}$. Then there are $y \in \|A\| \mathcal{M}$, $z \in \|B\| \mathcal{M}$, and $u \in \|C\| \mathcal{M}$ such that $Txyzu$. According to postulate 7, there is a v such that $Rxyv$ and $Rvzu$. Hence $v \in \|B \bullet C\| \mathcal{M}$, and thus $x \in \|A \bullet (B \bullet C)\| \mathcal{M}$.

Axioms 7–11 stipulate associativity of the ternary category. Here we will prove axiom 7 only.

Now prove $D \bullet [A\langle B\rangle C] \longrightarrow [(D \bullet A)\langle B\rangle C]$.

Suppose $x \in \|D \bullet [A\langle B\rangle C]\| \mathcal{M}$. Then there are $y \in \|D\| \mathcal{M}$ and $z \in \|[A\langle B\rangle C]\| \mathcal{M}$ such that $Rxyz$. Furthermore, there is a $y \in \|D\| \mathcal{M}$ such that $Rxyz$, and $u \in \|A\| \mathcal{M}$, $v \in \|B\| \mathcal{M}$, and $w \in \|C\| \mathcal{M}$ such that $Tzuvw$. MP8 entails that there is a t such that $Txtvw$ and $Rtyu$. Hence $t \in \|D \bullet A\| \mathcal{M}$, and then $x \in \|[(D \bullet A)\langle B\rangle C]\| \mathcal{M}$.

Last, we prove $A \bullet A \longrightarrow A$.

Suppose $x \in \|A \bullet A\| \mathcal{M}$. Then there is a $y \in \|A\| \mathcal{M}$ such that $Rxyy$. MP13 entails that $y \subseteq x$. Since $\subseteq$ is closed under W, thus $x \in \|A\| \mathcal{M}$.

The proof for deduction rules are similar. We will leave them for readers as exercise.   □

Theorem 2. *COMPLETENESS*

For all LLCM-models M, if $\|A\| \mathcal{M} \subseteq \|B\| \mathcal{M}$, *then* $LLCM \vdash A \longrightarrow B$. *(We may use* $\vdash$ *to stand for* $LLCM \vdash$ *when no misunderstanding arises.)*

Proof. We start with constructing a canonical model $CM = \langle W, R, S, T, \sim, f, g \rangle$, where W is simply a set of all LLCM categories.

For all atomic categories p, $f(p) = \{A \mid \vdash A \longrightarrow p\}$.

$RABC$ iff $\vdash A \longrightarrow B \bullet C$, and

$SABC$ iff $\vdash A \longrightarrow B | C$, and

$TABCD$ iff $\vdash A \longrightarrow [B\langle C\rangle D]$,

and $A \sim B$ iff $\vdash A \longrightarrow B$,

and $g(A) = A$ for all categories A and B.

$A \subseteq B$ iff $\vdash B \longrightarrow A$

Then we prove the truth lemma below.   □

3.1. Truth Lemma. In canonical model CM, it holds for all LLCM categories A and B that $A \in \|B\|_{C}\mathcal{M}$ iff $\vdash A \longrightarrow B$.

Proof. We prove this via induction over the complexity of *B*. If *B* is atomic, it follows directly from the construction of *f*. If *B* is constructed by one of the three Lambek connectives or LLC's anaphoric connective, the proof of the induction step follows that in [11, 21]. We will only show the induction steps for the ternary operator.

