Anaphoric Marking in Bare Noun Languages:
When and Why Demonstratives Are Needed

Min-Joo Kim

Department of English

Texas Tech University, USA

Setting the stage

- Languages like English have a definite article (e.g., *the*).

- But many languages don't (e.g., Mandarin Chinese, Japanese, Korean, Thai, Vietnamese, Shan).

- **The overarching question:** How do article-less languages express various types of definiteness?

Two types of definites (Schwarz 2009)

- **Unique/weak definites:**
    (1) Open <u>the door</u>!   (Immediate situation use in the sense of Hawkins 1978)
    (2) a. <u>The moon</u> has risen.   (Larger/immediate situation use)

        b. <u>The moon</u> orbits Earth.   (Larger situation use)

- **Anaphoric/familiar/strong definites:**
    (3) John bought a $book_1$ and a magazine. <u>The $book_1$</u> was expensive.
    (4) Every farmer that has a $donkey_1$ beats <u>the $donkey_1$</u>.
                                        (Schwarz 2009: 3, (3); 132, (95a))

What has been said about article-less languages

- They use bare nouns to mark both types of definites.

- But they may **optionally** use a demonstrative (Dem) to mark **anaphoric** definites (e.g., Ahn 2019; Moroney 2021; Park 2022; Simpson and Wu 2022; Dayal and Jiang 2023) (compare Jenks 2018).

- In brief:
  - Unique/weak definites: **Bare nouns**
  - Anaphoric/familiar/strong definites: **Bare/Dem-NPs**

Illustration from Mandarin

(5) Unique/weak definite

Yueliang	sheng	shang	lai	le.
moon	rise	up	come	LE

'<u>The moon</u> has risen.'

(Chen 2004: 1165)

(6) Anaphoric/familiar/strong definite

a.

Jiaoshi	li	zuo	zhe	yi	ge	nansheng	yi	ge	nüsheng,
classroom	inside	sit	PROG	one	CLF	boy	one	CLF	girl

'There is a boy and a girl sitting in the classroom...'

b.

Nüsheng	zuo	zai	**nansheng**	pangbian.
girl	sit	DUR	boy	side

'<u>The girl</u> is sitting next to <u>the boy</u>.'

(adapted from Dayal and Jiang 2023)

The problem

- Sometimes a Dem is necessary for an anaphoric definite, as shown in (7), and currently, there is no consensus as to when a Dem is needed and why.

- Illustration from Mandarin (Jenks 2018: (16a,b)):

(7) a. Jiaoshi   li   zuo zhe   yi   ge   nansheng yi   ge   nüsheng,
    classroom inside sit   PROG   one   CLF   boy   one CLF   girl
    'There is a boy and a girl sitting in the classroom…'

  b. Wo zuotian   yudao   **#(na   ge)   nansheng**.
    I   yesterday   meet   that   CLF   boy
    'I met <u>the boy</u> yesterday.'

The goal of this talk

- To provide an answer to when and why a Dem may be needed for marking anaphoric definites by looking at Korean in comparison with a few other bare noun languages (e.g., Mandarin Chinese, Japanese).

- **The main claim:**
  - Some Dem-NPs are "optional", but some are necessary, and the latter occur as "the last resort" (compare, a.o., Ahn 2019).
  - And they are needed not just for contrast or anti-uniqueness effects (compare Dayal and Jiang 2023).
  - And even the "optional" ones are not truly optional in that they carry pragmatic meanings different from bare nouns'.

Outline of the talk

- Identify environments where an anaphoric Dem is needed

- Brief explanation as to why

- Quick typological investigation

- Brief formal semantic analysis

- Conclusion

When is a DEM needed?

- **Dayal and Jiang's (2023) answer**:

  - In contexts like (6), the situation described by (a) is defined by two individuals, and the *same* two individuals define the situation in (b). Hence, bare nouns can be used in (6b).

  - In contexts like (7), the situations described by (a) and (b) have *different* participants. So, in (7b), a Dem-NP is preferred over a bare noun.

Problematic data to Dayal and Jiang (2023)

- Anaphoric bare nouns can appear even when the initial context is expanded to include *additional* participants.

(8) a. Jiaoshi   li   zuo zhe   yi   ge   nanhai he   yi   ge   nühai.
    classroom inside sit PROG one   CLF   boy   and   one   CLF   girl
    'A boy and a girl were sitting in the classroom.'
  b. Turan   yi   ge   xiaohai pao jin   jiaoshi   jiao **nanhai** gen   ta   chuqu.
    suddenly one CLF kid   run   in   classroom ask   boy   with   him   go-out
    'Suddenly, a kid ran into the classroom and asked <u>the boy</u> to go out with him.'

