Definite Discourse-New Reference in L1 and L2: A Study of Bridging in Mandarin, Korean,

and English

Peter Robert Crosthwaite

University of Cambridge

**Abstract**

Definite expressions may be used to introduce a referent into discourse when their familiarity (Prince, 1981) between speaker and listener can be inferred, a strategy known as bridging (Clark, 1975; Clark & Haviland, 1977). However, bridging may be difficult to acquire compared to the acquisition of indefinite introductions for non-inferable referent types, with the native language of the second language (L2) learners suggested to be a major factor influencing L2 acquisition. Two experiments were conducted into the use of bridging, using a controlled picture sequence where neutral, weakly, or strongly inferable referents were to be introduced. The first study investigated native speaker data from English, Mandarin Chinese, and Korean, and the second study focused on a cross-section of learner data from Mandarin and Korean L2 English learners from four proficiency levels from A1 (Basic user) to C2 (Proficient user) of the Common European Framework (CEFR) (Council of Europe, 2001). The native results show that the three languages introduce inferable referents in markedly different ways: definite articles for English, omission of numeral + classifier construction before the noun for Mandarin, and making no

apparent distinction for inferability in terms of NP form for Korean. The L2 results show that the acquisition of the definite article + noun construction to introduce inferable referents in L2 English occurs at lower CEFR levels for the Mandarin group (A2) than the Korean group (B2). Therefore, it is claimed that Mandarin L2 English learners may have less difficulty mapping L2 syntactic form to pragmatic function in their L2 narrative production than their Korean counterparts, as a possible consequence of positive transfer from their native language.

**Keywords**: Bridging, Inferability, Narratives, Second Language Acquisition, Mandarin, Korean, L2 English, Crosslinguistic Influence

**Author Note**

Correspondence concerning this article should be addressed to Peter Robert Crosthwaite, University of Cambridge, Department of Theoretical and Applied Linguistics, 9 West Road, CB3 9DP Cambridge, UK. E-mail: drprc80@gmail.com.

## Introduction

For most second language (L2) English learners, the article system is considered to be one of the most difficult aspects of English grammar to acquire due to its complex multifunctional nature, and as such, the appropriate use of articles is generally not acquired until quite late in the learning process. In terms of complexity, Hawkins (1978) presents a variety of different applications of the definite article, including anaphoric, visible situation ('pass me the water'), immediate situation ('don't go in there, the dog may bite you'), associative relations ('the book, the author'), and inclusivity ('bring (all of?) the wickets in after a game of cricket'), each of which have a wide range of appropriateness conditions including 'existence', 'identifiability', 'set

membership', and 'set composition'. Huebner (1983) defines article use as dependent on whether the reference is specific reference [+/-SR] or known to the hearer [+/-HK] with his taxonomy of article functions, including generic, referential indefinite, referential definite, and non-referential contexts. Hence, the English article system may be considered "a complex set of abstract distinctions which are, to some extent, arbitrarily mapped onto surface forms" (Ekiert, 2007, p. 1), and learners looking for a simple one-form one-function mapping for new/given reference (e.g., Andersen, 1984) may find it difficult to do so with English articles.

This article reports on two studies regarding the use of discourse-new definite referring expressions, known as bridging descriptions (Clark, 1975; Clark & Haviland, 1977), that are used to introduce what Prince (1981) called inferable referents into narrative discourse. Two parallel experiments were performed collecting both native speaker and L2 learner oral narrative production data. The first study determined how inferable referents are introduced into discourse in English, Mandarin, and Korean, with each native language handling the introduction of such referents in markedly different fashions. The findings lead to certain hypotheses regarding the learnability of inferable referent introductions by Mandarin and Korean L2 English learners in terms of potential transfer effects. The second study presents data regarding the introduction of discourse-new inferable referents by the L2 English learners, sampled from across 4 proficiency levels on the Common European Framework (CEFR) (Council of Europe, 2001), from A1 (Basic user) to C2 (Proficient user).

## Review of the Literature

### Bridging Descriptions

In English and other languages with both definite and indefinite articles, indefinite referring expressions may be used when a new entity (or referent) needs to be introduced by the speaker into the discourse. The speaker's choice of an indefinite expression informs their listeners

that they need to make a new representation of a referent (that is unfamiliar to them) into the shared mental model of the discourse as it unfolds, or what Stalnaker (1974) called common ground. Definite referring expressions such as pronouns or noun phrases with (or without) definite articles are then used to maintain reference to these referents, marking the referent's transition from 'new' to 'given' information in the discourse (Slobin, 1975).

Speakers mark the introductions of referents and maintain subsequent reference to them through the selection of particular referring expressions that are suitable for these functions within the typology of referring expressions available within their language. For maintaining discourse-old reference, the *identifiability* (Gundel, Hedberg, & Zacharski, 1993) or *accessibility* (Ariel, 1991, 2008, 2010) of a referent within the common ground is claimed to influence the choice of referring expression assigned to a referent, under economy/relevance based constraints (e.g., Sperber & Wilson, 1995). Such constraints predict that referents which are highly activated in the discourse are referred to with shorter referring expressions (such as pronouns, or zero anaphora; i.e., where the referent is not explicitly mentioned by the speaker when its identity is inferable) so that production remains optimally economical and relevant. Lesser-activated referents are referred to with longer, more informative expressions (such as proper names or full definite descriptions, e.g., 'the boy') so that production remains suitably informative and unambiguous.

Referent introductions are also claimed by Ariel (1996) to be subject to the concept of accessibility, with indefinite noun phrases such as '*this*+N'[1] used for more accessible referents than quantified indefinites (*another*+N), which in turn are used for more accessible referents than those referred to with indefinite-only NPs ('*a*+N'). However, in many languages, when referring to a referent for the first time, a distinction may be made regarding the perceived *familiarity* of a

---

[1] When used for discourse introductions (e.g., '...and then *this guy* arrives...') rather than the deictic use (e.g., 'This house looks rather big' with the speaker pointing at the house in question)

referent in the common ground between speaker and listener, as shown in Prince's (1981) familiarity scale for referents: New→Inferable→Evoked. Using the least activated NP form (e.g., the indefinite article + noun in English) is therefore unnecessary under an economy/relevance based approach, if the existence of the referent can be inferred from the context. One way that languages have dealt with this issue is by allowing definite descriptions to be used for discourse-new referents when their existence may be presupposed by the audience. References of this type are known as *bridging descriptions*, a term first coined by Clark (1975) and Clark and Haviland (1977). These represent an alternative form-function mapping to the use of definite expressions for reference maintenance, and do not always have a co-referential antecedent in the discourse (Gundel, 1996).

Examples of inferable referents include those in 'part/whole' relationships with each other (e.g., 'I walked into the room, *the chandeliers* sparkled brightly', Clark, 1975, p. 171), or 'cause/consequence' relationships (e.g., 'An earthquake…*the suffering* people are going through' Vieira, 1998, p. 23). A full taxonomy of bridging relations is found in Gardent, Manuélian, and Kow (2003).

These kinds of expressions use pragmatic world or general knowledge shared between speakers and their audience for their resolution, in the form of a "conceptual address" which is "a point of access to logical, encyclopaedic, and linguistic information in memory" (Matsui, 2000, p. 31) stored in long term memory as a single chunk or unit. For example, Sanford and Garrod (1981; cited in Hellman, 1996) found that reference to scenario-dependent definite NPs [such as 'the waiter' or 'the chef' in the context of a restaurant] "takes no longer to process [than discourse-old referents] in spite of the absence of explicit antecedents" (Hellman, 1996, p. 201). In another example, Wilson and Matsui (1998, p. 6) tested for acceptability the three passages given in (1) through (3) below, using a 7-point Likert scale questionnaire (7= the most acceptable for bridging):

5

(1) I was looking at van Gogh's self portrait. *The missing ear* made me feel sad.

(2) I went to see some impressionist paintings. *The missing ear* made me feel sad.

(3) I went to a gallery yesterday. *The missing ear* made me feel sad.

The weaker the familiarity of the target of the bridging description to the context (e.g., example 3), the less plausible the bridging inference becomes, leading to more inference being needed to make the bridge with an associated relevance cost for the audience.

To summarize, bridging descriptions come about through the speaker's need to be economical with creating new representations of entities through indefinite references, when, for inferable referents, a definite reference will suffice for optimal relevance if the context allows for bridging to occur.

**Bridging and L2 Acquisition**

In the course of L2 English acquisition, native speakers of article-less languages such as Mandarin and Korean have to acquire the definite article to mark inferable referent introductions appropriately in the L2. There is no one-to-one mapping of form and function with the English article system, in that even though the indefinite article is more commonly used for referent introductions, the definite article may be used when bridging introductions are possible. Moreover, as both indefinite and definite articles have a complex variety of associated functions, selecting the right article (or producing an article at all) when introducing a referent is likely to present difficulties for L2 English learners.

However, research has shown that L2 learners from particular source language backgrounds may find the acquisition of articles easier or more difficult in terms of transfer. Transfer, in this study, follows the definition of Ellis (1994) that includes "errors (negative transfer), facilitation (positive transfer), avoidance of target language norms, and their over-use"

(Ellis, 1994, p. 341). In their Complex Adaptive System Principles model, Filipovic and Hawkins (2013) have proposed that, in order to minimize processing effort and to maximize their communicative efficiency, L2 learners may maximize positive transfer of principles from their first language (L1) into their L2 production. These authors refer to this prediction as the Maximize Positive Transfer principle. Learners may also follow a Permit Negative Transfer principle from the L1 to the L2 in order to maximize their expressive power, as long as such transfer does not impede communication in the L2, and this last prediction is expressed as the Communicative Blocking of Negative Transfer principle.

Examples of positive transfer are seen in Nakahama's (2009) comparison of Korean and English L2 Japanese oral narratives, where the L2 requires zero anaphora for reference maintenance. She found that Koreans produced more zero anaphora in their Japanese narratives than their English counterparts did, and this held at each proficiency level tested. As the Japanese distribution of zero anaphora for reference maintenance is similar to that of Korean, the congruence between the native and target referential systems gave the Korean learners an advantage over their English counterparts. In their study of the L2 acquisition of English plural and possessive morphemes, Luk and Shirai (2009) found that native speakers of Japanese, Korean, and Mandarin Chinese acquired plural –s later than predicted by an established natural order of L2 morpheme acquisition (Krashen, 1977), with the possessive –'s morpheme acquired earlier than predicted, (see also Sasaki, 1987). They argued this pattern is dependent "on the presence or absence of the equivalent category in their L1" (Luk & Shirai, 2009, p. 721).

Ample evidence for both positive and negative effects has also been reported for article acquisition, specifically. Jarvis (2002) asked Finnish and Swedish speakers of English as a second language to mark newness with articles. The Finns, who do not mark newness in their native language, did not mark newness in their L2 English production, while the Swedes, who do mark newness in their native language, did so in their L2 production. Sleeman (2004) found positive

transfer of the French article system for Dutch L2 French learners (both of whom have an article system), while Japanese L2 French learners (who do not have an article system in their native language) struggled to acquire L2 French articles. L2 learners from article-less languages have been shown to have trouble assigning features of definiteness and/or specificity with the L2 article system, to overuse indefinite articles where definite articles are expected (Leung, 2001), to overuse definite articles where indefinite articles are expected (Brown, 1973; Huebner, 1983; Master, 1987; Young, 1996), to overuse articles in both definite and indefinite contexts (Ionin, Ko & Wexler, 2004) and to negatively transfer other lexical items from their L1, such as demonstratives or numerals, for the same function as English articles (Robertson, 2000). Another example of negative transfer was found for Korean learners of English at intermediate levels, who were not able to distinguish between the function of definite articles and demonstratives, using them analogously when referring to non-unique referents where demonstratives were expected, and unique referents where definite articles were expected (Ionin, Baek, Kim, Ko, & Wexler, 2012). In another example, Ekiert (2007, 2010) found that Polish L2 English learners' usage of articles was determined, in part, by the Polish system of encoding definiteness. She claimed that L2 acquisition "not only appears to be hindered by a limited, L1-motivated set of options (i.e., forms) for the grammatical encodings of characteristics of objects and events, but also by the lack of an equivalent conceptual system" (2010, p. 150).