⇒Suppose that $B = [C \langle E \rangle D]$ and suppose $A \in \| [C \langle E \rangle D] \|_{\mathrm{C}}\mathscr{M}$. Then there are $y \in \| C \|_{\mathrm{C}}\mathscr{M}$, $z \in \| E \|_{\mathrm{C}}\mathscr{M}$, and $u \in \| D \|_{\mathrm{C}}\mathscr{M}$ such that $TAyzu$ according to the definition of []. This means that there are $A_1 \in \| C \|_{\mathrm{C}}\mathscr{M}$, $A_2 \in \| E \|_{\mathrm{C}}\mathscr{M}$, and $A_3 \in \| D \|_{\mathrm{C}}\mathscr{M}$ such that $TAA_1A_2A_3$. It follows that $\vdash A_1 \longrightarrow C$, $\vdash A_2 \longrightarrow E$, $\vdash A_3 \longrightarrow D$, and $\vdash A \longrightarrow [A_1 \langle A_2 \rangle A_3]$ by the way the canonical model is defined. By D2, D3, and D4, it follows $[A_1 \langle A_2 \rangle A_3] \longrightarrow [C \langle A_2 \rangle A_3]$, $[C \langle A_2 \rangle A_3] \longrightarrow [C \langle E \rangle A_3]$, and $[C \langle E \rangle A_3] \longrightarrow [C \langle E \rangle D]$, respectively. Hence $\vdash A \longrightarrow [C \langle E \rangle D]$.

⇐ Suppose that $\vdash A \longrightarrow [C \langle E \rangle D]$. By the way the model is constructed, there is $TACED$. By the induction hypothesis, there are $C \in \| C \|_{\mathrm{C}}\mathscr{M}$, $E \in \| E \|_{\mathrm{C}}\mathscr{M}$, and $D \in \| D \|_{\mathrm{C}}\mathscr{M}$. Hence, there is an A such that $TACED$. By the way [] is defined, thus $A \in \| [C \langle E \rangle D] \|_{\mathrm{C}}\mathscr{M}$.

Then we have to prove that all postulates for LLCM in Definition 2 are fulfilled by the model. They follow directly from the model construction, monotonicity of the product, and the truth lemma above, thus will not be provided here. ∎

Finally, we will show that all LLCM valid formulae are derivable in LLCM. Let $\| A \|_{\mathrm{C}}\mathscr{M} \subseteq \| B \|_{\mathrm{C}}\mathscr{M}$. Suppose that A ⟶ B is not derivable in LLCM. There should be a $A \notin \| B \|_{\mathrm{C}}\mathscr{M}$ by the truth lemma. By identity axiom, $\vdash A \longrightarrow A$. Thus there is always $A \in \| A \|_{\mathrm{C}}\mathscr{M}$ and it is not the case $\| A \|_{\mathrm{C}}\mathscr{M} \subseteq \| B \|_{\mathrm{C}}\mathscr{M}$, which contradicts our assumption. Hence, A ⟶ B is derivable in LLCM.

Henceforth, the axiomatic system of LLCM is sound and complete. □

4. Sequent Presentation of LLCM

In order to characterize the decidability of LLCM, we need its Gentzen-style sequent presentation. The sequent presentation of LLCM extends that of LLC by proposing R and L for anaphora slot operator [], and monotonicity. For simplicity, we will omit the labeled λ-terms for the sequent presentation in the present section.

Definition 4. SEQUENT PRESENTATION OF LLCM

$$A \Rightarrow A \, [Ax] \qquad \frac{\Delta \Rightarrow A \quad \Gamma, \ A, \ \Gamma\prime \Rightarrow C}{\Gamma, \Delta, \Gamma\prime \Rightarrow C} \, [Cut],$$

$$\frac{\Delta, \ B \Rightarrow A}{\Delta \Rightarrow A/B} \, [/R] \qquad \frac{\Delta \Rightarrow B \quad \Gamma, \ A, \ \Gamma\prime \Rightarrow C}{\Gamma, A/B, \Delta, \Gamma\prime \Rightarrow C} \, [/L],$$

$$\frac{B, \ \Delta \Rightarrow A}{\Delta \Rightarrow B/A} \, [\backslash R] \qquad \frac{\Delta \Rightarrow B \quad \Gamma, \ A, \ \Gamma\prime \Rightarrow C}{\Gamma, \Delta, B/A, \Gamma\prime \Rightarrow C} \, [\backslash L],$$

$$\frac{\Gamma, A, B, \ \Gamma\prime \Rightarrow C}{\Gamma, A \bullet B, \ \Gamma\prime \Rightarrow C} \, [\bullet L] \qquad \frac{\Gamma \Rightarrow A \quad \Gamma\prime \Rightarrow B}{\Gamma, \Gamma\prime \Rightarrow A \bullet B} \, [\bullet R],$$