- Dayal and Jiang (2023: 163) state that (8a) is "minimally" expanded in (8b) via a "controlled" introduction of an individual, but they do not explain exactly how (8) differs from (7).

When is a DEM needed? Cont'd

- **Simpson and Wu's (2022) answer**:
  - Anaphoric bare nouns occur when the *same* spatiotemporal location is maintained in a narrative sequence.
  - Dem-NPs occur when there is a *shift* in the spatiotemporal location.

- This line of analysis captures the contrast between (6)/(8) and (7).

Problematic data to Simpson and Wu (2022)

- Even if the same spatiotemporal location is maintained in a narrative sequence, a Dem-NP may be required (Kim 2021).

(9) Illustration from Korean:
    a. Ecey   thipi   chwukkwu   cwungkey-eyse   Son Heung-Min-ul   pwassta.
       yesterday television   soccer   broadcasting-LOC   Heung-Min Son-ACC   saw
       'Yesterday, on the soccer broadcast, I saw Heung-Min Son.'
    b. **#(Ku)**   **senswu-ka**   mom-nollim-i   kacang   ppallassta.
       that   player-NOM   body-movement-NOM   most   was.quick
       '<u>That player</u> (i.e., Heung-Min Son) exhibited the fastest performance of all.'

Problematic data to Simpson and Wu Cont'd

- An anaphoric bare noun may be licensed even if the same spatiotemporal location is *not* maintained in the narrative sequence.

(10) Illustration from Korean:

  a. Kyosil   an-ey   sonyen-kwa   sonye-ka   ancaissta.

     classroom   inside-LOC   boy-and   girl-NOM   sitting

     'A boy and a girl are sitting in a classroom.'

  b. Ecey   **sonyen**-i   **sonye**-lul   mannassta.

     yesterday   boy-NOM   girl-ACC   met

     'The boy met the girl yesterday.'

Why is a Dem needed in (9)?

- A Dem is needed because the anaphoric definite's referent is *not* the only *x* that has *P* in the *resource situation* (s_r) in the sense of Schwarz (2009); if it is, then a Dem is not needed, as in (11).

(11) a. Ecey   nolay.calang-eyse   yeca.ai   namca.ai kuliko elun   han-myeng-ul   pwassta.

   yesterday singing.contest-LOC girl   boy   and   adult one-CLF-ACC   saw

   'Yesterday, in a singing contest, I saw a girl, a boy, and a grown-up person.'

   b. **(Ku) yeca.ai-ka**   nolay-lul   kacang cal   hayssta.

   that girl-NOM   singing-ACC   most   well   did

   'The/that girl sang the best among them.'

Why is a Dem not needed in (10)?

- In data like (10), bare nouns are used as *text-internally licensed quasi-names* (Kim 2021).

- **Support for this idea:** In story-telling contexts, anaphoric bare nouns can freely occur acting like proper names (cf. Jenks 2015; Ahn 2019).

(12) a. Yes-nal   enu   swup-sok-ey   thokki   han   mali-ka   sal-ko   iss-ess-e.
       old-day   some   forest-inside-LOC   rabbit   one   CLF-NOM   live-CONN   existed
       'Once upon a time, in some forest, there lived a rabbit.'
    b. **Thokki-nun**   meli-ka   acwu   coh-ass-e.
       rabbit-TOP   brain-NOM   very   was.good
       'The rabbit was very smart.'

Why is a Dem not needed in (10)? Cont'd

- Even in non-fable contexts, anaphoric bare nouns can be used as text-internally licensed quasi-names.

(13) a. Cinanswu-ey   kangaci han   mali-lul   ipyanghayssta.

     last.week-LOC   puppy   one   CLF-ACC   adopted

     'Last week I adopted a puppy.'

   b. Onul   **kanagci-ka**   salaciessta.

     today   puppy-NOM   disappeared

     'Today <u>the puppy</u> disappeared.'

