Toward Understanding Discourse Structure - Flow of Contextual Information in Spoken Korean

Suk-Jin Chang   and   *Jae-woong Choe*
Seoul National University   Koera University

1. Introduction

1.1. Contextual Information

Discourse information under consideration includes the following traditional pragmatic notions. Taken not into account in this paper is information concerning discourse participants' (DP) world knowledge, point of view, emotive states of affairs, and the like.[1]

```
(1) Contextual Information

    a. Deixis
    b. Speech Act (SA)  - Illocution
    c. Presupposition (PRSP)
    d. Implicature (IMPL)
    e. Discourse Function (DF)  - Focus (FOC)
                                - Topic (TOP)
```

1.2. Discourse Levels in Spoken Korean

A salient feature of spoken Korean is honorification. It is pragmatic in nature, imbued deeply into the grammar and the lexicon of the spoken language. It is grammaticalized in the form of sentence endings, reflecting the interpersonal relationship between the speaker and the hearer. Honorification also controls the subject-predicate agreement, constrained by the relationship between the speaker and the subject referent of the utterance. It also partially controls lexical selections. The interpersonal constraints between DPs are three dimensional: social, psychological and formal.

There are about half a dozen discourse levels in present-day Korean, each having a distinct sentence ending.

```
(2) Discourse Level[2] S-ending (Imperative)              Dimension
                                                Power  Solidarity  Formality
                                                (high) (loose)     (formal)
         Formal           -psio
         Polite           -yo
         Blunt            -o
         Familiar         -key
         Plain            -la
         Intimate         -e                   (low)   (tight)     (informal)
```

When the subject referent is honored by the speaker, relevant constituents in the discourse are realized in honorific form, including the systematic affixation of *si* to the verbal base.

(3) a. Kim-sensayng.nim-kkeyse   o. si. ess.upnita.   (Speaker < Hearer,
         teacher.HON-SBJ(HON) come.HON.PST.FO/DCL   Speaker < Subject)
      'Prof. Kim came.'
   b. Mia-ka   o. ass. ta.   (Speaker ≥ Hearer,
         SBJ come.PST.PL/DCL   Speaker ≥ Subject)
      'Mia came.'

2. Representing Discourse Information

2.1. Infons[3] and AVMs

We set up the following types of discourse infons (4) and speech-act infons (5). They are both represented in the AVM format in (4b) and (5b), respectively. The specific infons entered here will be made use of in the discourse analysis in Section 3.

(4) Discourse Infon Types[4]

- DP (= discourse participant), N (= name/address term),
  SX (= male/female), AT (= attribute)

a. Infons

DP: { ⟨person,a⟩ , ⟨person,b⟩ , ⟨person,c⟩ , ⟨person,d⟩ , ⟨person,e⟩ , ⟨person,f⟩ }

N: { ⟨named,a,kim-sensayng⟩ , ⟨named,b,tol⟩ , ⟨named,c,yong⟩
     ⟨named,d,mia⟩ , ⟨named,e,cheli⟩ , ⟨named,f,hia⟩ }

SX: { ⟨male,a⟩ , ⟨male,b⟩ , ⟨male,c⟩ , ⟨female,d⟩ , ⟨male,e⟩ , ⟨female,f⟩ }

AT: { ⟨teacher,a,b⟩ , ⟨student,b,a⟩ , ⟨friend,b,c⟩ , ⟨friend,c,d⟩ }

b. AVMs

DP: {RELN person, RELN person, RELN person, RELN person, RELN person, RELN person}
     ARG a   ARG b   ARG c   ARG d   ARG e   ARG f

N: {RELN   named   , RELN   named, RELN   named, RELN   named,
    BEARER a   BEARER b   BEARER c   BEARER d
    NAME   kim-sensayng   NAME   tol   NAME   yong   NAME   mia

    RELN   named, RELN   named}
    BEARER e   BEARER f
    NAME   cheli NAME   hia

SX: {RELN male, RELN male, RELN male, RELN female, RELN male, RELN female}
     ARG a   ARG b   ARG c   ARG d   ARG e   ARG f

AT: {RELN teacher, RELN student, RELN friend, RELN friend}
     ARG1 a   ARG1 b   ARG1 b   ARG1 c
     ARG2 b   ARG2 a   ARG1 c   ARG2 d

(5) Speech-act Infon Types

- DX (= deixis) , DL (= discourse level), IA (=illocutionary act)
- The prefix (*) (as in *x) is to indicate the prefixed object (x)
  is an indeterminate.

a. Infons

DX: { ⟨speaker,a⟩ , ⟨hearer,b⟩ , ⟨sbj-ref,c⟩ }

DL: { ⟨formal,a,b⟩ , ⟨polite,a,b⟩ , ⟨blunt,a,b⟩ ⟨familiar,a,b⟩ ,
      ⟨plain,a,b⟩ ⟨intimate,a,b⟩ , ⟨honor,a,b:0⟩ , ⟨honor,b,a⟩ ,
      ⟨honor,b,c:0⟩ }