$$\frac{X, A_1, Y_1, \ldots, A_n, Y_n \Rightarrow B}{X, A_1 | C, Y_1, \ldots, A_n | C, Y_n \Rightarrow B | C} | R \qquad \frac{Y \Rightarrow B \quad X, B, Z, A, W \Rightarrow C}{X, Y, Z, A | B, W \Rightarrow C} | L, \tag{10}$$

$$\frac{\Gamma, A_1, Y_1, \ldots, A_n, Y_n \Rightarrow A \quad B \Rightarrow D \quad \Delta, C_1, Z_1, \ldots, C_n, Z_n \Rightarrow C}{\Gamma, \Delta, [A_1 \langle B \rangle C_1], Y_1, Z_1, \ldots, [A_n \langle B \rangle C_n], Y_n, Z_n \Rightarrow [A \langle D \rangle C]} \langle \rangle R,$$

$$\frac{X, A, B, C, Y \Rightarrow D}{X, [A \langle B \rangle C], Y \Rightarrow D} \langle \rangle L,$$

$$\frac{X, (A, (B, C)), Y \Rightarrow D}{X, (A, C), Y \Rightarrow D} C \, (\text{Limited Contraction 2}),$$

$$\frac{X, A, Y \Rightarrow D}{X, A, A, Y \Rightarrow D} \, (\text{Monotonicity}).$$

To prove the sequent presentation is equivalent to the axiomatic version, Lemma 1 is needed. And also, a function σ that maps all commas in a sequent into products $\bullet$ is utilized to ensure categories to type correspondence. Thus, $\sigma(A) = A$; $\sigma(X, Y) = \sigma(X) \bullet \sigma(Y)$.

Lemma 1. *The arrow $\sigma(\Gamma, \Delta, [A_1\langle B\rangle C_1], Y_1, Z_1, \ldots, [A_n\langle B\rangle C_n], Y_n, Z_n)\longrightarrow[\sigma(\Gamma, A_1, Y_1, \ldots, A_n, Y_n)\langle B\rangle\sigma(\Delta, C_1, Z_1, \ldots, C_n, Z_n)]$ is derivable in LLCM's axiomatic version.*

Proof. For $n = 1$, the arrow $\sigma(\Gamma, \Delta, [A_1\langle B\rangle C_1], Y_1, Z_1) \longrightarrow [\sigma(\Gamma, A_1, Y_1)\langle B\rangle\sigma(\Delta, C_1, Z_1)]$ holds by axioms 7–10 and the way "σ" is defined.

Provided that the arrow holds when $n = k$. By induction hypothesis, arrow $\sigma(\Gamma, \Delta, [A_1\langle B\rangle C_1], Y_1, Z_1, \ldots, [A_k\langle B\rangle C_k], Y_k, Z_k) \longrightarrow [\sigma(\Gamma, A_1, Y_1, \ldots, A_k, Y_k)\langle B\rangle\sigma(\Delta, C_1, Z_1, \ldots, C_k, Z_k)]$ holds in axiomatic version. Obviously, there is $\sigma([A_{k+1}\langle B\rangle C_{k+1}], Y_{k+1}, Z_{k+1}) \longrightarrow [\sigma(A_{k+1}, Y_{k+1}) \langle B\rangle\sigma(C_{k+1}, Z_{k+1})]$ by A8 and A10. Via monotonicity of "$\bullet$", there is: $\sigma(\Gamma, \Delta, [A_1\langle B\rangle C_1], Y_1, Z_1, \ldots, [A_k\langle B\rangle C_k], Y_k, Z_k) \bullet \sigma([A_{k+1}\langle B\rangle C_{k+1}], Y_{k+1}, Z_{k+1}) \longrightarrow [\sigma(\Gamma, A_1, Y_1, \ldots, A_k, Y_k)\langle B\rangle\sigma(\Delta, C_1, Z_1, \ldots, C_k, Z_k)] \bullet [\sigma(A_{k+1}, Y_{k+1})\langle B\rangle\sigma(C_{k+1}, Z_{k+1})]$