BUT sometimes it may take some work

(14) Paykman-nyen-ina   cwuk-ci ahn-un   **koyangi-ka**
one.million-year-as.much.as   die-CONN neg-REL.PRF   cat-NOM
iss-ess-supni-ta.   (S_1)
exist-PST-ALLO.HON-DECL
'Once upon a time, there was a cat which had lived for as long as one million years.'
Ku koyangi-nun   paykman-pen-ina   cwuk-ko
that cat-TOP   one.million-time-as.much.as   die-CONN
paykman-pen-ina   sal-ass-supni-ta.   (S_2; Dem)
one.million-time-as.much.as   live-PST-ALLO.HON-DECL
'The cat had died as many as one million times but re-lived as many as one million times.'
Kuleten   han-ttay   **koyangi-nun**   …   (S_6; Bare)
then   one-time   cat-TOP   …
'Then one day, the cat …'
(https://market.bookshopmap.com/100만-번-산-고양이/, accessed May 24, 2021)

Interim summary

- In a narrative sequence instantiating discourse anaphora:

  - A Dem is not needed for an anaphoric definite if the sentences in the narrative sequence share the spatiotemporal location, but it may still be needed for *situation-internal identifiability* (e.g., (9)).

  - If the sentences do not share the spatiotemporal location, then a Dem is needed for an anaphoric definite *unless* the definite acts like a *quasi-name* at the *text level* (e.g., compare (7) and (10)).

- **In short:** The *spatiotemporal relation* between the situations matters, and *situation-internal* or *text-internal identifiability* matters.

Quick typological investigation

- I tested the idea I just outlined against five genetically-unrelated bare noun languages by using Korean as the point of departure (Kim 2022):

  - Korean (Language Isolate)
  - Mandarin (Sino-Tibetan)
  - Japanese (Language Isolate)
  - Slovenian (South Slavic)
  - Turkish (Turkic)

Templates used for data collection

#1: 'There's a boy and a girl sitting in the classroom. <u>The girl</u> is sitting next to <u>the boy</u>.'
(No spatiotemporal location change)

#2: 'There's a boy and a girl sitting in the classroom. I met <u>the boy</u> yesterday.'
(Spatiotemporal location change; object position)

#3: 'Yesterday I saw a dog. <u>The dog</u> was cute.'
(No spatiotemporal location change necessary)

#4: 'Yesterday I saw a dog. <u>The dog</u> came to my house today.'
(Spatiotemporal location change; subject position)

#5: 'Yesterday I saw Heung-Min Son on TV. <u>That player</u> was the fastest of all the players.'
(No spatiotemporal location change but a need for situation-internal identifiability)

Data obtained for Korean

(15) **Template 1:** No spatiotemporal location change
  kyosil   an-ey   sonyen han-myeng-kwa sonye han-nyeng-i ancaissta.
  classroom inside-LOC boy   one-CLF-and   girl one-CLF   are.sitting
  'There is <u>a boy</u> and <u>a girl</u> sitting in the classroom…'
  sonye-ka **sonyen** yep-ey   ancaissta.   (Bare)
  girl-NOM   boy   side-LOC   are.sitting
  '<u>The girl</u> is sitting next to <u>the boy</u>.'

(16) **Template 2:** Spatiotemporal location change; object position
  kyosil   an-ey   sonyen han-myeng-kwa sonye han-nyeng-i ancaissta.
  classroom inside-LOC boy   one-CLF-and   girl one-CLF   are.sitting
  'There is <u>a boy</u> and <u>a girl</u> sitting in the classroom…'
  na-nun ecey   **#(ku)** **sonyen**-ul   mannassta.   (Dem)
  I-TOP   yesterday that boy-ACC   met
  'I met <u>the boy</u> yesterday.'

Data obtained for Korean cont'd

(17) **Template 3:** No spatiotemporal location change necessary

    ecey   kay han mali-lul   pwassta. **(?ku)**   **kay**-ka   kwiyewessta.   (Bare)

    yesterday dog one CL-ACC   saw   that   dog-NOM   was.cute

    'Yesterday I saw <u>a dog</u>. <u>The dog</u> was cute.'

(18) **Template 4:** Spatiotemporal location change; subject position

    ecey   kay han mali-lul pwassta. onul **#(ku)** **kay**-ka   cip-ey   wassta. (Dem)

    yesterday   dog one CLF-ACC saw   today that dog-NOM house-to came

    'Yesterday I saw <u>a dog</u>. Today <u>the/that dog</u> came to my house.'

Data obtained for Korean cont'd

(19) **Template 5:** No spatiotemporal location change but a need for situation-internal identifiability

ecey thipi chwukkwu cwungkey-eyse Son Heung-Min-ul pwassta.

yesterday television soccer broadcasting-LOC Heung-Min Son-ACC saw

'Yesterday, I saw Heung-Min Son in a soccer game on TV.'

#(ku) **senswu**-ka mom-nollim-i kacang ppallassta. (Dem)

that player-NOM body-movement-NOM most was.quick

'That player (i.e., Heung-Min Son) was the fastest of all the players.'