2

IA: { ⟨assert,a,b,*ctt⟩ , ⟨question,a,b,*ctt⟩ , ⟨request,a,b,*ctt⟩ ,
⟨propose,a,b,*ctt⟩ }

b. AVMs

```
C-INDS SP a
       HR b
       UT. t
```

```
DL: {RELN formal,  RELN polite,  RELN blunt,  RELN familiar,  RELN plain,
     SP   a         SP   a        SP   a       SP   a          SP   a
     HR   b         HR   b        HR   b       HR   b          HR   b

     RELN intimate,  RELN      honor,   RELN      honor,  RELN      honor   }
     SP   a          HONORER   a        HONORER   b       HONORER   b
     HR   b          HONOREE   b        HONOREE   a       HONOREE   c
                     POL       0                          POL       0
```

```
IA: {RELN  assert,  RELN   question,  RELN   request,  RELN   propose}
    SP    a         SP    a           SP    a          SP    a
    HR    b         HR    b           HR    b          HR    b
    CTT   *ctt      CTT   *ctt        CTT   *ctt       CTT   *ctt
```

2.2. The CONTEXT Attribute

An expanded feature structure of the CONTEXT attribute is presented in full form in (6) and in short form in (7).

(6) CONTEXT (in full)

$$
\text{CONTEXT}\left[\begin{array}{l}
\text{SPEECH ACT}\left[\begin{array}{l}
\text{C-INDICES}\left[\begin{array}{ll}
\text{SPEAKER} & a \\
\text{HEARER} & b \\
\text{UTTERANCE TIME} & t_U
\end{array}\right] \\[1em]
\text{DISCOURSE LEVEL}\left[\begin{array}{ll}
\text{RELN} & \text{polite} \\
\text{SPEAKER} & a \\
\text{HEARER} & b
\end{array}\right] \\[1em]
\text{ILLOCUTIONARY ACT}\left[\begin{array}{ll}
\text{RELN} & \text{assert} \\
\text{SPEAKER} & a \\
\text{HEARER} & b \\
\text{CONTENT} & \ldots
\end{array}\right]
\end{array}\right] \\[1em]
\text{IMPLICATURE} \quad \{\ldots\} \\
\text{PRESUPPOSITION} \{\ldots\} \\
\text{DISCOURSE FUNCTION}\left[\begin{array}{l}
\text{FOCUS} \{\ldots\} \\
\text{TOPIC} \langle\ldots\rangle
\end{array}\right]
\end{array}\right]
$$
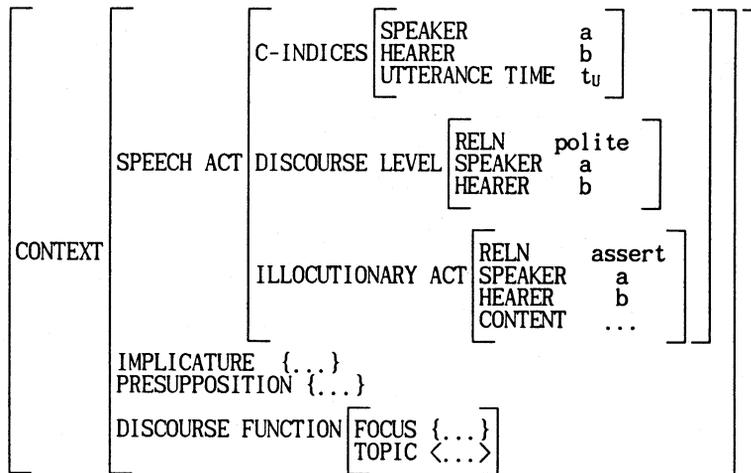
(7) CONTEXT (in short)

```
CXT SA C-INDS SP a
              HR b
              UT t_U
       DL RELN polite
          SP   a
          HR   b
       IA RELN assert
          SP   a
          HR   b
          CTT  ...
    IMPL {...}
    PRSP {...}
    DF FOC {...}
       TOP ⟨...⟩
```

3

3. Discourse Analysis

3.1. A Dialog

The following is a short dialog exchange between Kim (Teacher) and Tol (Student).

(8) Dialog
    a. Kim-sensayng: Nwu-ga Mia-lul ttayly.ess.nya?
                 who-SUBJ -OBJ   beat.PST.PL/INT
                 'Who beat Mia?'
    b. Tol   : Yong-i   ttaylyess.yo.
                   -SUBJ beat.PST.POL/DCL
                 'Yong did.'
    c.   Cheli-to ttayly.ess.eyo.
                     -too
                 'Cheli did, too.'
    d. Kim-sensayng: Ne-to ttayly.ess.nya?
                 you-too
                 'Did you beat her, too?'
    e. Tol   : Ce-nun an ttayly.ess.eyo.
                 I-SUBJ not
                 'I didn't beat her.'
    f.   Cheli-nun Hia-to ttayly.ess.eyo.
                 'Cheli beat Hia, too.'

3.2. Discourse Structure

Distinct from the structure analysis of a sentence or an utterance in isolation, that of a dialog makes use of a superimposed database, which we call discourse-initial data or CXT_{DI}. It consists of the discourse-infon types given in (4). They are the sort of background information minimally shared by the dialog participants, including their names, sexes, and some attributes or properties mutually known to them. We are not concerned with other types of background information such as DPs' world knowledge, their beliefs, their cognitive and emotive states. Those speech-act infons, which have to do with the DPs' role as speaker and hearer and their manner of speaking and illocutionary acts get in as soon as the dialog takes place. So they properly constitute SA, a subsort of CONTEXT in our system. And its subsort DL plays a crucial role in constraining the honorification of spoken Korean as overviewed earlier.