To summarize, the relative learnability of the L2 article system needed for referent introductions (including bridging) may be affected by the configuration of the learners' L1. Thus, if it could be shown that a particular native language is approaching an article-like system through other NP forms in a similar manner to English, then L2 learners of English from that native language may be able to maximize positive transfer when acquiring the definite article for bridging, and be at an advantage over a learner from a different native language that does not have an equivalent system. While bridging descriptions are well-documented for English (Clark, 1975;

Clark & Haviland, 1977; Gundel, 1996; Poesio & Viera, 1998; Viera & Poesio, 2000), there is less research on the introduction of inferable referents in Mandarin,[2] and at present the author is unaware of any research on this phenomenon in Korean. With this in mind, the next section provides some background information on how Mandarin and Korean might introduce inferable discourse-new referents into discourse, and whether speakers of either language may have an advantage when acquiring English articles in the L2.

**Language-Specific Strategies for Discourse-New Referent Introductions**

The choice of source languages for this study was motivated by Huang (2000), who argues that, unlike English, Mandarin and Korean are more pragmatic than syntactic languages. Under this definition, two differences with English are important. First, Mandarin and Korean have a far higher rate of zero anaphora for discourse-old reference than English[3]; in the latter language, the resolution of discourse-old reference has to be inferred. In Chinese, specifically, style topics are linked to their associated comment pragmatically (Li & Thompson, 1976) rather than bound syntactically. Second, Mandarin and Korean make little use of morphology to signal the (in)definiteness of referents, at least when compared to the English article system. On the other hand, there are a number of differences between how Mandarin and Korean introduce referents into discourse, and these differences appear to suggest that some pragmatic languages may be more or less pragmatic than others, at least when it comes to the marking of inferable referents. In what follows, I describe the introduction of referents into discourse in Mandarin and Korean in more detail.

Mandarin employs a word order preference for referent introductions, with a general preference for post-verbal referent introductions (Hickmann, Hendriks, Roland, & Liang, 1996),

---

[2] At least in the English literature; see Aiping (2003) for a Chinese language perspective on bridging in Mandarin.
[3] Although zero anaphora in English are restricted to 'diary drop' omission of subjects, while they are used for a wider variety of functions (e.g., objects) in Mandarin and Korean.

although the occasional use of pre-verbal referent introductions appears to be influenced by character status (e.g., main vs. supporting characters, Crosthwaite, 2014). Discourse-new introductions may also be locally marked with a numeral + classifier before the noun, as in illustration (4) [4]:

(4) 有　　　一　　个　　男孩　　走进　餐厅

　　yǒu　　**yī　gè　nánhái** zǒujìn cāntīng

　　There is **one　CLS boy**　　enter　restaurant
　　'A boy went to a restaurant'

Classifiers come in many forms depending on the referent to be introduced, but the most common classifiers for human referents are the so-called general or individual classifiers (个 – gè), or the honorable classifier (位 – wèi), where the speaker wishes to be polite about the referent. The use of the numeral + classifier construction before the noun is optional in both discourse-new and discourse-old contexts, yet the use of these constructions is preferred for discourse-new reference (Hickmann et. al., 1996), and they are used at high frequencies for this purpose (Crosthwaite, 2014). The use of constructions such as the numeral + classifier (as well as proximal and distal demonstratives) in Mandarin is claimed by Li and Thompson (1976, 1989), Hedberg (1996), Chen (2004), and Liu (2010) to be shifting diachronically toward an English-like article system, sharing a number of similar form/function relationships. For example, Hedberg (1996) provides the example in (5) as indicative of a distal demonstrative in Mandarin sharing the same function as the English definite article, as it can be used when the dog in question is unfamiliar to the listener:

---

[4] CLS = classifier, ATT = attributive particle, RES = result particle.

(5)　昨天　　晚上　　我　睡　不　着

　　　Zuotian　wanshang wo　shui　bu　zhao

　　　yesterday　evening　I　sleep not achieve

　　　隔壁　　　的　**那**　条　狗　　叫　得　　厉害

　　　Gebi　　　de　**nei** tiao gou jiao de　　lihai

　　　next.door ATT **that** CLS dog bark RES terribly

　　　'I couldn't sleep last night. The dog next door was barking.'

Liu (2010) claims that numeral + classifier + noun introductions are used primarily to introduce new but unfamiliar referents into the discourse. When a discourse-new referent is not marked with a numeral + classifier, such a referent may be considered as more familiar than one with an accompanying numeral + classifier, yet it may still be possible for the addressee to identify the unfamiliar referent as new "if the information encoded by the nominal is relevant and sufficient for identification in the context [...] on the basis of the nominal alone" (Hedberg, 1996, p. 184), and particularly if the referent is introduced pre-verbally (although not obligatorily). Thus, it is expected in this research that the unmarked bare nominal form should occur more often when a representation of the referent is inferable (as with definite article bridging descriptions in English), as shown in (6), where the waiter is inferable from the restaurant scenario:

(6) 有　　一　个　**男孩**　走进　餐厅.　　　**服务员**　给　他　一　份　　菜单.

　　 yǒu　**yī gè nánhái** zǒujìn cāntīng.　**fúwùyuán** gěi　tā　yī fèn　càidān.

　　 There is one CLS boy　enter　restaurant.　**Waiter**　give him one CLS menu

　　 'A boy walked into a restaurant. The waiter handed him a menu'

In summary, in Mandarin, discourse-new referents may be introduced pre- or postverbally and are usually accompanied by a numeral + classifier construction before the noun. However, for

inferable referents, it appears that the numeral + classifier construction is generally omitted before the noun.

In Korean, new referents may be introduced with numeral + noun or noun + numeral + classifier constructions, as shown in (7) [5]:

(7) 한 아이가    식당으로    들어갔습니다.

   **han ai -ga**    sikdangeuro    dureogassseumnida.

   **one boy**-SUB   restaurant-into   enter-PAST-POL-DEC

  'A boy walked into a restaurant'

While Korean does use numeral constructions (with or without classifiers) to introduce referents into the discourse, Koreans are more likely to use bare nominals to do so (Brown, 2011; Crosthwaite, 2014), as shown in (8):

(8) 선생님이       식당으로    들어갔습니다.

   **seonsaeng-nim**-i    sikdangeuro   deureogassseumnida.

   **teacher**-HON-SUB    restaurant-into  enter-PAST-POL-DEC

  '(A) teacher walked into a restaurant'

As the lowest accessibility NP form available to Korean should be the numeral + noun construction, the preference for a bare nominal introduction in example (8) is unusual in that this preference does not appear to follow the usual pattern of NP selection in terms of referent

---

[5] SUB = subject marker, OBJ = object marker, PAST = simple past, POL = polite verb ending, DEC = declarative marker, HON = honorific marker

accessibility (i.e., the lowest accessibility referent does not generally receive the lowest accessibility NP form in Korean). Crosthwaite (forthcoming) instead found that numeral introductions in Korean are generally reserved for the first-mentioned character in a narrative, and only if that character is not of a particular social status (such as a teacher). Word order cannot be used to establish discourse-newness in Korean either, as scrambling of subjects and objects is permitted (although rare, see Kwon, Polinsky, & Kluender, 2006). In addition, Crosthwaite (forthcoming) has shown that, for Koreans, first position in the utterance is preferred for both referent introductions and reference maintenance. Thus, in Korean, the inferability of a referent may not be realized in the NP form used to refer, given that a bare nominal is just as likely to be used when the referent cannot be inferred, as in example (8), than when the referent can be inferred, as in example (9):

(9) *Inferable*  한 아이가   식당으로    들어갔습니다.

Han ai-ga   sikdangeuro   deureogassseumnida.

One boy-SUB  restaurant-into   enter-PAST-POL-DEC.

**종업원이** 메뉴판을   주었습니다.

**Jongeobwoni** menyupaneul jueossseumnida.

**Waiter**-SUB   menu-OBJ   give-PAST-POL-DEC.

'A boy walked into a restaurant.  (The) waiter handed (him) a menu'

## Research Questions

As is clear from the review of the literature just offered, in order to generate hypotheses for the L2 acquisition of bridging descriptions in English, and assuming that the L1 will have some positive or negative transfer effects, more research is needed into how Mandarin and Korean introduce inferable referents. Therefore, two different studies were carried out, one with L1 data and one with L2 data.

For the L1 study, the following research question was posed:

Research Question 1: How do speakers of Mandarin, Korean, and English introduce inferable referents into discourse?

Of the two non-English native languages, Mandarin at least appears to make a distinction for inferability in the form of referring expression used for discourse-new mentions, while Korean does not appear to make any such distinction. Therefore, it is conceivable that Mandarin L2 English learners may have an advantage over their Korean counterparts when acquiring bridging descriptions in English. The advantage should come from two sources of positive transfer. Firstly, both Mandarin and English generally mark referent introductions using overt syntactic means (articles in English, numeral + classifier in Mandarin), while in Korean bare nominals are preferred for referent introductions in the majority of cases. Secondly, both Mandarin and English appear to make a distinction for inferability for referent introductions (definite articles in English, omission of numeral + classifier in Mandarin), while no such distinction appears to be made in Korean, where bare nominals for inferable and non-inferable referents alike. In light of these predicted transfer effects and the likelihood that L2 proficiency might moderate such effects, the following two research questions were posed for the L2 study:

Research Question 2: At which CEFR proficiency level will L1 Mandarin and Korean speakers learning L2 English acquire the use of 'bridging' definite article + noun constructions to introduce inferable referents into a discourse text?

Research Question 3: Will the acquisition of bridging in L2 English be facilitated by positive transfer in Mandarin L1 speakers as opposed to Korean L1 speakers?

**Study 1: The Introduction of Inferable Referents in English, Mandarin, and Korean**

The production of discourse-new referring expressions for inferable and non-inferable referents was investigated in English, Mandarin Chinese, and Korean, by collecting data from adult native speakers. These results from Study 1 allowed us to determine subsequently in Study 2 whether there is indeed transfer from L1 Mandarin to L2 English, and whether there is none from L1 Korean.

For Mandarin, it was predicted that neutral and non-inferable referents would be introduced through numeral + classifier NPs, which are commonly used to mark discourse newness, while bare nominal NPs would be used for inferable referents (regardless of word order). For Korean, given the reported preference for bare nominals for most referent types (Brown, 2011; Crosthwaite, forthcoming), no distinction for inferability is likely to be made in the introductions of both inferable and non-inferable referents, with either type of referent just as likely to receive a bare nominal introduction.

**Method**

*Participants*

Sixty adult native speakers of English, Mandarin and Korean ($n=20$ per language) participated in the native language study. Table 1 describes the native-language participants.

*Background of Native-Language Participants, Study 1*

| | Total *N* | Male/female | Age ranges | | | |
|---|---|---|---|---|---|---|
| | | | 16-20 yrs | 20-30 yrs | 30-40 yrs | 40-50 yrs |
| L1 English | 20 | 10/10 | 1 | 14 | 2 | 3 |
| L1 Mandarin | 20 | 7/13 | 5 | 11 | 2 | 2 |
| L1 Korean | 20 | 9/11 | 2 | 12 | 5 | 1 |

All of the English- and Mandarin-speaking participants and 8 of the Korean-speaking participants were recruited in Cambridge, UK. Twelve Korean participants were recruited by the researcher in South Korea when not enough Korean-speaking candidates could be found in Cambridge. The native English speakers were either English teachers working at language schools in Cambridge or students of Cambridge University. The native Mandarin and Korean speakers were either students at language schools in Cambridge or students of Cambridge University, with the exception of the twelve participants recruited in Korea, who were students at a private university. While the English level of the Mandarin and Korean participants in the native study was not tested, an effort was made to include participants who appeared to have a high level of English in the L2 study. As a result, the English ability of the native Mandarin and Korean group in Study 1 could be considered as relatively low (e.g., A2, possibly B1 level at the highest), and so it is unlikely that the participants' English ability would influence their L1 production, especially given that the listener of the narrative would be a native speaker of the L1 in question.

## Materials

The picture sequence, used in both Study 1 and 2, was designed to capture a variety of discourse-new and discourse-old phenomena, including the distribution and functions of zero anaphora, long distance pronominalization, switch-role reference, switch-reference, and other linguistic features of narrative discourse including the existence of inferable and non-inferable referents (Crosthwaite, 2013, forthcoming). A numbered version of the picture sequence can be found in the Supporting Information online (the materials used in the experiment were not numbered). The story grammar used in this experiment was normalized for narrative and episodic structure using Stein and Glenn's (1979) story grammar model. Within this structure, a main character tries to perform an action (such as flying a model aircraft, or playing with a ball), with the story focusing on how the character manages to perform that action in the face of some difficulty to overcome, and finishing with the character finally able to perform the action.