By σ's definition and A11, we get $\sigma(\Gamma, \Delta, [A_1\langle B\rangle C_1], Y_1, Z_1, \ldots, [A_k\langle B\rangle C_k], Y_k, Z_k, [A_{k+1}\langle B\rangle C_{k+1}], Y_{k+1}, Z_{k+1}) \longrightarrow [\sigma(\Gamma, A_1, Y_1, \ldots, A_k, Y_k, A_{k+1}, Y_{k+1})\langle B\rangle\sigma(\Delta, C_1, Z_1, \ldots, C_k, Z_k, C_{k+1}, Z_{k+1})]$ ☐

Lemma 1. *is thus proved.*

Theorem 3. *EQUIVALENCE OF AXIOMATIC AND GENTZEN PRESENTATIONS*
$LLCM \vdash X \Rightarrow A$ *iff* $\vdash \sigma(X) \longrightarrow A$ *is derivable in the axiomatic version.*

The proof is omitted for the current purpose. We prove the equivalence of LLCM's axiomatic and sequent versions for the same reason that Lambek proves the equivalence of L's sequent version and its axiomatic counterpart. The decidability is decidable by *Cut Elimination* in sequent presentation and this result can therein further percolate to its axiomatic version.

Theorem 4. *CUT ELIMINATION*
If $LLCM \vdash X \Rightarrow A$, then there is a Cut-free sequent proof of $LLCM \vdash X \Rightarrow A$.

To prove this theorem, we have to distinguish three cases: (1) at least one premise of the Cut is an identity axiom; (2) both premises are results of logical rules, and the Cut formula is the active formula in both premises; (3) both premises result from introducing logical rules, and the Cut formula is not the active formula in one premise. The proof is left as an exercise to the reader.

Theorem 5. *DECIDABILITY*
Decidability in LLCM is decidable.

Proof. For each rule of the Cut-free sequent calculus, the conclusion sequent of each rule contains more symbols than its premises because each formula in the premise occurs as a subformula in the conclusion and each logical rule introduces one connective. In addition, there are only finite ways to match certain sequents with the conclusion of some sequent rule. As a result, there are always at most finite choices to do a bottom-up proof search and every branch of the proof tree is finite. Decidability in LLCM is thus decidable. ☐

5. Tests on More Linguistic Phenomena

5.1. LLCM's Natural Deduction in Tree-format. Before testing on more linguistic phenomena, we offer LLCM's labeled natural deduction in tree-format here. As is shown in Sections 1 and 2, labeled deduction in tree-format helps to visualize the type-logical deduction over a sentence. Suppose Δ is a n-ary operator, ΔE and ΔI stand for its elimination rule and introduction rule, respectively. Here, we will offer <>I only.

Definition 5. <>INTRODUCTION RULE IN TREE-FORMAT

$$D: A_2 \frac{A: M \quad C: N}{\langle B: A_1|A_2\rangle}\langle\,\rangle I. \qquad (11)$$

$\langle\,\rangle$ I enables the covert or elided pronominals to be inserted first before a $[A\langle B\rangle C]$ structure is constructed. E, however, is not given because is a temporary notational strategy to show where zero pronominal is. The notational will disappear when the zero pronoun finds its antecedent and multiplies its interpretation with $|E$. Now we can show the charm of LLCM with more linguistic constructions that allow zero pronominals.

5.2. Deriving Pros. The anaphora and control constructions we list in (1)–(6) exhibit only the tip of the iceberg of the zero pronominals used in natural languages. Generative grammar distinguishes two types of zero pronominals. In addition to the PROs that are limited to the subject position of a non-finite clause as in control constructions of (5) and (6), there are zero pronouns that occur elsewhere in less restricted manners than PROs. The linguistic economy allows pronouns to be dropped in many languages. They are called little *pro*. For instance, a subject pronoun in Spanish may be dropped from a tensed clause as in (9), and in Chinese, both the subject and object pronouns may be dropped in similar circumstances as in (10) and (11). Apparently, a type-logical system ready to derive correct readings for sentences in these languages is expected to be capable of inserting the elided pronouns in the anaphora slots and multiplying the semantic resource legitimately.