According as who the speaker is, the dialog D is divided up into two subsets: Kim's utterance set (U_k) and Tol's (U_T).

(9) a. $D = \{u_1, u_2, \ldots u_6\}$
    b. $U_k = \{u_1, u_4\}$
    c. $U_T = \{u_2, u_3, u_5, u_6\}$

The sequence of the dialog exchange is presented in (10).

(10)

$$
\begin{array}{ccc}
 & D & \\
 \swarrow & & \searrow \\
U_k & & U_T \\
\downarrow & & \downarrow \\
u_1 & \rightarrow & u_2 \\
 & & \downarrow \\
u_4 & \leftarrow & u_3 \\
 \searrow & & \\
 & \rightarrow & u_5 \\
 & & \downarrow \\
 & & u_6
\end{array}
$$
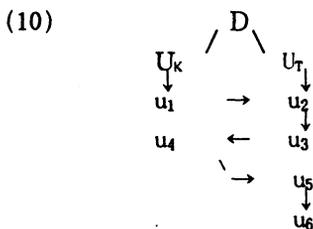
Making use of the information in the discourse-initial database, C-INDS and DL values of U_K and U_T get assigned as the dialog proceeds with the speaker-hearer role properly marked. They are the default values. The utterance times vary, either K's utterance time (t_{UK}) preceding T's utterance time (t_{UT}) or the other way around (i.e. '$t_{UK} < t_{UT}$' or '$t_{UK} > t_{UT}$').

(11) a. U_K: SA C-INDS SP a   b. U_T: SA C-INDS SP b

$$
\begin{array}{ll}
\text{(11) a. } U_K\text{: SA C-INDS SP a} & \text{b. } U_T\text{: SA C-INDS SP b} \\
\qquad\qquad\quad \text{HR b} & \qquad\qquad\quad \text{HR a} \\
\qquad\qquad\quad \text{UT } t_{UK} & \qquad\qquad\quad \text{UT } t_{UT} \\
\qquad\quad \text{DL RELN plain} & \qquad\quad \text{DL RELN polite} \\
\qquad\qquad \text{SP} \quad \text{a} & \qquad\qquad \text{SP} \quad \text{b} \\
\qquad\qquad \text{HR} \quad \text{b} & \qquad\qquad \text{HR} \quad \text{a}
\end{array}
$$

With the SA values inherited from the database, the IA value of each utterance u_i is assigned (i.e. asserted, questioned, etc.) with regard to the content expressed. The anaphoric or elliptical content (e.g. the pronominal or zero-pronominal) of the utterance is determined in part by the information in SA and also by the discourse functions, FOCUS and TOPIC, which mark the locus of information center and the hierachical order of centering (Chang 1986, 1989; Kameyama 1985, 1986; Walker et al 1990). The content (CTT) asserted or questioned in IA is then cumulatively augmented with implicatures and presuppositions in keeping with a default inheritance mechanism (Gazdar 1987). The principle of contextual inheritance and consistency (Chang 1990, Pollard and Sag (in press)[5]), which licenses flow of contextual information from Mother to Daughters within a sentence, is also at work here across the utterance and speaker-hearer boundaries. All the augmented information of both CONTENT and CONTEXT of u_i is formed into CXT_i^+, the cumulative CXT_i, which is input to u_{i+1}. The structure of information flow is quadri-directional in the sense that it is up-and-down and left-and-right between the two DPs engaged in dialog exchange. When the dialog terminates at u_6, the contextual information incrementally shared by the DPs amounts to the same. In this case it is the teacher who comes to share the information To1 already had.

The information flow in the dialog exchange is schematized roughly as in (12). (See next page.)

5

(12)

Discourse Initial Data

CXT_{DI} DP   ...
     N   ...
     SX   ...
     AT   ...

Kim's Discourse Data

...
...
...