As shown in the Supporting Information online, school setting is introduced in Picture 1, and is seen again in Pictures 2, 3, and 5.  By repeating the setting in the background in this way, it should be "retained in peripheral consciousness as background orientation for the particular, localized events which may then be focused on" (Chafe, 1980, p. 42).  Given that the school scenario has been established, this allowed for the referents to be selected as either neutral, non-inferable or inferable, as listed in Table 2.

*Description of Target Referents to Be Introduced*

| Neutral | Non-inferable | Inferable |
|---|---|---|
| BOY (Picture 1) | SOLDIER (Picture 15) | PLAYGROUND/COURT (Picture 2 - English only) |
| BALL (Picture 1) | CANNON (Picture 16) | HOOP/NET (Picture 5) |
| GIRL (Picture 7) | DOG (Picture 20) | BACKBOARD (Picture 11) |
|  |  | TEACHER (Picture 12) |

*Note.* The pictures are shown in the Supporting Information online.

The *neutral* referents were defined as not scenario-dependent (Sanford & Garrod, 1981), not in a part/whole relationship with an antecedent (e.g., Clark, 1975), nor in a cause/consequence relationship (e.g., Vieria, 1998), but were also not considered entirely implausible given the school scenario.  The *non-inferable* referents were considered unexpected and implausible given the scenario.   For the *inferable* referents, the PLAYGROUND/COURT and TEACHER were considered scenario-dependent to the school setting, while the HOOP/NET and BACKBOARD (or variations, such as BOARD, FRAME) were considered to be in a part/whole relationship to the basketball court or inferable as part of the equipment needed for a basketball game.  The PLAYGROUND/COURT was only used as a target referent in English as locatives are treated

differently in Mandarin or Korean[6] and the school itself must have been introduced explicitly beforehand for reference to the PLAYGROUND to be counted. It was predicted that both neutral or non-inferable referents should be referred to with the indefinite article + noun in English, numeral (+/-classifier) + noun in Mandarin and bare nominals in Korean, while inferable referents should be referred to with a bridging definite article + noun introduction in English, and with bare nominals in both Mandarin and Korean. The second boy seen in Picture 3 was not included in the scoring, as this referent may be introduced differently from the others due to what Arnold and Griffin (2007) call a similarity-based interference effect (- e.g., '...and then *his friend* arrived', '...and *another boy* showed up').

*Procedure*

The participants were met by the researcher individually in a quiet room. Three identical copies of the picture sequence were placed in three separate envelopes marked A, B, and C on a desk in the room. A group of naïve listeners were recruited who were informed about the purpose of the study and were native speakers of their respective source languages (English, *n*=3, Mandarin, *n*=4, or Korean, *n*=3)). The experimenter (the author of this paper for the English group and a native-speaking assistant for the Korean and Mandarin groups) and a single naïve listener of the native language from the groups described above were present in the room, and the participant was asked to enter the room once everything was ready. The experimenter (either the author or a native-speaking assistant) read out the instructions to the participants in their native language.

The instructions stated that the participant was to choose from one of the envelopes marked A, B or C while the naïve listener looked away, and to take out the picture sequence

---

[6] Locatives in Mandarin and Korean are assigned postnominal locative markers (e.g., 'qiang-shang' - 'wall-upon' regardless of inferability, rather than pre-nominal numerals / numeral + classifiers, or bare nominal NPs.

inside, being careful to hold the picture sequence toward them so that the listener could not see it once they turned their back around. The experimenter then removed all of the envelopes from the table and told the listener to face the participant so that the listener was unable to see which envelope had been removed. The participant was then told that the picture sequence they had chosen was different from the ones hidden in the other envelopes and that they should use the picture sequence they had chosen to tell a story to the naïve listener. The participants were told that the listener had not yet seen any of the picture sequences in the envelopes at the time, and that the listener would attempt to guess from the participant's story which envelope from the three the participant had picked, at a later time when the listener would compare the three pictures together with the recorded narratives. This was done to ensure no mutual knowledge of the sequence between the naïve listener and the participant could be assumed by the participant before telling the narrative, so as to avoid deictic reference from the participants.

The participant was told that they could take as much time as they needed before they started to tell the story, so that they had time to organize the narrative in a series of sequential episodes (or macrostructures, following van Dijk, 1976) to achieve global coherence. To emphasize this point, they were told that they should try to make a story with a beginning, middle and end, rather than simply read off what they saw in the picture sequence. They were also told to make the story as interesting as possible so that the listener would enjoy it. On average, participants in each native group took 2-3 minutes to observe the sequence before they started telling the story, suggesting the participants required a pre-planning stage where the macrostructures of the actions performed by the referents were being worked out.

The participants were told that they could take as long as they liked to tell the story once they had begun, so they did not feel any pressure of time constraints. They were told that they could continue to look at the pictures as they told their story, so that they were not under pressure to memorize any part of the sequence. When the participant signaled that they were ready to

begin, the tape recorder was started and the experimenter left the room, leaving only the participant and the naïve listener.

During the telling of the story, the naïve listener actively listened to the story, offering both linguistic (e.g., 'hmm', 'OK') and paralinguistic signals (e.g., nods, smiles) to reassure the speaker that their story was being listened to. Once the narrative was complete, the recorder was stopped and the participant was thanked and offered a gift voucher worth £5.

*Scoring*

The target referents were divided between the neutral, non-inferable, or inferable referent types outlined in Table 2. Only the first mention of a referent was considered in the scoring. Whenever one of the target referents was explicitly introduced by the participant, the NP form used to introduce that referent was noted, and the proportion of particular NP forms used by a language group for a particular referent type was calculated. All scoring was performed by the researcher after the narratives were transcribed by a native-speaking assistant and translated by the researcher. References where the narrator had not identified the character correctly from the picture sequence (for example, introducing the teacher as 'an old man') or had used a proper name ('Jack') were not included in the scoring. Occasionally the soldier (non-inferable, Picture 15) was referred to as a 'police officer', perhaps from difficulty in ascertaining the identity of the soldier from the picture. However, this interpretation was also decided to be non-inferable in a school context, so references of this kind were still included in the final scoring for the non-inferable referent type.

**Results**

*English*

20

Table 3 describes the NP forms used to introduce each referent type into the discourse, after correcting for forms that did not match the referent type (e.g., 'an old man', 'Jack').

*NP forms Used to Introduce Each Referent Type in the Native English Data*

| Referent type | NP form | Frequency | Percentage |
|---|---|---|---|
| Neutral (*k*=42) | Definite article + noun (e.g., 'the boy') | 3 | 7.1 |
| | Indefinite article + noun (e.g., 'a boy') | 31 | 73.8 |
| | Possessive pronoun + noun (e.g., 'his ball') | 6 | 14.3 |
| | Other (numeral+noun 'one ball', demonstrative+noun 'this girl') | 2 | 4.8 |
| Non-inferable (*k*=54) | Definite article + noun (e.g., 'the soldier') | 4 | 7.4 |
| | Indefinite article + noun (e.g., 'a dog') | 45 | 83.3 |
| | Possessive pronoun + noun (e.g., 'his cannon') | 3 | 5.6 |
| | Other (numeral+noun 'one dog', demonstrative+noun 'this dog') | 2 | 3.7 |
| Inferable (*k*=45) | Definite article + noun (e.g., 'the teacher', 'the playground') | 39 | 86.7 |
| | Indefinite article + noun (e.g., 'a teacher') | 5 | 11.1 |
| | Possessive pronoun + noun (e.g., 'his teacher') | 1 | 2.2 |

*Note. k* is total number of introductions per referent type

It can be seen in Table 3 that the neutral or non-inferable referents were introduced mainly through indefinite article + noun NPs (73.8% and 83.3% respectively), while the inferable referents were introduced with the definite article in 86.7% of cases[7].These patterns are illustrated in (10) through (12):

(10) *Neutral* '**A little boy** arrived at the house (.) spinning **a basketball** on his finger'

(11) *Non-inferable* 'Along came **a soldier** with **a cannon** to shoot the ball down'

(12) *Inferable* 'So the children go to find **the teacher**' / 'he shot the ball to **the hoop**'

---

[7] A reviewer of an earlier version of this article suggested that the English speakers may be using the definite article to refer to the uniqueness of these referents, rather than their familiarity for bridging. However, I agree with another anonymous reviewer in that it is the familiarity of the referent that allows for the hearer to infer the uniqueness of that referent felicitously when the definite article is used, i.e., the hearer cannot accommodate the existence of a unique *soldier* on a playground, since soldiers aren't normally a familiar fixture at playgrounds.

As discourse-new data is of a binomial distribution, a binary logistic regression was performed in SPSS to see whether the inferability of the referent was a predictor of definite article + noun bridging first mentions. Both the non-inferable and neutral referents were significant negative predictors of definite article + noun mentions ($\beta = -2.551$, Wald = 18.122, exp $\beta$ odds ratio = .078, $p <0.001$ for the neutral referents, $\beta = -2.826$, Wald = 22.614, exp $\beta$ odds ratio = .059, $p <0.001$ for the non-inferable referents), while the inferable referent type was a significant positive predictor of referent introductions with this NP form ($\beta = 2.046$, Wald = 18.807, exp $\beta$ odds ratio = 7.740, $p <0.001$). The predictors as a set reliably predicted bridging use (*chi* 116.945, $p <0.001$, *df*=3) and a Nagelkerke's $R^2$ of .758 indicated a strong relationship between prediction and grouping. In comparison, a subsequent logistic regression regarding referent type and indefinite article + noun introductions showed that the non-inferable referent was a significant positive predictor of introductions with that NP form ($\beta = 1.784$, Wald = 21.453, exp $\beta$ odds ratio = 5.952, $p <0.001$), as was the neutrally inferable referent ($\beta = 0.927$, Wald = 7.342, exp $\beta$ odds ratio = 2.527, $p <0.05$). Therefore, the results follow the literature on bridging in English: indefinite is preferred for neutral and non-inferable referents, definite for inferable referents .

***Mandarin***

Table 4 describes the NP forms used to introduce each referent type into the discourse

*NP forms Used to Introduce Each Referent Type in the Native Mandarin Data (n=20)*

| Referent type | NP form | Frequency | Percentage |
|---|---|---|---|
| Neutral (*k*=38) | Numeral + classifier + noun (e.g., '一个男孩' - 'one CLA boy') | 27 | 71.1 |
| | Bare nominal (e.g., '篮球' - 'basketball') | 9 | 23.7 |
| | Demonstrative+ classifier + noun ('这个球' - 'this CLA ball') | 1 | 2.6 |
| | Possessive pronoun + noun ('他的球' - 'he ATT[1] ball') | 1 | 2.6 |
| Non-inferable (*k*=58) | Numeral + classifier + noun (e.g., '一个士兵' - 'one CLA soldier') | 53 | 91.3 |
| | Bare nominal (e.g., '士兵' - 'soldier') | 5 | 8.7 |
| Inferable (*k*=40) | Numeral + classifier + noun (e.g., '一个老师' - 'one CLA teacher') | 1 | 2.5 |
| | Bare nominal (e.g., '老师' - 'teacher') | 36 | 90 |
| | Demonstrative + classifier + noun ('那个铁篮筐' - 'that CLA iron basket') | 3 | 7.5 |

*Note. k* is total number of introductions per referent type; *n* is total number of Mandarin L1 participants. ATT = attributive marker (的 - 'de'), ASP = aspect marker (了 - 'le').

