Are L2 English Article Choices UG-regulated?*  

Nattama Pongpairoj  
University of York

The paper reports English article substitutions by L1 Thai (-articles) and L1 French (+article) speakers. The study was a semi-replication of the forced-choice elicitation task from Ionin, Ko & Wexler (IKW) (2004). IKW report that Russian and Korean speakers overuse a(n) in [+def; spec] and the in [-def; +spec] contexts, and argue that this pattern is a result of the learners fluctuating between the two settings of the Article Choice Parameter, one linked to [+definiteness], the other to [+specificity]. The problem is that the claim is based on materials in which ‘specificity’ was operationalised as the speaker’s explicitly stated knowledge (ESK) of the referent being talked about (see Trenkic 2007b). The values of [spec] and ESK were conflated in all contexts ([+spec; +ESK]; [-spec; -ESK]). However, a referent can be [+spec], even when personal knowledge is denied. It is thus possible that in IKW’s study the article choice was not influenced by [+spec] (UG-regulated), but by what was explicitly claimed about the referent. This possibility was investigated by adding cases where [spec] and ESK values were separated: [+spec; -ESK]. Repeated-measures ANOVA confirmed that, on this test, the L1 Thai learners’ article distinctions were influenced by ESK, and not by [+spec]. This is in line with the proposal in Trenkic (2007a) that L2 speakers from (-article) languages misanalyse and use English articles as lexical elements, attributing them referential, commonsense meanings of definite/indefinite (’that can be/cannot be identified’). In contrast, L1 French learners’ article choices were mainly appropriate, suggesting a correct analysis of English articles as determiners, and therefore having syntactically-triggered production.

1 INTRODUCTION

A particular problem to L2 speakers from (-article) L1 backgrounds is persistent variability in article production, even at advanced levels. L2 speakers sometimes use a(n) instead of the, and vice versa. L2 English article substitution problems are the focus of this study.

This paper is organised as follows. Section 2 presents Ionin, Ko and Wexler (IKW)’s (2004) account, which is related to L2 English article substitutions. Section 3 points to some problematic issues in IKW’s study. Section 4 introduces the present study. Section 5 reports the results and section 6 discusses these results, including their implications. Finally, section 7 reports the conclusions of the study.

2. IKW’S (2004) ACCOUNT

There have been several studies into English article substitutions (cf. Trenkic 2002; Leung 2005; among others). A recent study exploring this issue is IKW (2004). IKW (2004:12)

* I would like to express my gratitude to Danijela Trenkic, my supervisor, for her valuable advice and guidance, including the use of materials in this study. I also sincerely thank anonymous reviewers of this article for their useful comments. Thanks are also due to the audience at CamLing 2007 for their constructive comments and questions. Any errors remain my own. I am also grateful to the Department of Language & Linguistic Science at the University of York for generously providing financial support in terms of payment to the participants in this study.
Nattama Pongpairoj

proposed the Article Choice Parameter (ACP) for languages with two articles. According to the ACP, ‘definiteness’ and ‘specificity’ are cross-linguistic article semantic features evidenced in two-article languages:

(1) The Definiteness Setting: Articles are distinguished on the basis of definiteness.
The Specificity Setting: Articles are distinguished on the basis of specificity.

For example, the definiteness setting is found in English while the specificity setting is evidenced in Samoan. Article systems encoding [+def] are claimed to cut across the distinction of [+spec] (cf. table 1), and vice versa (cf. table 2):

<table>
<thead>
<tr>
<th></th>
<th>[+def]</th>
<th>[-def]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+spec]</td>
<td>the</td>
<td>a(n)</td>
</tr>
<tr>
<td>[-spec]</td>
<td>the</td>
<td>a(n)</td>
</tr>
</tbody>
</table>

Table 1
[+def] cutting across [+spec] in languages encoding [+def]

<table>
<thead>
<tr>
<th></th>
<th>[+def]</th>
<th>[-def]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+spec]</td>
<td>le</td>
<td>le</td>
</tr>
<tr>
<td>[-spec]</td>
<td>se</td>
<td>se</td>
</tr>
</tbody>
</table>

Table 2
[+spec] cutting across [+def] in languages encoding [+spec]