u_1

CTT
CXT SA C-INDS
      DL
       IA
   IMPL
   PRSP
   DF FOC
     TOP

u_2

CTT
CXT SA C-INDS
      DL
       IA
   IMPL
   PRSP
   DF FOC
     TOP

u_3

CTT
CXT SA C-INDS
      DL
       IA
   IMPL
   PRSP
   DF FOC
     TOP

u_4

CTT
CXT SA C-INDS
      DL
       IA
   IMPL
   PRSP
   DF FOC
     TOP

u_5

CTT

CXT SA C-INDS
      DL
       IA
   DF FOC

     TOP

u_6

CTT
CXT SA C-INDS
      DL
       IA
   IMPL
   PRSP
   DF FOC
     TOP

Tol's Discourse Data

...
...
...
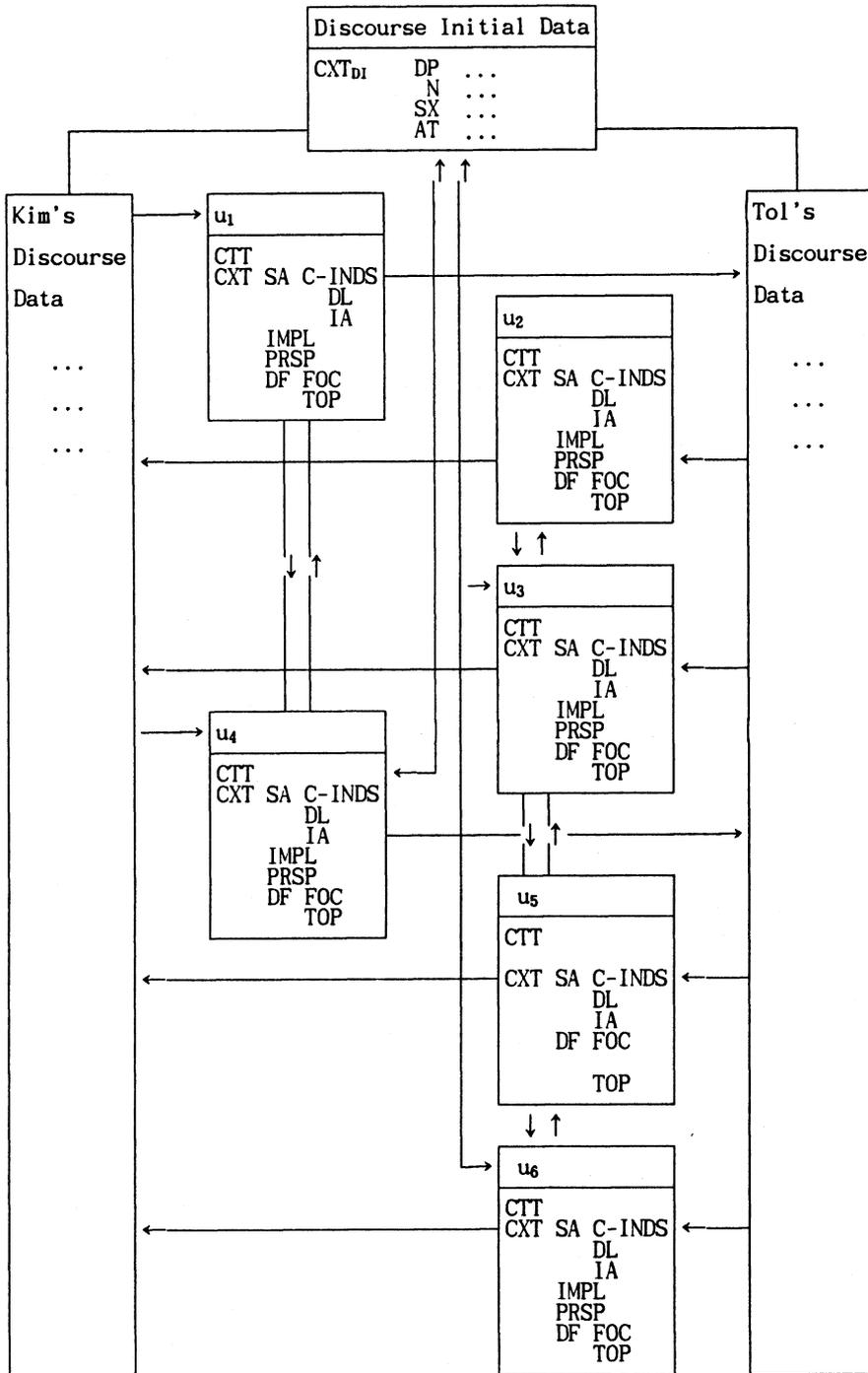
Figure: Flow of Contextual Information

3.3. Analysis and Annotations

A circled symbol (e.g. ©) is a tag indicating a token identity of feature structures. Some remarks are entered on the right column.

(13)

CXT_{DI}: DP, N, SX, AT — % Discourse Initial Database

u_1: *Nwu-ka Mia-lul ttayly.ess.n ya?* — % from CXT_{DI}
  from U_k

```
CXT SA C-INDS SP   a
             HR   b
             UT   t_U1
        DL RELN plain
           SP   a
           HR   b
        IA RELN question          % grammaticalized: nya-
           SP   a                    ending, together with the plain DL
           HR   b
           CTT ©                  % CTT is what is presupposed.
      PRSP©RELN ∧
         SOA-ARG1 RELN beat
                  BEATER *x
                  BEATEN d
                  ET     t_E1
         SOA-ARG2 RELN precede    % grammaticalized: past tense ess
                  ET     t_E1
                  UT     t_U1
      DF FOC {nwuka}
         TOP ⟨mia-lul, ttaylyessnya⟩   % mia-lul: possible Cb in u_2
```

CXT_1^+ (= CXT_{DI} + SA_1 + $PRSP_1$ + DF_1) — % CXT_1^+ (cumulative CXT_1)

u_2: *Yong-i ttaylyesseyo.* — % from U_T (speaker-hearer shift.
  from CXT_{DI}

```
CTT©RELN ∧                        % from Tol's database
     SOA-ARG1 RELN beat
              BEATER c
              BEATEN d            % from DF
              ET     t_E1
     SOA-ARG2 RELN precede        % grammaticalized: past tense
              ET     t_E1
              UT     t_U2
CXT SA C-INDS SP b
             HR a
             UT t_U2
        DL RELN polite
           SP   b
           HR   a
        IA RELN assert            % grammaticalized: yo-ending,
           SP   b                   together with polite DL[6]
           HR   a
           CTT ©                  % assertion of CTT_2
        DF FOC {yong-i}           % FOCUS_1 value-assigned
           TOP ⟨ φ, ttaylyesseyo⟩ % zero = Cb:
                                    i.e. SLASH: {mia-lul}in u_1[7]
```