It can be seen in Table 4 that the neutral or non-inferable referents were introduced with numeral + classifier + noun NPs in most cases, while the inferable referents were introduced with bare nominals in 90% of cases. These patterns are illustrated in (13) through (15)[8]:

(13) *Neutral*　　　　　一 个　小朋友　　去 学校

yī　gè　xiǎopéngyǒu qù xuéxiào.

one CLA little child　go school

'a little child goes to school'

---

[8] ATT = attributive marker (的 - 'de'), ASP = aspect marker (了 - 'le'), CLA = classifier (个- 'ge')

(14) *Non-inferable*    于是 两 个 小朋友    叫 来    了 一 个    士兵.

yúshì liǎng gè   xiǎopéngyǒu jiàolái   le **yī gè shìbīng**.

so   both CLA child       call over ASP **one CLA soldier**

'so both children called a soldier over'


(15) *Inferable*    这 个    时候 他们 叫 来    了    老师.

zhè gè   shí hòu tāmen jiào lái   le   **lǎo shī**.

this     time   they   call over ASP **teacher**

'this time they called the teacher over'


The logistic regression used for the Mandarin data showed that the non-inferable and neutral referents were significant negative predictors of introductions with bare nominals (with regression statistics of $\beta = -1.170$, Wald = 9.403, exp $\beta$ odds ratio = .310, $p < 0.005$ for the neutral referents and $\beta = -2.361$, Wald = 25.466, exp $\beta$ odds ratio = .094, $p < 0.001$ for the non-inferable referents), while the inferable referent was a significant positive predictor of referent introductions with this NP form ($\beta = 1.946$, Wald = 16.566, exp $\beta$ odds ratio = 7.000, $p < 0.001$). The predictors as a set reliably predicted bridging use (*chi* 82.725, $p < 0.001$, *df*=3) and a Nagelkerke's $R^2$ of .608 indicated a strong relationship between prediction and grouping. In comparison, a subsequent logistic regression regarding referent type and numeral + classifier + noun introductions showed that the non-inferable referent was a significant predictor of introductions with that NP form ($\beta = 1.986$, Wald = 24.275, exp $\beta$ odds ratio = 7.286, $p < 0.001$), while the neutrally inferable referent was a positive predictor yet not a significant one ($\beta = 0.539$, Wald = 2.569, exp $\beta$ odds ratio = 1.174, $p = 0.109$). This is likely due to the inanimate BASKETBALL referent receiving a relatively high number of bare nominal introductions (57.1% of 14 mentions, compared with

28.6% numeral + classifier + noun introductions)[9], while the animate neutral referents (BOY, GIRL) were introduced with numeral + classifier + noun introductions in almost all cases. The inferable referent type was a significant negative predictor of this NP form being used for introductions ($\beta = -3.664$, Wald $= 13.086$, exp $\beta$ odds ratio $= .026$, $p < 0.001$). The word order of all introductions was 76.5% postverbal / 23.5% preverbal, with no correlation between word order and referent type. From the results, there is a clear effect of inferability on local newness markings in Mandarin, with the numeral + classifier newness marking usually omitted for inferable referents.

*Korean*

Table 5 describes the NP forms used to introduce each referent type into the discourse.

---

[9] Another inanimate referent (the cannon) was introduced through numeral+classifiers in 16 out of 18 (88.9%) introductions, so animacy cannot be said to be a significant predictor of bare nominal introductions. There does not appear to be any effect of animacy on bare nominal introductions in Korean either, with both animate and inanimate referents likely to be introduced through bare nominals.

*NP forms Used to Introduce Each Referent Type in the Native Korean Data (n=20)*

| Referent type | NP form | Frequency | Percentage |
|---|---|---|---|
| Neutral (*k*=53) | Bare nominal (e.g., '소년' - 'boy') | 34 | 64.1 |
| | Numeral + noun (e.g., '한아이' - 'one boy') | 11 | 20.8 |
| | Quantifier + noun (e.g.'어떤농구공' - 'some basketball') | 6 | 11.3 |
| | Noun + numeral + classifier (e.g., '아이 한명' - 'boy one CLA') | 2 | 3.8 |
| Non-inferable (*k*=50) | Bare nominal (e.g., '장교님' - 'officer') | 43 | 86.0 |
| | Quantifier + noun (e.g '어떤 군인 – 'some soldier') | 4 | 8.0 |
| | Noun + numeral + classifier (e.g., '개한마리' - 'dog one CLA') | 3 | 6.0 |
| Inferable (*k*=32) | Bare nominal (e.g., '선생님' - 'teacher', '그물' - 'net') | 32 | 100 |

*Note.* *k* is total number of introductions per referent type; *n* is total number of Korean L1 participants. CLA = classifier (e.g. 명 – person)

It can be seen in Table 5 that both inferable and non-inferable referents were introduced through bare nominals, while the neutral referents (BOY/GIRL) were the only referents to occasionally receive a numeral before the noun when being introduced (in 44.4% of cases with BOY). These patterns are illustrated in (16) through (18):

(16) *Neutral*       **한 소년**이    학교를 가고 있었습니다.

**han sonyeon**-i hakgyoreul gago isseossseumnida.

**one child**-SUB school-OBJ go-and is-PAST-POL.

'a child was going to school'

(17) *Non-inferable*    그러자 **군인**이    와서.

geureoja **gunin**-i    waseo.

then    **soldier**-SUB came.

'then a soldier came'

(18) *Inferable*     선생님이      와서.

seonsaengnim-i waseo

**teacher**-SUB    came

'the teacher came'

A binary logistic regression analysis performed with referent type as a predictor for bare nominal introductions showed that both non-inferable and inferable referents were significant positive predictors of introductions with this form, with 100% cases for the inferable referents, and regression statistics of $\beta$ =1.516, Wald = 16.969, exp $\beta$ odds ratio = 4.456, $p$ <0.001 for the non-inferable referents, while the neutral referent type was not a significant predictor of this NP form. The predictors as a set reliably predicted bridging use (*chi* 67.009, *p*<0.001, *df*=3) and a Nagelkerke's $R^2$ of .522 indicated a strong relationship between prediction and grouping. The results of a subsequent logistic regression analysis showed that each referent type was a significant negative predictor of numeral + noun / noun + numeral + classifier introductions, with none of these NP forms attributed to inferable referents, and regression statistics of $\beta$ =−3.982, Wald = 14.843, exp $\beta$ odds ratio = 0.020, $p$ <0.001 for the non-inferable referents and $\beta$ =−1.587, Wald = 18.817, exp $\beta$ odds ratio = 0.205, $p$ <0.001 for the neutral referents.  These results show that there is no clear effect of inferability on local newness markings in Korean, with both inferable and non-inferable referents introduced through bare nominals.

***Summary***

Table 6 summarizes how the different referent types were introduced by the three native language groups.

*Summary of native language introductions*

| Referent Type | English | Mandarin | Korean |
| --- | --- | --- | --- |
| Neutral | Indefinite article + noun | Numeral + classifier + noun | Bare nominal |
| Non-inferable | Indefinite article + noun | Numeral + classifier + noun | Bare nominal |
| Inferable | Definite article + noun | Bare nominal | Bare nominal |

**Discussion**

Regarding Research Question 1, the data point toward the confirmation of the hypotheses formed for Study 1.  In English, definite article + N was used for inferable referents.  In Mandarin, optional newness markings were omitted for the inferable referents but were used for both the neutral and non-inferable referents.  In Korean, no distinction was made for the perceived inferability of the referent in the NP form used, with both inferable and non-inferable referents just as likely to be introduced through bare nominals.

These findings appear to corroborate Huang's (2000) distinction that Mandarin and Korean are more pragmatic than English in terms of the way that the inferability of referents is handled in these languages for referent introductions, as neither language marks inferability overtly, in contrast to English.  While Mandarin omits numeral + classifier constructions before the noun for inferable referents, in Korean, the (un)inferability of a particular referent in a particular context is all to be inferred by the audience through the context alone. However, as inferability appears to be a factor influencing the presence and absence of local newness marking in the Mandarin data, it can be argued that Mandarin may not be as pragmatic as Korean for the introduction of inferable referents, and may in fact be more similar to the syntactic mode of English in this regard.  If the numeral + classifier construction in Mandarin is indeed taking on the function of the English indefinite article, Mandarin learners may map this form to the indefinite article in the L2 in contexts where the numeral sounds natural in Mandarin (i.e., with neutral and non-inferable

referents). While Mandarin does not have a definite article, learners may map the bare NP used with bridging in Mandarin with the English definite article, once they realize that bridging introductions must be both marked for newness and marked differently to other referent types. If this is the case, the potential for positive transfer may aid the Mandarin learners of English, while the Korean learners of English would not be able to maximize positive transfer of this linguistic feature from their L1 to the L2, as bare NPs are used in both indefinite and definite contexts in Korean.

## Study 2: The Acquisition of Bridging First Mentions in L2 English

**Predictions**

In Study 2, two additional research questions were addressed. For Research Question 2, it was predicted that L2 learners would initially follow one of two possible tracks. Firstly, the learners may use the indefinite article for all referent introductions (regardless of inferability) and reserve the use of the definite article for discourse-old subsequent mentions. L2 learners are expected to notice that the indefinite article + noun construction is more frequent for referent introductions, and that the definite article is more frequent for discourse-old reference. Hence, they may wish to stick to this initial form/function mapping in their lower-level L2 production. Secondly, however, learners could use their native language strategy to introduce inferable referents through bare nominals (inappropriately for the L2). In either case, as proficiency level increases, the learners should gradually incorporate the use of the definite article for bridging descriptions. It was therefore predicted that the use of indefinite introductions to introduce the neutral and non-inferable referent types would be attested before definite introductions for inferable referents, due to the higher frequency of indefinite articles over definite articles for referent introductions generally, and given that definite article NPs are typically used for reference

maintenance rather than referent introductions (i.e., equating indefinite NPs with a one-to-one mapping of indefinite=new). For Research Question 3, it was predicted that as Mandarin makes a distinction for the inferability of referents for referent introductions, Mandarin L2 English learners may acquire the use of bridging at lower proficiency levels than their Korean counterparts, who first have to notice that newness markings are required for the majority of referents and secondly that the L2 makes a distinction for inferability in the form-to-function mapping for those markings.

**Method**

*Participants*

Sixty native Korean and 58 native Mandarin Chinese adults provided the L2 English narratives, with none of these participants providing the native data in Study 1. The L2 participants were all pre-tested for English proficiency level using the Oxford Quick Placement Test, a 30-minute multiple-choice standardized language test which can be used to ascertain a participant's CEFR level, from A1 to C2. In the Korean group, 10 participants were recruited for each CEFR level; in the Mandarin group, there were 10 participants each at levels A1 and B2, and 9 participants each at levels C1 and C2. Table 7 describes the L2 participants.

*Background of L2 Participants, Study 2*

|  |  | L1 Mandarin | L1 Korean |
|---|---|---|---|
| Total *N* |  | 58 | 60 |
| Male/female |  | 20/38 | 23/37 |
| Age ranges | 16-20 yrs | 6 | 11 |
|  | 20-30 yrs | 45 | 40 |
|  | 30-40 yrs | 4 | 6 |
|  | 40-50 yrs | 3 | 3 |
| CEFR Proficiency[a] | A1 | 10 | 10 |
|  | B2 | 10 | 10 |
|  | C1 | 9 | 10 |
|  | C2 | 9 | 10 |

[a]Common European Framework's four levels (Council of Europe, 2001) from Basic user (A1) to Proficient user (C2), obtained using the Oxford Quick Placement Test.

All of the Mandarin L2 English participants and 18 of the Korean L2 English participants were recruited in Cambridge, UK, and 42 of the Korean L2 English participants were recruited by the researcher in South Korea when not enough Korean-speaking candidates could be found in Cambridge. The Mandarin and Korean L2 English learners were either students at language schools in Cambridge or were students of Cambridge University, with the exception of the 42 participants recruited in Korea who were students at a private university (ages 20+) or at a high school in Korea specializing in English-language education (ages 16-19).

*Procedure*

The procedure for the L2 study followed the procedure for the native language study except for the following changes. All instructions were read to the participants in English, and the naïve listener was a native speaker of English in all cases, with glosses in English allowed if the participant did not understand certain words in the instructions (participants were excluded if they

were deemed not to have understood the instructions). During the telling of the story (see Supporting Information online), questions about vocabulary could be glossed with single words by the naïve listener if a gesture was made (e.g., a climbing motion for 'ladder') or if the participant was able to provide an oral description (e.g., 'a thing that shoots a bullet' for 'cannon'), except in cases where the question referred to one of the target referents, and the participant was not allowed to show the listener the picture sequence. The naïve listener would attempt to answer such questions with single word glosses (e.g., saying 'throw?' when the student made a throwing action) to help move the narrative forward, while any questions about the target referents were answered with a phrase such as 'let's move on' or 'if you are not sure, please say what happened next' by the naïve listener. The listener could also ask the participant to move on if they were unable to continue, although the listener was instructed only to do this if the participant had said nothing for 30 seconds or more. As with the L1 study, the naïve listener actively listened to the story, offering both linguistic (e.g., 'hmm', 'OK') and paralinguistic signals (e.g., nods, smiles) to reassure the speaker that their story was being listened to. Doing so was deemed particularly important for L2 speakers, who are less likely to produce or repair the appropriate markings in the L2 if the listener does not provide such signals (see Crosthwaite, 2011).