  (9) José sabe [que él/*pro* ha sido visto por María]. [22].
     José know that he/∅ has been see by María.
     'José knows that [he] has been seen by Maria.'

(10) Zhangsan$_i$ shuo [Lisi hen xihuan *pro* $*_{i/j}$]. [23].
     Joe said Lee very liked [him].
     'Joe$_i$ said that Lisi liked [him $*_{i/j}$].'

(11) Zhangsan$_i$ shuo [ta/*pro* hen xihuan *pro* $*_{i/j}$].

Figure 5: LLCM derivation of (9).

Figure 6: LLCM derivation of (10).

Joe said he/∅ very like ∅

'Joe said that [he] liked [him/it...].'

Thus, when the Lambek system is equipped with anaphoric slash and an anaphora slot operator, it becomes as powerful as we would expect it to be. For example, derivation of (9) is illustrated in Figure 5.

In Chinese, both subject and object pronominals in a tensed clause can be dropped. However, the dropped object cannot take the matrix subject as its antecedent, but as some other person known in the discourse. For example, in the Chinese discourse (12) below, *pro* in speaker B's answer can only refer to the object in A's question. Ideally, if our system allows the inserted pronominal to search its antecedent across the sentence border, it can derive (10) in the same way as that of (12) in Figure 6. Derivation for sentences like (11) that drops both subject and object is likewise.

(12) Speaker A: Zhangsan xihuan huaju ma?

Joe like stage-play Q?

'Does Joe like stage play?'

Speaker B: Zhangsan shuo ta hen xihuan *pro*.

Joe said he very like [it].

6. Discussion

LLCM's labeled natural deduction in tree-format shows that in LLCM, sentences with different degrees of zero pronominals are all derivable, be it a PRO with the restricted occurrence, or a pro-drop with less restrictions in syntax. It exhibits a very promising picture. Nevertheless, poetic as it seems to be, the situation of pros is more complicated than we assume because they are under different restrictions in different pro-drop languages. How to tailor the system according to the requirements of different languages remains to be a question. It might be a good idea to set up a universal model and assign different parameters for different languages as in multimodal CCG [24]. We will leave it for future work. [25].

As seen from substructural logics, system LLCM with the ternary complex category $[A\langle B\rangle C]$ is a specific substructural logic system. It not only shares structural rules such as associative law and commutative law, but also contains monotonicity and a variant of limited contraction, whose axiomatic counterpart is $A \bullet C \longrightarrow A \bullet (A \bullet C)$. This rule is capable of inserting the left-hand side category A and is exactly what is needed for characterization on zero pronominals. However, linguistic facts also demonstrate that category B in the anaphora slot may relate to categories outside the slot. In other words, the elided or covert category may bear an anaphoric relation with the category outside the slot. Thus, we stipulate those three categories in the anaphora slot are different from each other and propose the variant of contraction. Further research is needed on theoretical significance of this variant from the perspective of substructural logics.

Data Availability

The data used to support the analysis of the study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. Fox, *Economy and Semantic Interpretation*, The MTI Press, Cambridge, MA, USA, 2000.

[2] A. Nasta, "Variaties of Linguistic Economy," Doctoral dissertation, University of East Anglia, Norwich, England, 2014.

[3] W. Moortgat, W. Sijtsma, and A. von Horck, "Generalized quantification and discontinuous type constructors," *Discontinuous Constituency*, pp. 181–208, De Gruyter, Berlin, Germany, 1996.

[4] M. Moortgat, M. Sadrzadeh, and G. Wijnholds, "A frobenius algebraic analysis for parasitic gaps," *Journal of Applied Logics-IfCoLog Journal*, vol. 7, no. 5, pp. 823–852, 2020.

[5] G. Morrill, "Discontinuity in categorical grammar," *Linguistics and Philosophy*, vol. 18, no. 2, pp. 175–219, 1995.

[6] G. Morrill and O. Valentín, "Displacement logic for anaphora," *Journal of Computer and System Sciences*, vol. 80, no. 2, pp. 390–409, 2014.