CXT_2^+ (= CXT_{DI} + CTT_2 + SA_2 + DF) — % CXT_2^+ (cumulative CXT_2)

u_3: *Cheli-to ttaylyesseyo.* — % from U_T
  from CXT_{DI}

```
CTT©RELN ∧                        % from Tol's database
     SOA-ARG1 RELN beat
              BEATER e
              BEATEN d            % from DF
              ET     t_E3
```

```
            SOA-ARG2 RELN precede              % grammaticalized: past tense
                     ET    t_E3
                     UT    t_U3
CXT SA C-INDS SP b
              HR a
              UT t_U3
     DL RELN polite
        SP    b
        HR    a
     IA RELN assert                            % grammaticalized: yo-ending,
        SP    b                                  together with polite DL
        HR    a
        CTT Ⓒ                                  % assertion of CTT_3
IMPL:⁸ QUANTS ⟨(∃ⓧ¦{person(ⓧ)})⟩              % implicature: someone other
      NUCLEUS RELN ∧                             than Cheli beat Mia;
              SOA-ARG1 RELN beat                 the indeterminate (*x) is
                       BEATER ⓧ*x                anchored to Yong in u_2; so
                       BEATEN d                  the implicature is simply
              SOA-ARG2 RELN ≠                    CXT_2 in u_2.
                       ARG1 ⓧ
                       ARG2 e
     DF FOC {chelito}
        TOP ⟨ φ, ttaylyesseyo⟩                 % = TOP_2 (zero = Cb; SLASH: {mia-lul})
```

```
CXT_3⁺ ( = CXT_DI + CTT_3 + SA_3 + IMPL_3 + DF_3)    % CXT_3⁺ (cumulative CXT_3)

u_4: Ne-to ttaylyessnya?                        % from U_k (speaker-hearer shift)
```

```
CXT SA C-INDS SP a
              HR b                              % from SA, 2nd PER pronoun
              UT t_U4                             ne is identified as referring to b.
     DL RELN plain
        SP    a
        HR    b
     IA RELN  question                          % grammaticalized: nya-ending
        SP    a
        HR    b
        CTT   Ⓒ
CTT Ⓒ RELN ∧                                    % CTT questioned
     SOA-ARG1 RELN beat
              BEATER b                           % from CXT_1 and SA: 2nd PER
              BEATEN d                            pronoun ne identified as
              ET    t_E4                          referring to b
     SOA-ARG2 RELN precede                       % grammaticalized: past tense
              ET    t_E4
              UT    t_U4
     IMPL QUANTS ⟨(∃ⓧ¦{person(ⓧ)})⟩             % implicature: someone other
          NUCLEUS RELN ∧                           than Tol beat Mia; the
                  SOA-ARG1 RELN beat               indeterminate (*x) is
                           BEATER ⓧ*x             anchored to Yong and Cheli
                           BEATEN d                in the cumulative CXT_4.
                  SOA-ARG2 RELN ≠
                           ARG1 ⓧ
                           ARG2 b
     DF FOC {ne-to}
        TOP: ⟨ φ, ttaylyessnya⟩               % = TOP_3 (zero = Cb; SLASH {mia-lul})
```

```
CXT_4⁺ ( = CXT_4 + CTT_4 + IMPL_4 )            % CXT_4⁺ (cumulative CXT_4)

u_5:  Ce-nun an ttaylyesseyo.                   % from U_T (speaker-hearer shift
```

```
CTT Ⓒ RELN ∧
     SOA-ARG1 RELN beat
              BEATER b                           % from CXT_DI and SA; 1st PER
              BEATEN d                            pronoun ce identified as
              ET    t_E5                          referring to b.
              POL   0
     SOA-ARG2 RELN precede                       % grammaticalized: past tense
              ET    t_E5
              UT    t_U5
CXT SA C-INDS SP b
              HR a
```

8

```
                    UT  t_U5
            DL  RELN  polite
                SP      b
                HR      a
            IA  RELN      assert
                SP        b
                HR        a
            CTT      ©
    IMPL  QUANTS ⟨( ∀ⓧ|{person(ⓧ)})⟩          % implicature: Tol is not
          NUCLEUS RELN ∧                          one of those who beat Mia;
                SOA-ARG1 RELN beat                the indeterminate (*x) is
                         BEATER ⓧ*x               anchored to Yong and Cheli
                         BEATEN d                 in CXT_2 and CXT_4, respectively.
                         POL    1
                SOA-ARG2 RELN ⧧
                         ARG1 ⓧ
                         ARG2 b
    DF  FOC {ce-nun, an}                        % ce-nun: contrastive[9]
        TOP ⟨0, ttaylyessyo⟩                    % = TOP_4 (zero = Cb; SLASH {mia-lul})
```