**Results**

*Mandarin L2 English*

Figures 1 and 2 display the NP forms used for the neutral and non-inferable referent types by the Mandarin L2 English group (henceforth ML2E).
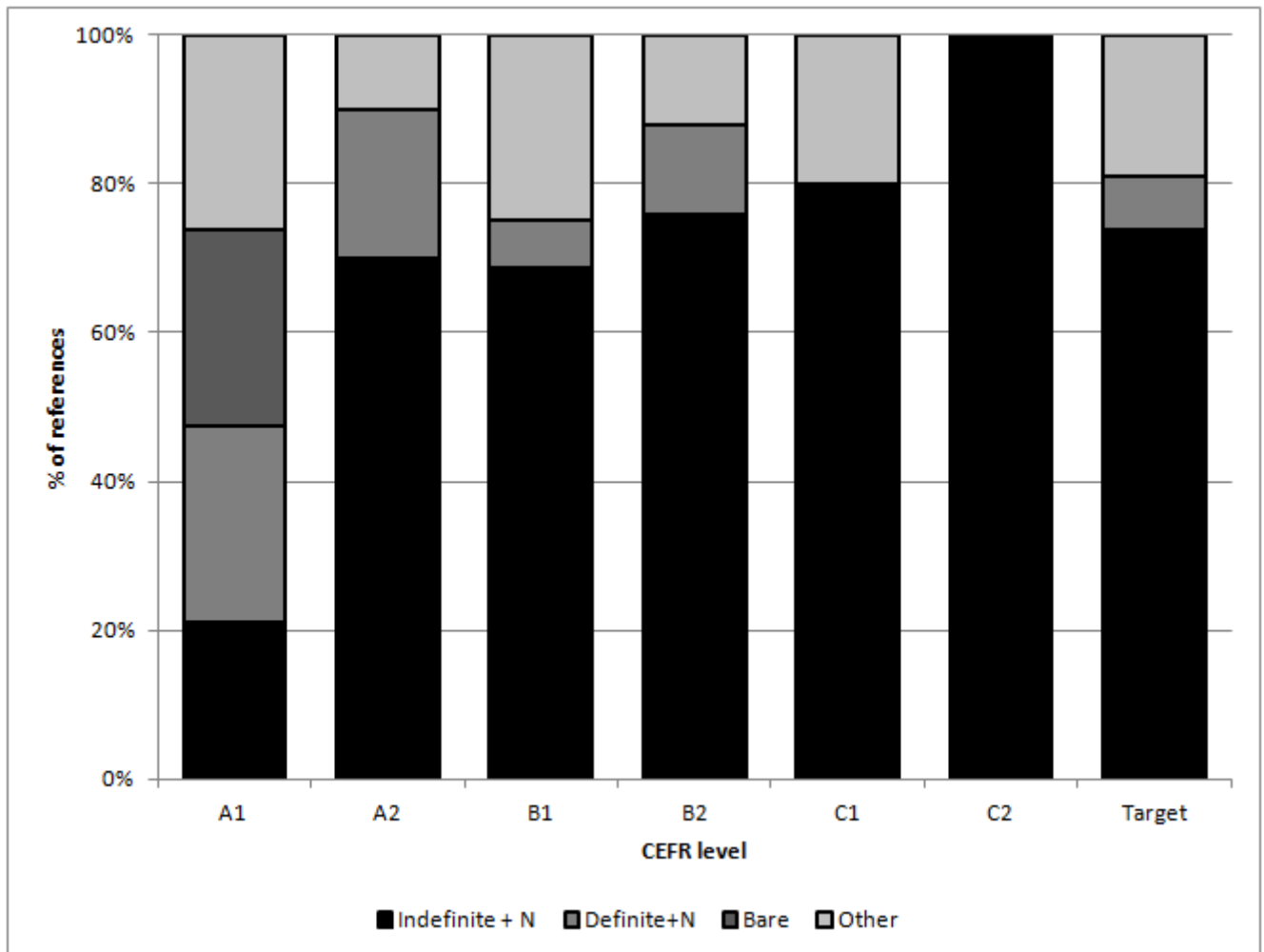
*Figure 1 - NP forms to introduce NEUTRAL referent, Mandarin L2 English*
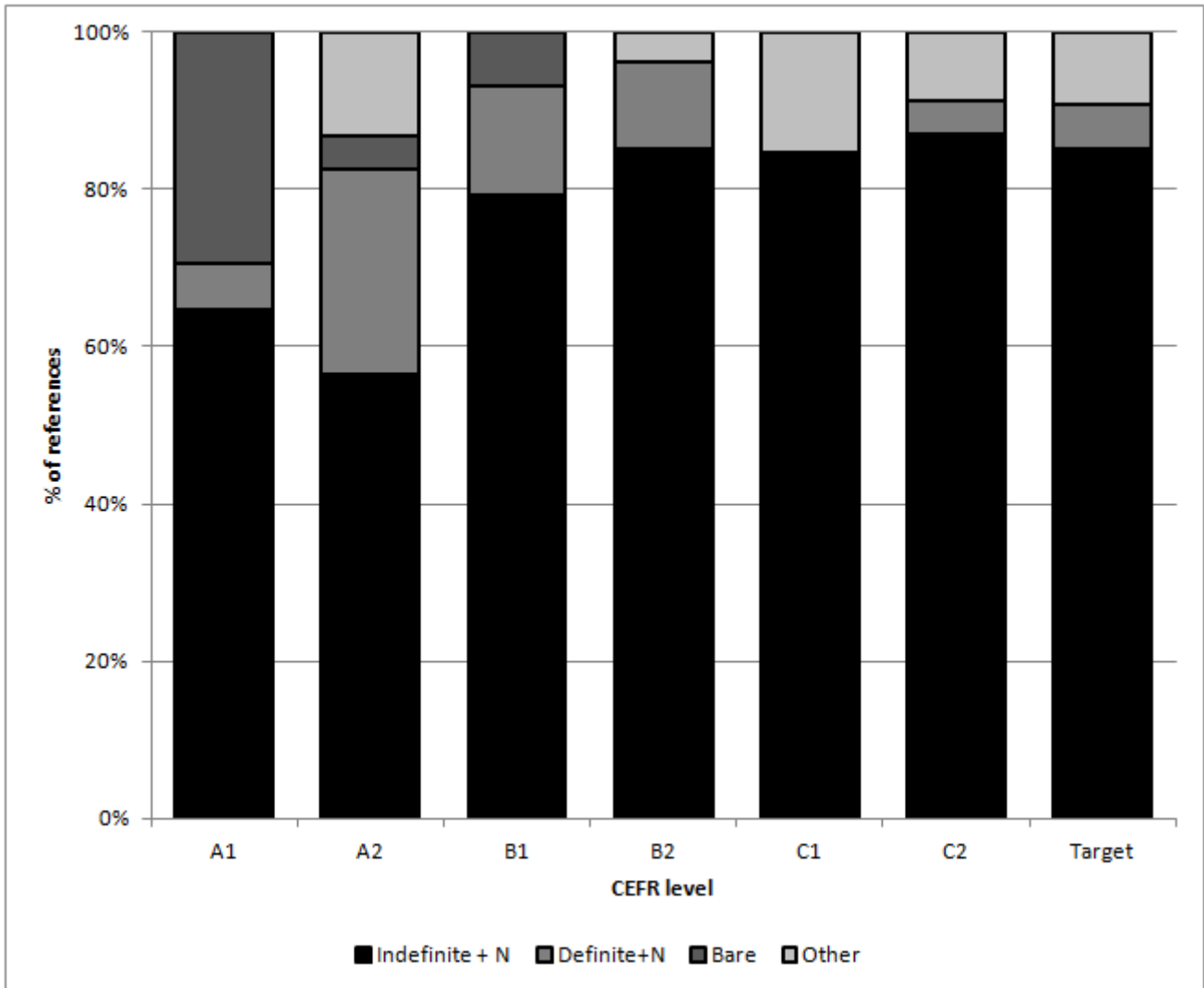
*Figure 2 - NP forms to introduce NON-INFERABLE referent, Mandarin L2 English*

At the A1 level, the proportion of indefinite article + noun introductions for the neutral referents is low (21.1%), but it increases with proficiency from the A2 level onward. For non-inferable referents, 56.5% of introductions are made with indefinite article introductions at the A1 level, rising to 64.7% at the A2 level, then remaining above 75% from the B1 level onward.

A pair of logistic regression analyses were performed (for the neutral and non-inferable referents) to determine whether an indefinite article + noun introduction was a significant predictor of the introductions of these referents at each CEFR level, measured against the likelihood of the same construction being used to introduce these referents in the native English

data. Target-like use was determined when the difference between the likelihood of the L2 group

and the native English group using that form in that context had a significance value of $p > 0.05$.

The results suggested that the learners were as likely as native English speakers to use indefinite

article introductions for these referents from the A2 level onward for the neutral referents

($\beta = -.551$, Wald $= 0.803$, exp $\beta$ odds ratio $= 0.600$, $p = 0.370$) as well as the non-inferable

referents ($\beta = -1.143$, Wald $= 3.232$, exp $\beta$ odds ratio $= 0.319$, $p = 0.072$).

Figure 3 shows the NP forms used to introduce the inferable referents by the ML2E group.
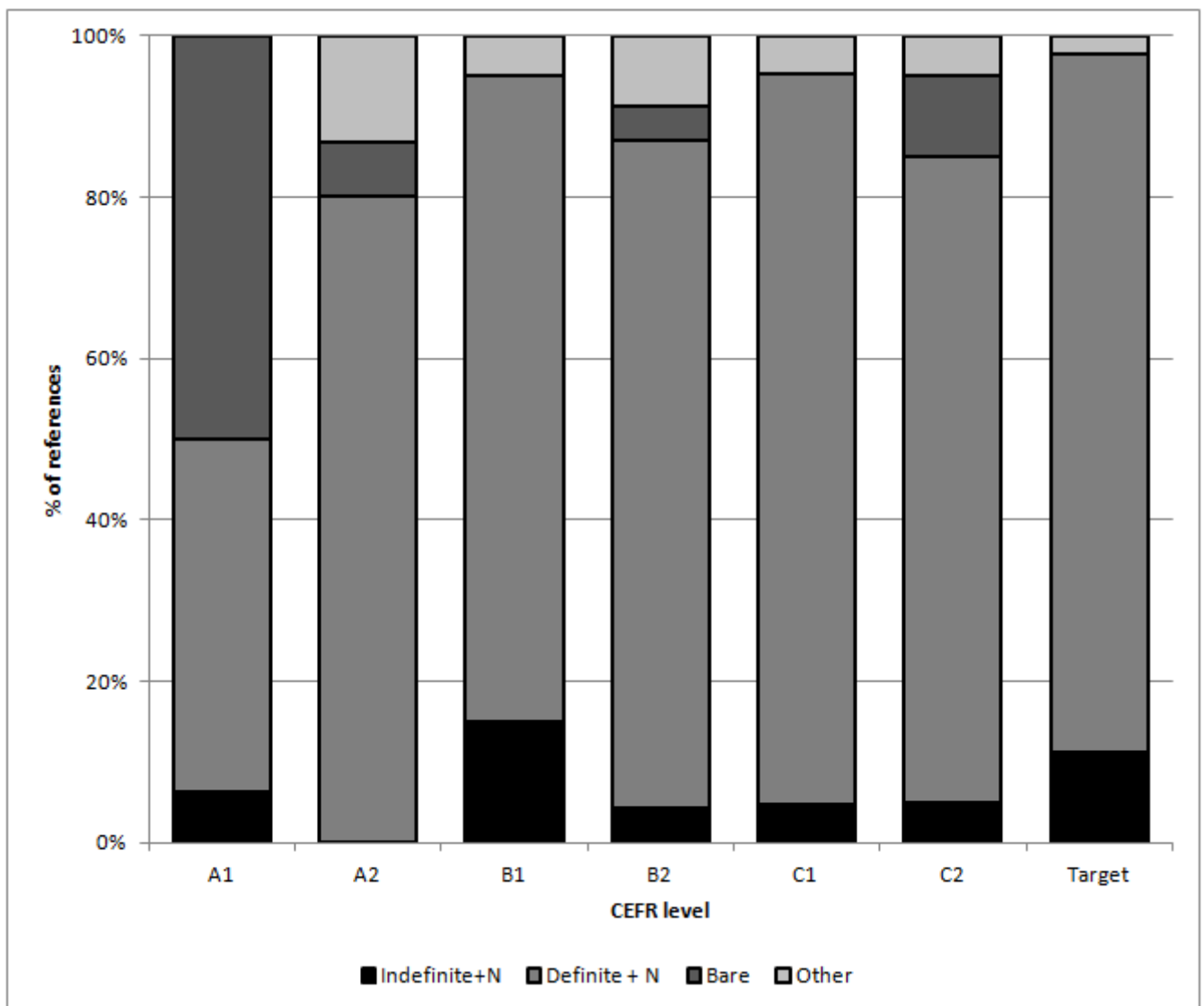


*Figure 3 - NP forms to introduce INFERABLE referent, Mandarin L2 English*

At the A1 level, a high proportion of bare nominals was used to introduce the inferable referents (50% of introductions). The use of indefinite articles to introduce neutral/non-inferable referents was not significantly different from the L2 target from A2 level. This suggests that the ML2E learners are following their native language strategy for introducing the inferable referents at the A1 level, as they are indeed not marking the inferable referents in the same way that they are marking the non-inferable referents (neutral referents notwithstanding). In fact, many learners at the A1 level do use the appropriate definite article + noun introduction for inferable referents (43.8% of cases), although they are still less likely to do so overall when compared to the L2 target usage. Beyond the A1 level, the use of definite articles for bridging is as high as 80%. A third logistic regression analysis was performed to determine whether a definite article + noun introduction was a significant predictor of the introductions of these referents at each CEFR level, measured against the likelihood of the same construction being used to introduce these referents in the native English data. The results show that the ML2E group were as likely as English native speakers to use definite article bridging introductions from the A2 level onward ($\beta = -0.486$, Wald $= 0.387$, exp $\beta$ odds ratio $= 0.615$, $p = 0.534$ at A2 level).