The findings

- The five languages behaved *alike,* but there were a few unexpected *differences*.

	Korean (n=2)	Mandarin (n=2)	Japanese (n=3)	Slovenian (n=2)	Turkish (n=3)
Template 1	Bare	Bare	Bare/Dem	Bare	Bare
Template 2	Dem	Dem	Dem	Dem	Dem
Template 3	Bare	Bare	Dem	Bare/Dem	Bare
Template 4	Dem	Dem	Dem	Dem	Dem
Template 5	Dem	Dem	Dem	Dem	Dem (n=2) Bare/Dem (n=1)

What to make of the findings?

- In bare noun languages, spatiotemporal shift between sentences instantiating discourse anaphora ordinarily requires using a Dem-NP.

- If there is no spatiotemporal shift, then a Dem is "optional" but languages or speakers may vary as to how they use a Dem:
  - My Japanese speakers seemed to use a Dem (along with the "nominative" marker *ga*) to focus a *discourse-salient* entity.
  - My Slovenian speakers seemed to use a Dem to mark *emotivity*.

- As to the behavior of one of my Turkish speakers regarding Template 5, additional discussion suggests that they were using a bare noun as a *quasi-name/epithet*.

Towards a formal semantic analysis

- Based on her analysis of Shan (Tai-Kadai, Myanmar), Moroney (2021) posits that bare noun languages have a *null anaphoric type shifter ι^x*, which has the semantics of the strong definite article proposed by Jenks (2018), given in (20b).

- Dayal and Jiang (2023) make essentially the same claim, as shown in (21b).

(20) Jenks (2018: (22))
    a. **Unique definite article**: $[\![\iota]\!] = \lambda s_r \lambda P: \exists !x[P(x)(s_r)].\iota x[P(x)(s_r)]$
    b. **Anaphoric definite article**: $[\![\iota^x]\!] = \lambda s_r \lambda P \lambda Q: \exists !x[P(x)(s_r) \,\&\, Q(x)].\iota x[P(x)(s_r)]$

(21) Dayal and Jiang (2023: (19))
    a. $[\![Dem]\!] = \lambda s \lambda P: \exists s'\ s \leq s'\ |P_{s'}| > 1.\iota x[P_s(x) \,\&\, x = y]$
    b. $[\![the_{strong}]\!]/[\![\iota]\!] = \lambda s \lambda P: |P_s \cap \lambda x[x = y]| = 1.\iota x[P_s(x) \,\&\, x = y]$

My proposal

- Bare noun languages have *two* different null anaphoric type shifters, one for *situation-internally licensed* definites, $\iota^x{}_S$, and one for *quasi-names* licensed at the *text* level, $\iota^x{}_T$.

- Dems encode a locative relation (*LR*) between an individual *x* and another individual *y*.

- Below, s_s stands for 'source situation', the situation where the immediate antecedent of an anaphoric definite is found; $\propto$ stands for 'spatiotemporal overlap'; *T* stands for 'text'.

(22) a. $[\![\iota^x{}_S]\!]^g = \lambda P: \exists!x[P(x)(s_s) \ \& \ s_s \propto s_r \ \& \ \exists y[y \neq x \ \& \ Q(y)(s_s) \ \& \ Q \neq P]].$
$\qquad\qquad \iota x[P(x)(s_r) \ \& \ \exists y[y \neq x \ \& \ Q(y)(s_r) \ \& \ Q \neq P] \ \& \ x = g(1)]$

   b. $[\![\iota^x{}_T]\!]^g = \lambda P: \exists!x[\text{Known-as-}P(x)(T)].\iota x[\text{Known-as-}P(x)(T) \ \& \ x = g(1)]$

   c. $[\![Dem]\!]^g = \lambda P: \exists!x[P(x)(s_r) \ \& \ LR(x)(y)(s_s)].\iota x[P(x)(s_r) \ \& \ LR(x)(y)(s_r) \ \& \ x = g(1)]$

My proposal cont'd

- The present analysis lets us explain data like (23), improving previous analyses (e.g., Ahn 2019; Moroney 2021; Dayal and Jiang 2023):
  - In (23), the anaphoric definite in (b) meets the uniqueness presupposition of the strong determiner assumed in the literature, but it cannot be a bare noun because there is no y different from x such that y has property Q that is different from P in s_r.
  - It can be a Dem-NP because it satisfies the presupposition of (22c).