CXT₅⁺ (= CTT₅ + SA + IMPL₅ + DF₅)	% CXT₅⁺ (cumulative CXT₅)

u_6: *Cheli-nun Hia-to ttaylyesseyo.*   % from CXT_DI
                                          from U_T

```
CTT©RELN ∧
    SOA-ARG1 RELN beat
             BEATER e
             BEATEN f
             ET      t_E6
    SOA-ARG2 RELN precede                    % grammaticalized: past tense
             ET      t_E6
             UT      t_U6
CXT SA C-INDS SP b
              HR a
              UT t_U6                        % UT overrides UT_5
    IMPL  QUANTS ⟨( ∃ ⓧ|{person(ⓧ)})⟩       % implicature: Cheli beat
          NUCLEUS RELN ∧                        someone other than Hia;
                SOA-ARG1 RELN beat              the indeterminate (*x)
                         BEATER e               is anchored to Mia in
                         BEATEN ⓧ*x             CXT_3.
                SOA-ARG2 RELN ⧧
                         ARG1 ⓧ
                         ARG2 f
    DF  FOC {hia-to}
        TOP ⟨cheli-nun, ttaylyesseyo⟩        % Cb shift;
                                               Cb = cheli-nun, from DF_3 in u_3
```

CXT₆⁺ (= CXT₆ + CTT₆ + SA + IMPL₆ + DF₆)	% CXT_T⁺ (cumulative CXT₆)

```
Information shared by Kim and Tol (as a result of the dialog):
  Plus CONTEXT-1:
    - Yong beat Mia.
    - Cheli beat Mia.
    - Tol didn't beat Mia.
    - Cheli beat Hia.
```

In concluding the analysis of the dialog, the last utterance, u_6 is given a full-scale AVM, including CATEGORY, CONTENT, and CONTEXT, followed by a list of annotations.

(14) Utterance-6

Tol: *Cheli-nun Hia-to ttayly.esse.yo.*

```
PHON /cheli,nun,hia,to,ttayli,esse,yo/
SYNSEM LOC CAT HEAD ⓗ
               SUBCAT ⟨ ⟩
           CTT ©
```

```
                CXT ⊗
DTRS H-DTR PHON /ttayli,esse.yo/
           SYNSEM LOC CAT HEAD ⓗPOS V
                            VFORM TENSE PST:ⓣ
                                  FIN   POL/DCL:①, ⓘ
                        SUBCAT ⟨①NP, ②NP⟩¹⁰
                    CTT ©{RELN beat, ⓣRELN precede}
                        BEATER e      ET      t_E6
                        BEATEN f      UT      t_U6
                        ET     t_E6
                    CXT⊗{SA C-IND SP b
                               HR a
                               UT t_U6
                         DL①RELN polite
                           SP   b
                           HR   a
                         IAⓘRELN assert
                           SP   b
                           HR   a
                           CTT  ©
                    IMPL QUANT ⟨(∃⊗!{person(⊗)})⟩
                         NUCLEUS RELN ∧
                                 SOA-ARG1 RELN beat
                                          BEATER e
                                          BEATEN ⊗≠x
                                 SOA-ARG2 RELN ≠
                                          ARG1 ⊗
                                          ARG2 f
                    PRSP { }
                    DF TOP
     C1-DTR PHON /cheli,nun/
         ①SYNSEM LOC CAT HEAD ⓗ
                         SUBCAT ⟨ ⟩
                         MARKING ⓜ
                   CTT ©
                   CXT ⊗
           DTRS H-DTR PHON /cheli/
               SYNSEM LOC CAT ⓗHEAD POS N
                                 GF SBJ
                             SUBCAT ⟨ ⟩
                   CTT ©PARA IND e PER 3
                                   NUM sg
                   CXT⊗CXT_DI(e)
           M-DTR PHON /nun/
               SYNSEM LOC CAT HEAD POS P
                               SPEC N[DF TOP]
                             SUBCAT ⟨ ⟩
                             MARKING ⓜnun
                   CTT { }
                   CXT {DF TOP}
     C2-DTR PHON /hia,to/
         ②SYNSEM LOC CAT HEAD ⓗ
                         SUBCAT ⟨ ⟩
                         MARKING ⓜ
                   CTT ©
                   CXT ⊗
           DTRS H-DTR PHON /hia/
               SYNSEM LOC CAT ⓗHEAD POS N
                                 GF OBJ
                             SUBCAT ⟨ ⟩
                   CTT´PARA IND f PER 3
                                   NUM sg
                   CXT ⊗CXT_DI(f)
               M-DTR PHON /to/
```

10

```
SYNSEM LOC CAT HEAD POS P
                    SPEC N[DF FOC]
                    SUBCAT < >
                    MARKING ⓜto
              CTT { }
              CXT {...DF FOC}¹¹
```

Annotations:

Most remarkable in the cross-utterance analysis of u_6 is the inheritance of the discourse-initial context CXT_{DI} by default. The $CXT_{DI}(e)$ and $CXT_{DI}(f)$ given as context values to C-1 and C-2 in the AVM indicate that the contextual information about the third-person referents (e, f) are inherited from CONTEXT-1, which is anchored to CONTENT as indices e and f. The context values of e and f then get inherited locally to the respective phrases *Cheli-nun* and *Hia-to* by the Context Inheritance Principle as tagged by the prefix ⊗.