### *Korean L2 English*

Figures 4 and 5 display the NP forms used for the neutral and non-inferable referent types by the Korean L2 English group (henceforth KL2E).
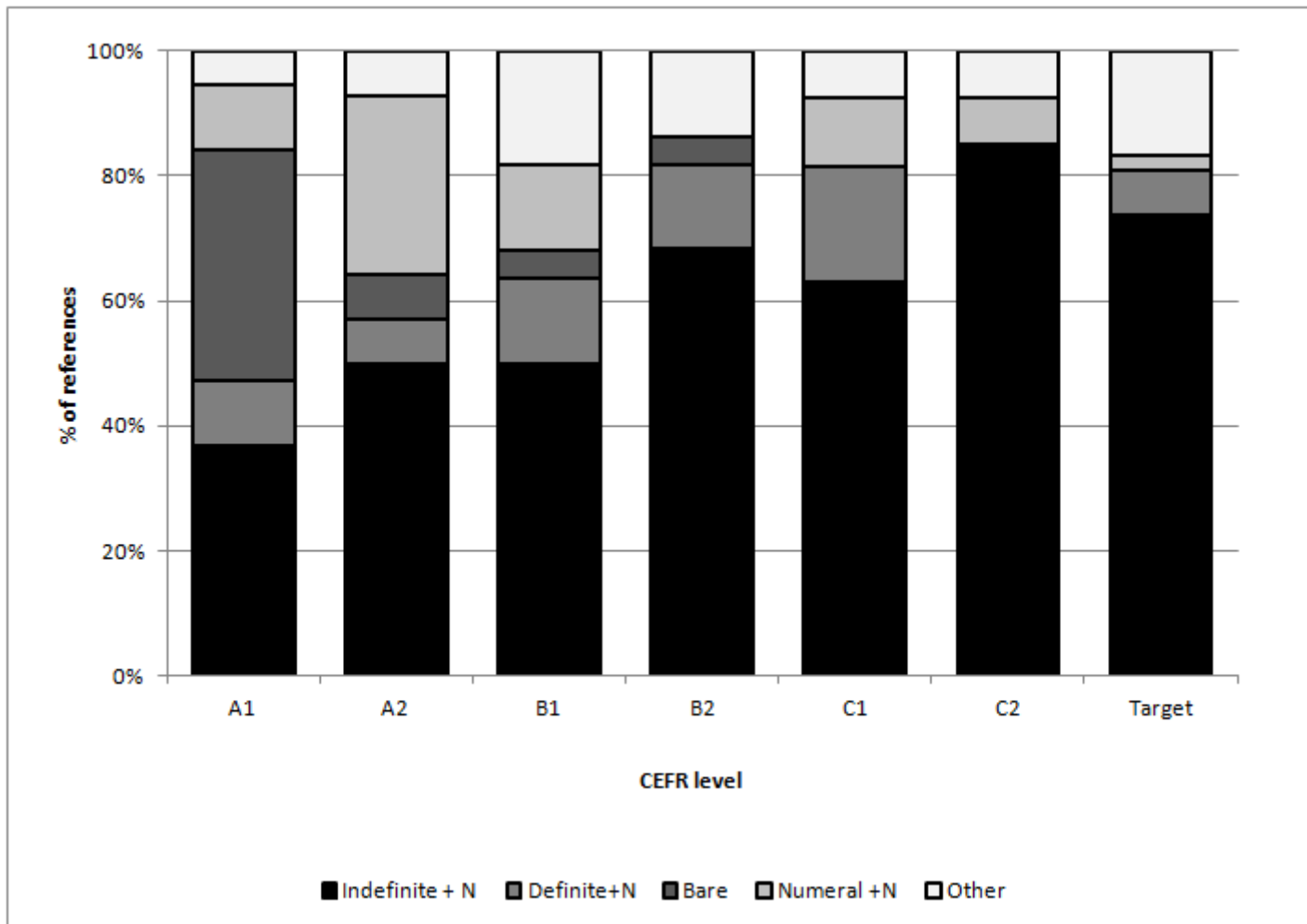
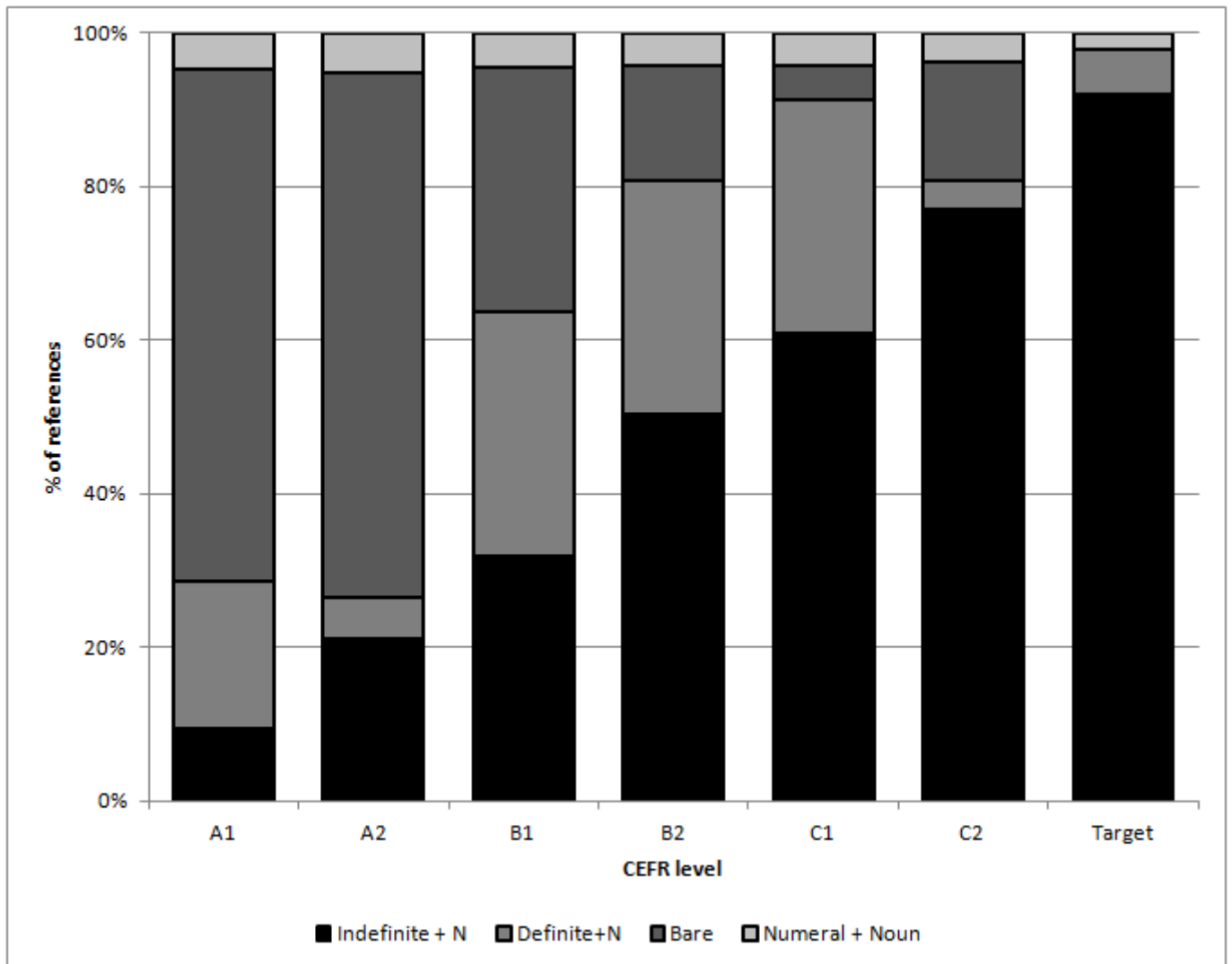*Figure 4 - NP forms to introduce NEUTRAL referent, Korean L2 English*

*Figure 5 - NP forms to introduce NON-INFERABLE referent, Korean L2 English*

For the neutral referents, at the A1 level, the proportion of indefinite article + noun introductions is low (36.8%) with an equal proportion of bare nominals. The use of indefinite article + noun introductions increases with proficiency from A2, with an accompanying rise in numeral + noun introductions at that level (28.6%) which are acceptable in the L2[10] as well as in the native Korean data (cf. the results for Study 1 shown in Table 5). For the non-inferable referents, native-like production of bare nominals is prevalent until the C1 level.

---

[10] Although unusual and only marginally acceptable, as seen in the native English data in Study 1.

ACQUISITION OF REFERENCE IN L2 NARRATIVE DISCOURSE

The results of the KL2E set of logistic regression analyses regarding indefinite article + noun introductions and neutral/non-inferable referents showed that the KL2E learners were as likely as English native speakers to use indefinite articles for neutral referents from the A2 level onward ($\beta = -.916$, Wald = 2.087, exp $\beta$ odds ratio = 0.400, $p = 0.149$) but as high as the B2 level for the non-inferable referents ($\beta = -0.936$, Wald = 1.906, exp $\beta$ odds ratio = 0.124, $p = 0.062$ at B2 level). The data suggests that the KL2E learners' L2 production was influenced by their native language preference for bare nominals until intermediate CEFR levels.

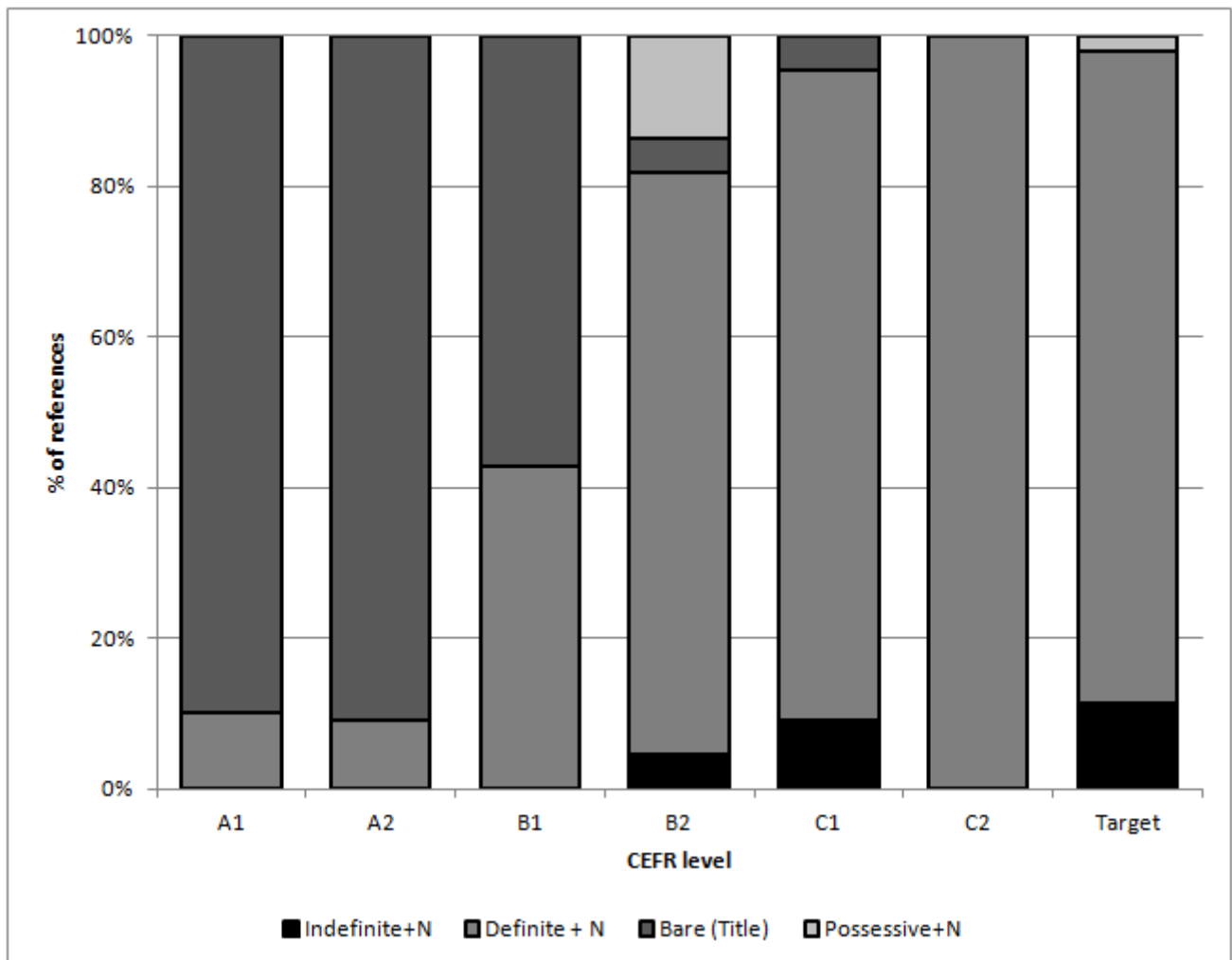Figure 6 shows the NP forms used to introduce the inferable referents by the KL2E group.