(23) a. Tongmwul-pyengwon-eyes   kay han   mali-lul   po-ass-ta.
        animal-hospital-LOC   dog one   CLF-ACC   see-PST-DECL
        'I saw a dog at a/the veterinary hospital.'
    b. #(**Ku**)   kay-ka   kacang   kwiyewu-ess-ta.
        that   dog-NOM   most   be.cute-PST-DECL
        Intended: 'The dog was the cutest of all (the animals there).'

My proposal cont'd

- The present analysis also explains the contrast between (23) and (24) :
  - In (24b), a bare noun is used because here, the situation-internal identifiability presupposition of (22a) is satisfied.
  - Using a Dem-NP is judged infelicitous because here, a comparison is being made among the three animals that have been introduced into the discourse *in s_r*.

(24) a. Tongmwul-pyengwon-eyes   **kay**   **han**   **mali-wa**,   **koyangi** **han mali**,
       animal-hospital-LOC   dog   one   CLF-and   cat   one   CLF
       kuliko   **haymsutha**   **han**   **mali-lul**   po-ass-ta.
       and   hamster   one   CLF-ACC   see-PST-DECL
       'I saw a dog, a cat, and a hamster at a/the veterinary hospital.'
   b. Ku   cwung-eysey   (#**ku**)   kay-ka   kacang   kwiyewu-ess-ta.
       that among-LOC   that   dog-NOM   most   be.cute-PST-DECL
       Intended: 'The dog was the cutest among the three.'

Conclusion

- The goal of this talk was to provide an answer to when and why a Dem may be needed for marking anaphoric definites in bare noun languages like Korean.

- Along the way, I also mapped out when bare anaphoric definites may be licensed.

- **Central claims:**
  - There are two types of anaphoric bare nouns and one of them behaves like quasi-names.
  - Anaphoric Dems that are necessary occur as "the last resort" to locate a referent (compare Sybesma and Sio 2008; Ahn 2019; Dayal and Jiang 2023).
  - Dems encode *LR*s between two individuals but other discourse-pragmatic effects (e.g., anti-uniqueness effects) may come about due to the interplay between their semantics and other information-structural/cognitive factors (cf. Kim 2019, in revision).

Implications

- Anaphoric bare nouns are licensed by the salient property *P* their referents have in contrast to another individual *y* with a different property *Q*.

- Dem-NPs are licensed by the salient *LR* they have in contrast to another individual *y* with the same *P* but with a different *LR*.

- In both cases, *contrast* plays a role but in different ways (compare Jiang 2012; Robinson 2005; Dayal and Jiang 2023).

- Anaphoric definites with ι^x_S are comparable to immediate situation uses of definites.

- Anaphor definites with ι^x_T are comparable to larger situation uses of definites.

- This may be the reason why they have the same form as truly unique/weak definites, i.e., bare noun.

- Bare nouns and Dem-NPs are not really "in competition" (compare Ahn 2019).

Acknowledgements

- Many thanks to the following linguists for their help with judgments and discussion:

  - C.-T. James Huang
  - Feng-hsi Liu
  - The late Chisato Kitagawa
  - Shigeto Kawahara
  - Asako Higurashi
  - EunHee Lee
  - Lanko Marušič
  - Ebru Türker
  - Deniz Özyildiz
  - Jaklin Kornfilt

Selected references

Ahn, Dorothy. 2019. THAT thesis: A competition mechanism for anaphoric expressions. Doctoral dissertation, Harvard University.

Dayal, Veneeta and Julie Li Jiang. 2023. The puzzle of anaphoric bare nouns in Mandarin: A counterpoint to *Index! Linguistic Inquiry* 54(1): 147-167.

Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49: 501-536.

Kim, Min-Joo. 2021. Anaphoric definiteness in Korean: Situation-internal/-external reference. Korean Linguistics Workshop. University at Buffalo.

Moroney, Mary. 2021. Updating the typology of definiteness: Evidence from bare nouns in Shan. *Glossa: a journal of general linguistics* 6(1): 56. DOI: https://doi.org/10.5334/gjgl.1221.

Schwarz, Florian. 2009. Two types of definites in natural language. Doctoral dissertation, University of Massachusetts, Amherst.

Simpson, Andrew and Zoe Wu. 2022. Constraints on the representation of anaphoric definiteness in Mandarin Chinese: A reassessment. In *New explorations in Chinese theoretical syntax: Studies in honor of Yen-Hui Audrey Li*, ed. Andrew Simpson, 301-330. Amsterdam: John Benjamins.

Sybesma, Rint and Joanna Ut-Seong Sio. 2008. D is for demonstrative. Investigating the position of the demonstrative in Chinese and Zhuang. *The Linguistic Review* 25: 453-478.

Thank you for listening!

min-joo.kim@ttu.edu