Operations of unification applied in (14), which are theory-internal to Head-driven Phrase Structure Grammar (HPSG), include the following:

1. ①,②: Head daughter's SUBCAT value is the concatenation of Mother's SUBCAT list with the list (in order of increasing obliqueness) of the SYNSEM values of the COMP daughters--by the Subcategorization Principle.

2. ⓗ: Mother's Head feature value is structure shared with that of the Head daughter--by the Head Feature Principle.

3. ©: Mother's CONTENT value is structure shared with that of the (semantic) Head daughter--by the Semantics Principle.

4. ⊗: Mother's CONTEXT value is structure shared with those of the daughters-- by the Context Inheritance Principle (Chang 1990) or equivalently, by the Principle of Contextual Consistency (Pollard and Sag (in press)).

5. ⓜ: Mother's Marking value is structure shared with that of the Marker daughter--by the Marker-Head Schema.

6. ⓣ: The time relation in CONTENT is mapped to TENSE|PST in CATEGORY as its content value (Chang 1990).

7. ⓛ,ⓘ: The DL and IA values are mapped respectively to the verb-final SL and ST forms in CATEGORY as their context values (Chang 1990).

8. DF marking: C-2 daughter is marked as FOC; C-1 and Head daughters are marked as TOP. The Context Inheritance Principle licenses unification of daughters' DF values with mother's, i.e.:

    CXT|DF   FOC {mia-to}
              TOP <cheli-nun,ttaylyesseyo>

9. Morphological operations involved are:
    a. Past Tense: f_{PST}(VFORM[PREFINAL]) = *ess*
    b. S-ending: $f_{POL/DCL}$(VFORM[FINAL]) = *yo*

4. Conclusions

The dialog analysis presented here is an attempt to expand HPSG beyond a sentence boundary. Conspicuous for the approach taken here are the following.

First, flow of contextual information in a dialog exchange is quadri-directional, that is: up-and-down and left-and-right between the discourse participants, originated in the discourse-initial database, CXT_{DI}. It reduces the amount of grammatical description in each constituent utterance. Second, in keeping with the mechanism of default inheritance the contextual information cumulatively expands across the utterance and speaker-hearer boundaries, assembling

content and various sorts of context, speech acts, implicatures, and presuppositions. Third, two discourse functions, focus and topic are made available for an account of anaphora and ellipsis as well as preference of interpretation. Fourth, two subsorts of speech acts, the discourse level and the illocutionary act are mapped to sentence endings in spoken Korean.

Details of HPSG-based Korean grammar are not dealt in here, as we are concerned not with a sentence-based grammar but a discourse grammar. HPSG itself has been constantly revised and expanded. As it stands, HPSG is an information-based grammar well-founded for exploration of discourse understanding. We have begun to implement the HPSG-based Korean grammar (Chang 1992) using GULP, a preliminary version of which will come out soon.[12]

Notes

1. The 'presupposition' is delimited to cover only what is called 'pragmatic' presupposition' (Keenan 1971), which is cancellable and nonconventional; by contrast, the 'implicature' is used here to cover only the 'conventional' one, which is noncancellable and conventional, although the distinction between the two has not always been clear-cut (Karttunen and Peters 1977, Gazdar 1979, Levinson 1983). The 'point of view' is also known as 'empathy' (Kuno 1976), and as 'identification' (Kameyama 1985) or 'ID' (Chang 1986). The 'topic', misleading as it may be, is 'presuppositional' and in complementary to 'focus' (Chang 1989, 1990). A list of topics, hierarchically ordered, correspond functionally to the 'forward-looking centers' (Cfs) of the centering theory (Grosz et al 1983; Kameyama 1985, 1986; Chang 1986; Walker et al 1990), of which the most prominent center becomes the 'backward-looking center' (Cb) in the succeeding utterance.
   <topic, subject, object, object2,...>
2. The terms used for discourse levels are originally from Martin (1954), modified in Chang (1973). Martin's Yale Romanization is used here.
3. The 'infon', a unit of information (Devlin 1991), is formerly known in situation semantics (Barwise and Perry 1983, Barwise 1989) as 'soa (states of affairs)'.
4. It is only a partial list of discourse infon types. Polarity is not entered when its value is positive.
5. The Context Inheritance Principle guarantees that the CONTEXT value of a phrase in a headed structure is the union of the CONTEXT values of its daughters. In Pollard and Sag (in press) it is called the Principle of Contextual Consistency.
6. The polite declarative ending $-yo$ must be further marked as falling ($\downarrow$) in terminal contour (TC), in contrast to the interrogative ending yo$\uparrow$ (marked as rising in TC).
7. The gap is filled with a topic in u_1 (i.e. *Mia-lul*). So it is inherited as an element of TOPIC. In terms of centering, the Cb of u_2 is the zero pronominal (topic) referring to *Mia-lul* in u_1. In HPSG's term it is the value of the NONLOCAL|INHERITED|SLASH attribute.
8. In the implicature 'some X beat Mia', X is 'anchored' to Yong in u_2. Another possible interpretation of u_3 is: 'Yong beat Cheli, too'. This reading, however, is not preferred as it requires a Cb shift -- a shift from Mia (the Cb in u_2) to Yong. Continuation of the same Cb is preferred according to the theory of centering (Grosz et al 1983). The Cb, the null element in u_3 continues to be *Mia-lul,* the object referent. For the second reading, 'Yong-un' (as the new Cb) must be mentioned overtly as in 'Yong-un Cheli-to ttaylyesseyo.'
9. The *nun*-marked subject NP (*ce-nun*) is a contrastive focus, not a topic and as such it bears a contrastive and exclusive implicature in U_5:
   $\forall x[person(x) \rightarrow beat(x,mia) \wedge x \neq b]$
   Notice also that the *nun*-marked subject NP in u_6 (*Cheli-nun*) is the primary topic, the Cb of U_6. It looks back to 'Cheli' mentioned in u_3.
10. In Pollard and Sag (in press), the SUBCAT list does not contain the subject. It is no longer considered a complement. This view is not adopted in this paper, though.
11. The implicature meaning carried in the particle *to* 'also' (e.g. $\lambda Q[\lambda P[\exists x,y[Qx \wedge Px] \wedge \neg(x=y)]])$ should be entered here. Instead, it is given in the head daughter's CONTEXT.
12. For an implementation of HPSG/K (HPSG-based Korean grammar for NLP -Chang 1992) GULP 2.0 (Graph Unification Logic Programming; Covington 1989) is being used in collaboration with Y. Min (University of Arizona).