*Figure 6 - NP forms to introduce INFERABLE referent, Korean L2 English*

At the A1 through B1 levels, a high proportion of bare nominals was used to introduce the inferable referents (90%, 90.1%, and 57% of introductions at A1, A2, and B1 levels, respectively). As bare nominals were also used with the non-inferable referents it is unlikely that the use of bare nominals to introduce the inferable referents implies any notion of these referents' inferability. The proportion of use of the definite article to introduce inferable referents is quite high from B2 level onward. The results of the KL2E set of logistic regression analyses regarding definite article + noun introductions and inferable referents show that the KL2E group were as likely as native English speakers to use bridging from B2 level onward ($\beta = -0.648$, Wald $= 0.931$, exp $\beta$ odds ratio $= 0.523$, $p = 0.335$ at B2 level). Thus, the KL2E group used bare nominals for inferable referents from A1-B1 levels, then used the definite article + noun to appropriately introduce inferable referents from B2 level onward when compared to the native target usage, two CEFR levels *higher* than their ML2E counterparts.

Figure 7 compares the performance of both learner groups across each CEFR level when using bridging descriptions for inferable referents.
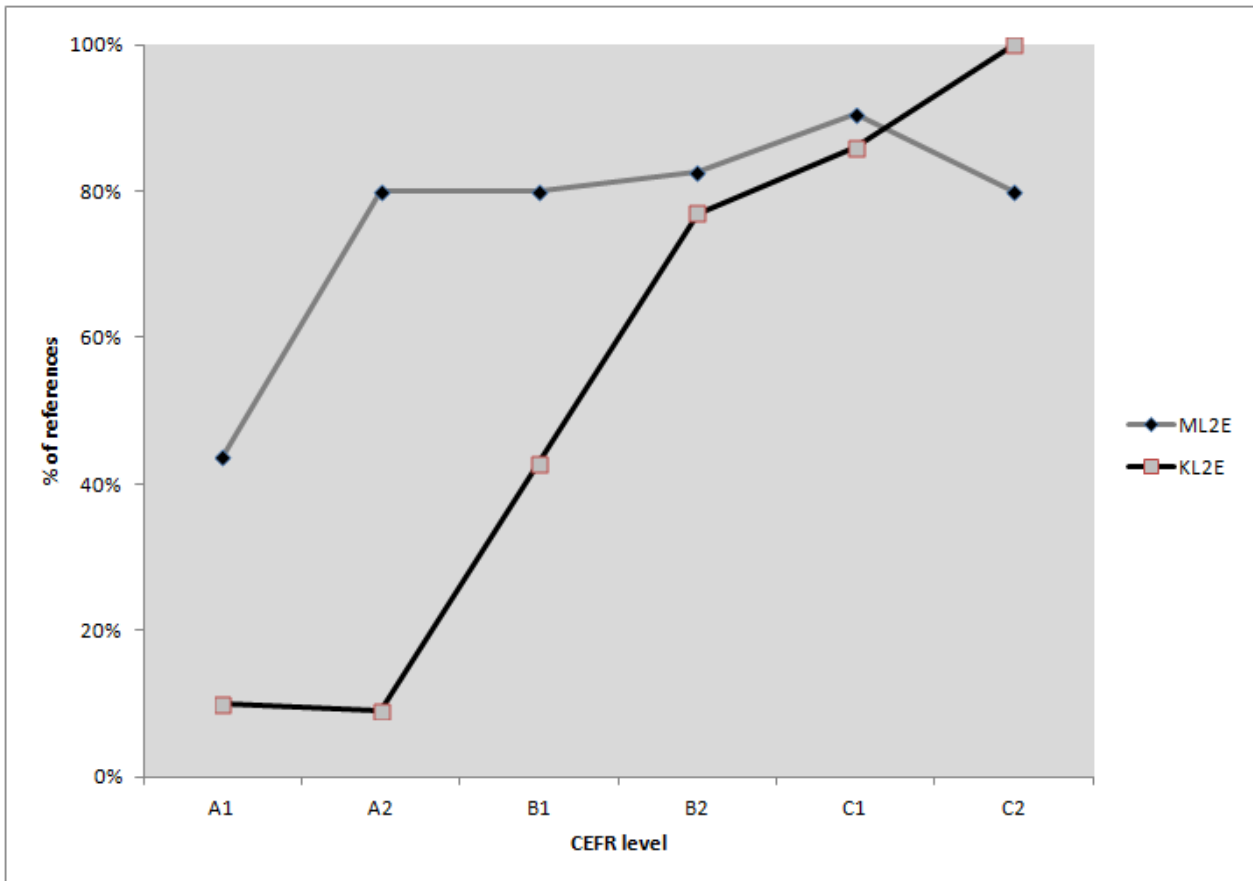
*Figure 7 - Comparison of bridging between Mandarin and Korean L2 English groups*

For Research Question 3, the data suggests that the ML2E group did manage to use the definite article for bridging descriptions at higher frequencies than their Korean counterparts at least until the B1 level, while the KL2E group eventually began to use this form more than the ML2E group from the C1 level onward (although both sets of learners had reached L2 target-like use before C1). A final set of logistic regression analyses was performed to see whether or not the likelihood of a definite article + noun being used to introduce the inferable referent was significantly different between the L2 groups at each CEFR level. The results showed significant differences between the L2 groups in favor of the ML2E group at the A1 level ($\beta = -2.197$, Wald $= 4.345$, exp $\beta$ odds ratio $= 0.111$, $p < 0.05$), A2 level ($\beta = -2.303$, Wald $= 4.820$, exp $\beta$ odds ratio $= 0.100$, $p < 0.05$), and B1 level ($\beta = -1.163$, Wald $= 5.154$, exp $\beta$ odds ratio $= 0.313$, $p < 0.05$).

Therefore, the ML2E group outperformed the KL2E group at statistically significant values from levels A1 and B1 when compared with each other, not taking into account the native English data.

*Summary*

The L2 data collected in Study 2 appear to show that for Research Question 2, L1 Mandarin speakers learning L2 English are as likely as native English speakers to use the definite article for bridging descriptions from the A2 level of the CEFR, with their Korean counterparts as likely as native speakers to use bridging from the B2 level. The use of the definite article for bridging descriptions was attested at higher CEFR proficiencies than the use of the indefinite article for neutral and non-inferable referent introductions in both L2 groups. For Research Question 3, the data suggest that the ML2E group has acquired bridging definite introductions at lower CEFR proficiencies than the KL2E group (from the A2 level compared to the B2 level) when compared with the native English data, with significant differences in performance between the two learner groups from A1 to B1 levels when compared directly with each other.

**Discussion**

Regarding Research Question 2, the data suggest that the English syntactic strategy of introducing inferable referents through definite bridging descriptions is attested at an intermediate stage in language acquisition (A2/B2 levels of the CEFR), and that indefinite introductions (of other referent types) are attested before definite introductions. These findings were predictable in terms of the relatively low frequency of definite introductions compared to indefinite introductions in English, and given the relative complexity of the form/function relationships for the definite article compared with those of the indefinite.

For Research Question 3, I would like to suggest that positive transfer effects played a role in the relative success of the ML2E learners compared to the KL2E learners when acquiring

bridging. At the lowest proficiency level (A1), both sets of L2 learners used bare nominals to introduce the inferable referents, and this finding is indicative of a general early developmental stage in L2 acquisition that is attested regardless of whether the source and target languages have articles or not (e.g., the Basic Variety; Klein & Perdue, 1992, 1997). However, even at this low level (and as seen in the native Mandarin data in Study 1), the ML2E group introduced inferable referents differently from the other referent types. Despite this, given that in native Mandarin numerals and classifiers were omitted for inferable referents, it is uncertain whether or not the ML2E group took into account the inferability of the referents when omitting the article at A1. As the ML2E group appeared to have acquired the use of the indefinite article + noun to introduce the non-inferable referents from the A1 level and did not use indefinite articles to introduce the inferable referents, I suggest that learners in the ML2E group were indeed making a distinction for inferability when introducing inferable referents, but were not aware at the A1 level that they should be using the definite article for this function - instead omitting the article altogether as they would omit the numeral + classifier in their native language. From A2 onward, if the ML2E group did not take the inferability of the referent into account during referent introductions, then they would likely have transferred their use of the numeral + classifier + noun construction into using the English indefinite article for the inferable referents, as they did with the neutral and non-inferable referents. The ML2E group did not do this, however, and instead marked referent introductions for inferability in an L2-like manner at lower CEFR levels than the KL2E group, as they were able to positively transfer their native-like strategy of introducing inferable referents differently to other referent types. This would imply that the ML2E group is already on a 'semi-syntactic' path toward the English 'syntactic' system, following Huang's (2000) taxonomy, while this is questionable in Korean (e.g., Ionin, et al., 2012). The present data suggest that L1 Mandarin speakers may therefore have less difficulty mapping L2 (English) syntactic form to

pragmatic function in their L2 (English) narrative production due to the features shared between each language.

The same cannot be said for the KL2E group, who continued to refer to both the inferable and non-inferable referents through bare nominals until B1/B2 levels with no apparent distinction for inferability. As the use of the indefinite article to refer to the neutral referents by the KL2E group was also not attested in the data until the A2 level, with bare nominals preferred instead for both neutral and non-inferable referents, it is apparent that the KL2E group struggled to use articles at all (regardless of inferability) at lower CEFR levels. At higher CEFR levels, the KL2E group eventually began to use indefinite articles for the other referent types (despite a fairly high frequency of definite article introductions for the non-inferable referent until C1 level). At this stage, it may then be argued that using the correct determiner for the inferability of the referent is difficult for these learners, once the native Korean-like preference for bare nominals was no longer a factor for any of the referent types. However, the KL2E group appears to be at a distinct disadvantage over their ML2E counterparts when acquiring bridging in the L2, as they are not able to positively transfer any strategy for marking inferability from their L1, instead negatively transferring their bare-nominal led strategy when introducing all referent types regardless of inferability.

**Limitations and Suggestions for Future Research**

The main limitation of the present twin study is that the number of inferable scenarios in the picture sequence was limited to a single animate scenario-dependent referent (the teacher) and three further inanimate referents (the playground, the hoop, and the backboard) in part/whole relationships, although animacy was a not a significant predictor of NP form in the data. The materials were designed as part of ongoing research into L2 discourse reference more generally, rather than being designed strictly with inferable contexts in mind. Further research is needed to

investigate whether L2 learners find it easier or more difficult to acquire the L2 target bridging of other kinds of inferables (such as cause/consequence relationships, or optional/necessary relationships). It would also be interesting to observe how English-speaking L2 Mandarin or L2 Korean learners manage to introduce inferable referents in these languages, from a 'syntactic' to a 'semi-syntactic' (Mandarin) or 'pragmatic' (Korean) strategy. The relatively small L2 dataset (*n*=10 for each CEFR level per language group) is also a confounding factor when attempting to generalize from the results, although the method of statistical analysis used still features significant differences in the likelihood of certain NP forms being used for certain referent types.