By combining the assumption that UG is available to L2 speakers with the theoretical construct of the ACP, IKW formulated the Fluctuation Hypothesis (FH) (2004:15). That is, the two article settings in the ACP are fully accessible to L2 speakers. L2 speakers will fluctuate between these settings until sufficient input causes the appropriate semantic parameter to be established as the correct setting for the language. The following patterns were thus predicted to occur for L2 speakers of (-article) L1, when speaking English:

<table>
<thead>
<tr>
<th></th>
<th>+ specific</th>
<th>-specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>+definite (target: the)</td>
<td>correct use of the</td>
<td>overuse of the</td>
</tr>
<tr>
<td>-definite (target: a(n))</td>
<td>overuse of a(n)</td>
<td>correct use of a(n)</td>
</tr>
</tbody>
</table>

Table 3
Predictions of L2 article use based on semantic properties of [def] and [spec]

If [def] and [spec] are of the same value, correct use of the for [+def; +spec] and a(n) for [-def; -spec] is predicted. However, if the two values clash, overuse of a(n) should be found in [+def; -spec] and the in [-def; +spec] contexts. IKW tested their predictions on L2 speakers of English whose L1s were Korean and Russian (-article) languages, by using a forced-choice elicitation task. The results reported supported their predictions.

3 PROBLEMS WITH IKW (2004)’S STUDY

According to IKW, an NP is specific when the speaker (a) has a referent in mind and (b) intends to refer to it. However, they also claim that the speaker must also ‘consider this individual to possess some noteworthy property’ (IKW 2004:5). For example, in ‘Peter intends to marry a/this merchant banker—even though he doesn’t get on at all with her’ (Lyons 1999, quoted in IKW 2004:6), the NP a/this merchant banker is specific because the speaker intends to refer to a particular merchant banker with the noteworthy property that ‘Peter does not get on at all with her’.

However, as Trenkic (2007b) notes, from a closer observation of the forced-choice elicitation task materials in IKW (2004), there might be some problems with their claim, due to how the ‘specificity’ was operationalised. In their materials, [+spec] was operationalised through the speaker explicitly stating or denying knowledge of the referent being talked about.
Are L2 English Article Choices UG-regulated?

(i.e. ‘explicitly stated/denied knowledge’ or ±ESK). For example, (4) describes how [-definite; +specific] context was operationalised:

(4)  [-indefinite, +specific]
    Meeting on a street
    Roberta: Hi, William! It’s nice to see you again. I didn’t know that you were in Boston.
    William: I am here for a week. I am visiting (a, the, --) friend from college—his name is Sam Brown, and he lives in Cambridge now. (IKW 2004:22)

IKW reported that the L2 speakers were likely to overgenerate the in such contexts, and they attributed this tendency to the positive value of specificity. For now, note that the specificity was operationalised as the speaker claiming to know the referent being talked about, i.e. that this person’s name is ‘Sam Brown’ and that ‘he lives in Cambridge now’.

Example (5) operationalises a [-definite; -specific] context:

(5)  [-definite,-specific]
    Chris: I need to find your roommate Jonathan right away.
    Clara: He is not here—he went to New York.
    Chris: Really? In what part of New York is he staying?
    Clara: I don’t really know. He is staying with (a, the, --) friend—but he didn’t tell me who that is. He didn’t leave me any phone number or address. (IKW 2004:22)

IKW reported that the L2 speakers tended not to overuse the in such contexts, and attributed this tendency to the negative value of specificity. Note that the non-specificity was operationalised as the speaker denying the knowledge of the identity of a friend.

So, the contexts from IKW’s materials covered only cases where the two unrelated variables of ‘specificity’ and ‘explicitly stated knowledge’ were conflated and of the same value. The context was either [+spec; +ESK] or [-spec; -ESK]. However, as Trenkic (2007b) observes, it is possible for a [-def] context to be [+spec] and [-ESK]:

(6)  [-definite, + specific] (- explicit speaker knowledge)
    Office gossip
    Gina: and what about the others?
    Mary: Well, Dave is single, Paul is happily married, and Peter…he is engaged to a/ this merchant banker, but none of us knows who she is or what she is like.