[7] G. Morrill and O. Valentín, "A reply to Kubota and Levine on gapping," *Natural Language & Linguistic Theory*, vol. 35, no. 1, pp. 257–270, 2017.

[8] M. Hepple, "Hybrid categorial logics," *Logic Journal of IGPL*, vol. 3, no. 2-3, pp. 343–355, 1995.

[9] P. Jacobson, "Towards a variable-free semantics," *Linguistics and Philosophy*, vol. 22, no. 2, pp. 117–185, 1999.

[10] F. L. Jacobson, P. F. Judy, U. Feldman, and S. E Seltzer, "Perceived features reported as nodules: interpretation of spiral chest CT scans," *Academic Radiology*, vol. 7, no. 2, pp. 77–82, 2000.

[11] G. Jäger, *Anaphora and Type Logical Grammar*, Springer, Dordrecht, The Netherland, 2005.

[12] M. Corbalán and G. Morrill, "Overtly anaphoric control in type logical grammar," in *Proceedings of the 20th International Conferences on Formal Grammar*, pp. 183–199, Barcelona, Spain, August 2015.

[13] M. Corbalán, "A syntactic treatment for free and bound pronouns in a type-logical grammar," in *Proceedings of the Distributed Computing and Artificial Intelligence 14th International Conference*, pp. 223–230, Porto, Portugal, June, 2017.

[14] G. Wijnholds and M. Sadrzadeh, "A type-driven vector semantics for ellipsis with anaphora using lambek calculus with limited contraction," *Journal of Logic, Language and Information*, vol. 28, no. 2, pp. 331–358, 2019.

[15] L. McPheat, G. Wijnholds, M. Sadrzadeh, A. Correia, and A. Toumi, "Anaphora and ellipsis in lambek calculus with a relevant modality: syntax and semantics," *Journal of Cognitive Science*, vol. 22, no. 2, pp. 1–34, 2021.

[16] B. Coecke, M. Sadrzadeh, and S. Clark, "Mathematical foundations for distributed compositional model of meaning," *Linguistic Analysis*, vol. 36, no. 1, pp. 345–384, 2010.

[17] J. Krishnamurthy and T. Mitchell, "Vector space semantic parsing: a framework for compositional vector space models," in *Proceedings of the Workshop on Continuous Vector Space Models and their Compositionality*, pp. 1–10, Sofia, Bulgaria, August 2013.

[18] J. Maillard, S. Clark, and E. Grefenstette, "A type-driven tensor-based semantics for CCG," in *Proceedings of the EACL 2014 on Type Theory and Natural Language Semantics (TTNLS)*, pp. 46–54, Gothenburg, Sweden, April 2014.

[19] B. Coecke, E. Grefenstette, and M. Sadrzadeh, "Lambek vs. lambek: functorial vector space semantics and string diagrams for lambek calculus," *Annals of Pure and Applied Logic*, vol. 164, no. 11, pp. 1079–1100, 2013.

[20] G. Wijnholds, "A proof-theoretic approach to scope ambiguity in compositional vector space models," *Journal of Language Modelling*, vol. 6, no. 2, pp. 261–286, 2019.

[21] J. Lambek, "The mathematics of sentence structure," *The American Mathematical Monthly*, vol. 65, no. 3, p. 154, 1958.

[22] C. T. J. Huang, "On the distribution and reference of empty pronouns," *Linguistic Inquiry*, vol. 15, no. 4, pp. 531–574, 1984.

[23] O. Huang, O. Gaeggli, and K. Safir, "Pro-drop in Chinese: a generalized control theory," in *The Null Subject Parameter*, pp. 185–214, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1989.

[24] J. Baldridge and G. J. Kruijff, "Multi-modal combinatory categorial grammar," in *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics*, pp. 211–218, Budapest, Hungary, April 2003.

[25] F. Kong and H. Man, "Type logical grammar and anaphora in Mandarin," *Journal of Hunan University of Science & Technology (Social Science Edition)*, vol. 14, no. 2, 2011, in Chinese.