## Conclusion

The evidence offered here suggests that bridging descriptions for referent introductions are acquired between the A2 (Mandarin L2 English) and the B2 (Korean L2 English) levels of the CEFR proficiency scale. The data also suggest that definite article introductions for inferable referents are attested at higher CEFR proficiencies than indefinite article introductions for other referent types. Moreover, the findings support the claim that the learnability of definite discourse-new reference is strongly influenced by L1 transfer, in that in the present study, learnability of bridging appeared easier if a distinction for inferability determined NP selection in both the native and target languages. This is realized in the adoption of definite articles to introduce inferable referents in L2 English narrative discourse by Mandarin learners of L2 English earlier on the cline of L2 proficiency than Korean learners.

**References**

Aiping, M. (2003) *Bridging reference in Chinese discourse: A pragmatic-cognitive approach.* Unpublished doctoral dissertation. Guangdong University of Foreign Studies, China.

Andersen, R. W. (1984). The one-to-one principle of interlanguage construction. *Language Learning*, *34*, 77–95.

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics*, *16*, 443–463.

Ariel, M. (1996). Referring expressions and the +/- coreference distinction. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility*. Amsterdam: John Benjamins.

Ariel, M. (2008). *Pragmatics and grammar*. Cambridge, UK: Cambridge University Press.

Ariel, M. (2010). *Defining pragmatics*. Cambridge, UK: Cambridge University Press.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*, 521–536.

Brown, L. (2011). *Korean honorifics and politeness in second language learning*. Amsterdam: John Benjamins.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Chafe, W. L. (Ed.). (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Chen, P. (2004). Identifiability and definiteness in Chinese. *Linguistics*, *42*, 1129–1184.

Clark, H. H. (1975). Bridging. In R.C. Schank, & B. L.Nash-Webber, (Eds.), *Theoretical issues in natural language processing* (pp. 169-174). New York: Association for Computing Machinery.

Clark, H. H., & Haviland, S. (1977). Comprehension and the given-new contrast. In R. Freedle, (Ed.) *Discourse Production and Comprehension* (pp. 1-40). Norwood, N.J.: Ablex.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.

Crosthwaite, P. (2011). The effect of collaboration on the cohesion and coherence of L2 narrative discourse between English NS and Korean L2 English users. *Asian EFL Journal*, *13*(4), 135-166.

Crosthwaite, P. (2013). An error analysis of L2 English discourse reference through learner corpora analysis. *Linguistic Research, 30*(2), 163-193.

Crosthwaite, P. (2014). *Differences between the coherence of Mandarin and Korean L2 English learner production and English native speakers: An empirical study*. Unpublished doctoral dissertation, University of Cambridge, UK.

Ekiert, M. (2007). The acquisition of grammatical marking of indefiniteness with the indefinite article a in L2 English. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, *7*(1). Retrieved 8 May 2014 from: http://journal.tc-library.org/index.php/tesol/article/viewFile/265/224

Ekiert, M. (2010). Linguistic effects on thinking for writing: The case of articles in L2 English. In Z. Han & T. Cadierno (Eds.), *Linguistic relativity in SLA: Thinking for speaking* (pp. 125–153). Bristol, UK: Multilingual Matters.

Ellis, N. C. (1994). *Implicit and explicit learning of languages*. London: Academic Press.

Filipovic, L., & Hawkins, J. A. (2013). Multiple factors in second language acquisition: The CASP model. *Linguistics*, *51*, 145–176.

Gardent, C., Manuélian, H., & Kow, E. (2003). Which bridges for bridging definite descriptions? In *Proceedings of the EACL 2003 Workshop on Linguistically Interpreted Corpora*, Budapest (pp. 69–76). Retrieved 8 May 2014 from: http://www.cnrtl.fr/corpus/dede/gardent-manuelianLINC.pdf

Gundel, J. K. (1996). Relevance Theory meets the Givenness Hierarchy. An account of inferables. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 155–179). Amsterdam: John Benjamins.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*, 274–307.

Hawkins, J. A. (1978). *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croom Helm.

Hedberg, N. (1996). Word order and cognitive status in Mandarin. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 179–192). Amsterdam: John Benjamins.

Hellman, C. (1996). The'price tag'on knowledge activation in discourse processing. In T. Fretheim & J. K. Gundel (Eds.), *Reference and referent accessibility* (pp. 193–211). Amsterdam: John Benjamins.

Hickmann, M., Hendriks, H., Roland, F., & Liang, J. (1996). The marking of new information in children's narratives: A comparison of English, French, German and Mandarin Chinese. *Journal of Child Language*, *23*, 591–620.

Huang, Y. (2000). *Anaphora: A cross-linguistic study*. New York: Oxford University Press.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Ionin, T., Baek, S., Kim, E., Ko, H., & Wexler, K. (2012). That's not so different from the: Definite and demonstrative descriptions in second language acquisition. *Second Language Research*, *28*, 69–101.

Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, *12*, 3–69.

Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, *24*, 387–418.

Klein, W., & Perdue, C. (Eds.). (1992). *Utterance structure: Developing grammars again*. Amsterdam: John Benjamins.

Klein, W., & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research*, *13*, 301–347.

Krashen, S. (1977). Some issues relating to the Monitor Model. In H. Brown, C. Yorio, & R. Crymes (Eds.), *On TESOL'77* (pp. 144–158). Washington D.C.: TESOL.

Kwon, N., Polinsky, M., & Kluender, R. (2006). Subject preference in Korean. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th West Coast conference on Formal Linguistics (WCCFL25).* (pp. 1–14). Somerville, MA: Cascadilla Press.

Leung, Y. I. (2001). The initial state of L3A: Full transfer and failed features? In X. Bonch-Bruevich, W. Crawford, J. Hellerman, C. Higgins, & H. Nguyen (Eds.), *The past, present, and future of second language research: Selected proceedings of the 2000 Second Language Research Forum* (pp. 55–75). Somerville, MA: Cascadilla Press.

Li, C. N., & Thompson, S. A. (1976). Subject and topic: A new typology of language in subject and topic. In C. N. Li (Ed.), *Subject and Topic* (p. 475). New York: Academic Press.

Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. California: University of California Press.

Liu, M. (2010). Emergence of the indefinite article. In A. van Linden, J. Verstraete, & K. Davidse (Eds.), *Formal Evidence in Grammaticalization Research* (p. 275).

Luk, Z. P., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning*, *59*, 721–757.

Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation. University of California at Los Angeles.

Matsui, T. (2000). *Bridging and relevance*. Amsterdam: John Benjamins.

Nakahama, Y. (2009). Cross-linguistic influence on referent introduction and tracking in Japanese as a second language. *Modern Language Journal*, *93*, 241–260.

Poesio, M., & Vieira, R. (1998). A corpus-based investigation of definite description use. *Computational Linguistics*, *24*, 183–216.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223–256). New York: Academic Press.

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, *16*, 135.

Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language: Explorations of comprehension beyond the sentence*. New York: Wiley.

Sasaki, M. (1987). Is Uguisu an exceptional case of 'idiosyncratic variation'? Another counterexample to the 'natural order.' *Chugoku-Shikoku Academic Society of Education Research Bulletin*, *32*, 170–174.

Sleeman, P. (2004). Guided learners of French and the acquisition of emphatic constructions. *International Review of Applied Linguistics in Language Teaching*, *42*, 129–152.

Slobin, D. I. (1975). The more it changes…On understanding language by watching it move through time. *Papers and Reports on Child Language Development*, *10*, 1–30.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Cambridge, MA: Blackwell.

Stalnaker, R. (1974). Pragmatic presuppositions. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and Philosophy: Essays* (pp. 197–213). New York: New York University Press.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. *New Directions in Discourse Processing*, *2*, 53–120.

Van Dijk, T. A. (1976). Narrative macro-structures. Cognitive and logical foundations. *PTL: A Journal for Descriptive Poetics and Theory of Literature*, *1*, 547–568.

Vieira, R. (1998). *Definite description resolution in unrestricted text*. Unpublished doctoral dissertation. University of Edinburgh, Centre for Cognitive Science.

Vieira, R., & Poesio, M. (2000). An empirically based system for processing definite descriptions. *Computational Linguistics*, *26*, 539–593.

Wilson, D., & Matsui, T. (1998). Recent approaches to bridging: Truth, coherence, relevance. *UCL Working Papers in Linguistics*, 10, 173–200.

Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D. R. Preston (Eds.), *Second Language Acquisition and Linguistic Variation* (pp. 135–175). Amsterdam: John Benjamins.

**Supporting Information**

Additional Supporting Information may be found in the online version of this article at the publisher's website:
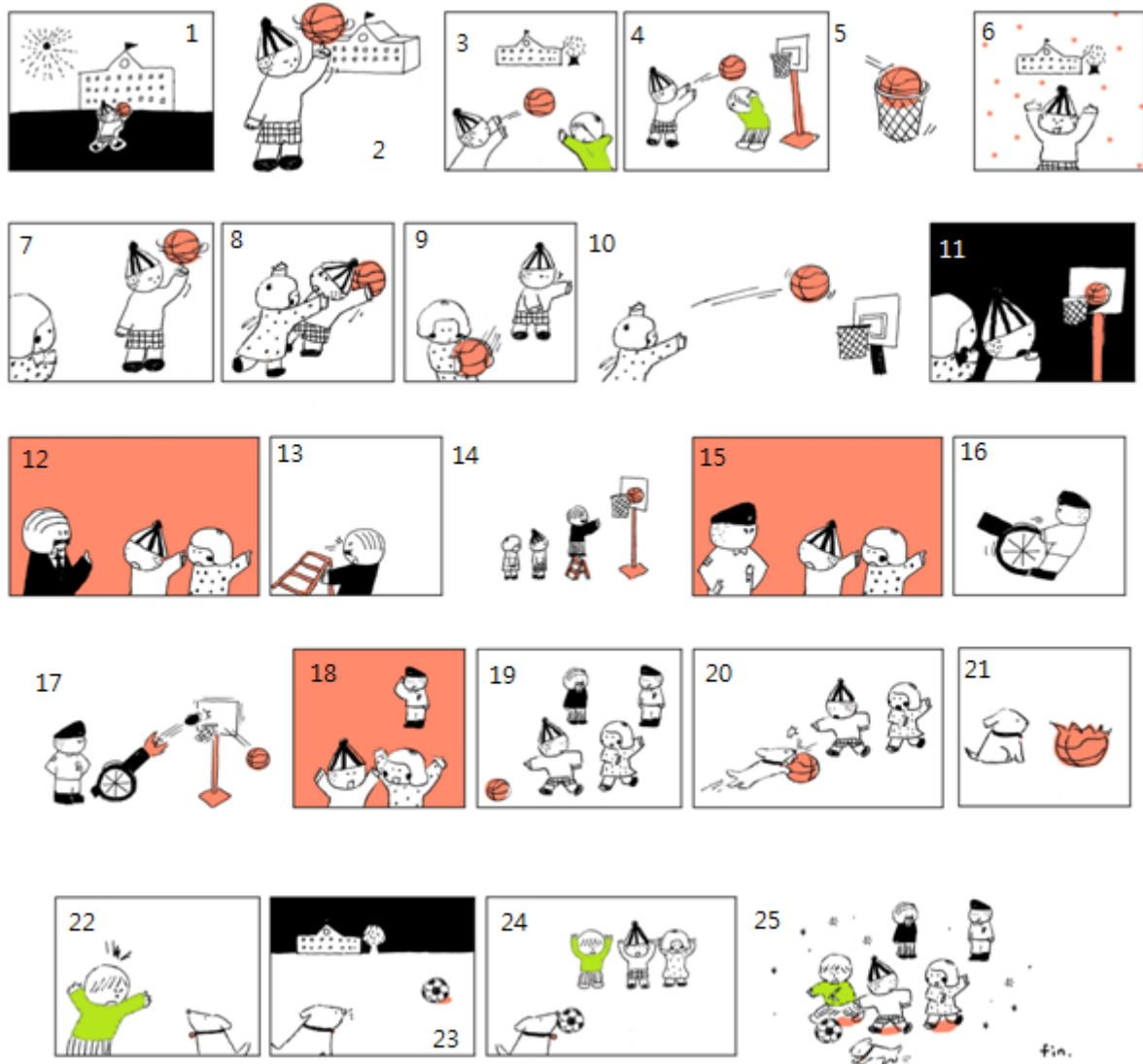
**Appendix S1**: Elicitation Materials

*Figure 8 - Elicitation materials*