References

Barwise, J. (1989) *The Situation in Logic.* Stanford: CSLI, Stanford University.

Barwise, J. and J. Perry. (1983) *Situations and Attitudes.* Cambridge, MA: MIT Press.

Brennan, S. E., M. W. Friedman, C. J. Pollard. 1987. A Centering Approach to Pronouns. in *Proceedings of the 25th Annual Meeting of the American Associatioin of Computational Linguistics*, 156-162.

Chang, Suk-Jin (1973) *A Generative Study of Discourse: Pragmatic Aspects of Korean with Reference to English Language Research* 9:2 (Supplement).

Chang, Suk-Jin. (1986) "Discourse Function of Anaphora: on Reflexives" [in Korean]. *Hangul* 194. 121-155.

Chang, Suk-Jin. (1989) "Understanding Korean Discourse: Topic and Focus," Paper presented at the International Conference on Korean Language Education, Monteray, 7/19-21/1989.

Chang, Suk-Jin. (1990) "Discourse Context and Grammar" [in Korean]. *Eno* 15.499-538.

Chang, Suk-Jin. (1992) *Korean Grammar: HPSG/K for NLP* [in Korean]. Report NO. COSMOS-92-2.

Covington, M. A. (1989) *GULP 2.0: An Extension of Prolog for Unification-Based Grammar.* Research Report AI-1989-01.AI Research Group, University of Georgia.

Devlin, K. (1991) *Logic and Information, Volume 1: Infons and Situations.* Cambridge: Cambridge University Press.

Gazdar, G. (1987) "Linguistic Applications of Default Inheritance Mechanism," in Whitelock et al. eds. *Linguistic Theory and Computer Applications.* (New York: Academic Press), 37-67.

Grosz, B. J., A. K. Joshi and S. Weinstein. (1983) "Providing a Unified Account of Definite Noun Phrases in Discourse," in *Proceedings of the 21st Annual Meeting of the American Computational Linguistics*, 44-50.

Grosz, B. J. and C. L. Sidner. (1985) *The Structure of Discourse.* CSLI 85-39. Stanford: CSLI, Stanford University.

Kameyama, M. (1985) *Zero Anaphora in Japanese.* Ph.D. dissertation, Stanford University.

Kameyama, M. (1986) "A Property-Sharing Constraint in Centering," in *Proceedings of the 24th Annual Meeting of the American Association of Computational Linguistics*, 200-206.

Karttunen, L. and P. S. Peters. (1977) "Conventional Implicature," in Oh and Dineen, eds. *Syntax and Semantics 11* (New York: Academic Press), 1-56.

Keenan, E. L. 1971. "Two Kinds of Presupposition in Natural Language," in C. Fillmore and Langendoen, eds.(1971). *Studies in Linguistic Semantics* (New York: Holt), 45-54.

Kuno, S. (1976) "Subject, Theme, and the Speaker's Empathy: A Reexamination of Relativization Phenomena," in C.N. Li, ed., *Subject and Topic* (New York: Academic Press), 417-444.

Levinson, S. C. (1983) *Pragmatics.* Cambridge: Cambridge University Press.

Martin, J. E. (1954) *Korean Morphophonemics.* Baltimore: Linguistic Society of America.

Pollard, C. and I. Sag. (1987) *Information-Based Syntax and Semantics, Volume 1: Fundamentals.* Stanford: CSLI, Stanford University.

Pollard, C. and I. Sag. (in press). *Information-Based Syntax and Semantics, Volume 2: Agreement, Control and Binding.* Stanford: CSLI, Stanford University.

Searle, J. R. 1969. *Speech Acts.* Cambridge: Cambridge University Press.

Walker, M., M. Lida, and S. Cote. (1990) "Centering in Japanese Discourse," in *Proceedings of the 28th Annual Meeting of the American Association of Computational Linguistics.*