The referent a merchant banker is [-def] and [-ESK]. However, in English, when an indefinite NP can be felicitously introduced in discourse by this, that NP is treated as specific (cf. Prince 1981; Foder & Sag 1982:360). In this indefinite context, the speaker has a particular referent in mind and has an intention to refer to this person but denies personal knowledge of her. Such a context was not covered by IKW (2004). In their materials, the noteworthy property of a [+spec] NP was always linked to [+ESK], and vice versa. The speaker’s knowledge was concerned with the referent’s identifying attributes (name, appearance, characteristics, etc). There was no [+spec;-ESK] context. It is therefore important to use the contexts where the values of [spec] and [ESK] do not match, to test the FH. If learners associate article choice with [+spec], overuse of the is expected in [+spec; +ESK] and [+spec; -ESK] indefinite contexts. However, if their article choice depends on ±ESK), overuse of the should be found in only [+spec; +ESK], not in [+spec; -ESK] contexts.

4 The present study
4.1 Participant groups

Proficiency-matched intermediate and advanced English speakers from two language backgrounds were compared: two L1 Thai (-article) and two L1 French (+article) groups. There were 20 participants per group. The experiment was conducted in Thailand. The Thai participants’ age range was 16;9 to 19;6 years (mean = 18;2) and the French participants’ age range was 17;2 to 19;3 years (mean = 18;2). The levels of English proficiency were determined by the Oxford Placement Test (Allen 2004).

4.2 Materials & Procedures

The participants were tested on the forced-choice elicitation task materials, a semi-replication of IKW’s (2004) materials. They were a discrete-item test, consisting of 24 items. Each item contained a short English dialogue based on items from the forced-choice elicitation task in IKW. In the modified version here, the introduced contexts were those where the two features were separated, i.e. [+spec; -ESK]. This was done for both [+def] and [-def] contexts. To test whether [+spec] or [±ESK] influenced the L2 speakers’ article choice, 3 combinations of [±spec] and [±ESK] values were made in each (in)definite context, yielding 6 contexts: 3 [-def] and 3 [+def] equivalents:

<table>
<thead>
<tr>
<th></th>
<th>[-definite]</th>
<th>[±definite]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+spec; +ESK]</td>
<td>[±spec; +ESK]</td>
<td>[+spec; -ESK]</td>
</tr>
<tr>
<td>[-spec; -ESK]</td>
<td>[-spec; -ESK]</td>
<td>[+spec; -ESK]</td>
</tr>
</tbody>
</table>

Table 4
Indefinite & definite equivalent contexts

There were four items per context type. The test items were arranged in a random order. Based on the context in each dialogue, the participants were asked to fill in a blank in front of the target singular nominal phrase in each item with the, a(n) or leave the blank empty if they decided that no article should be used.

4.3 Predictions

If L2 speakers associate article choice with [±spec](based on the ACP and FH), then:

<table>
<thead>
<tr>
<th>CONTEXTS</th>
<th>+spec; +ESK</th>
<th>-spec; -ESK</th>
<th>+spec; -ESK</th>
</tr>
</thead>
<tbody>
<tr>
<td>+definite(target: the)</td>
<td>correct use of the</td>
<td>overuse of a(n)</td>
<td>correct use of the</td>
</tr>
<tr>
<td>-definite(target:a(n))</td>
<td>overuse of the</td>
<td>correct use of a(n)</td>
<td>overuse of the</td>
</tr>
</tbody>
</table>

Table 5
Predicted English article choices if article choice is influenced by [±spec]

If L2 speakers’ article choice depends on [±ESK], then:

<table>
<thead>
<tr>
<th>CONTEXTS</th>
<th>+spec; +ESK</th>
<th>-spec; -ESK</th>
<th>+spec; -ESK</th>
</tr>
</thead>
<tbody>
<tr>
<td>+definite(target: the)</td>
<td>correct use of the</td>
<td>overuse of a(n)</td>
<td>overuse of a(n)</td>
</tr>
<tr>
<td>-definite(target:a(n))</td>
<td>overuse of the</td>
<td>correct use of a(n)</td>
<td>correct use of a(n)</td>
</tr>
</tbody>
</table>

Table 6
Predicted English article choices if article choice is influenced by [±ESK]

5 RESULTS

1 Logically, it is impossible to have a combination of [-spec] and [+ESK].
5.1 Results from the Thai groups

5.1.1 The intermediate Thai group

<table>
<thead>
<tr>
<th></th>
<th>[+spec; +ESK]</th>
<th>[-spec; -ESK]</th>
<th>[+spec; -ESK]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-definite] (target: a(n))</td>
<td>28.75% the 68.75% a(n) 2.5% Ø</td>
<td>7.5% the 86.25% a(n) 6.25% Ø</td>
<td>10% the 82.5% a(n) 7.5% Ø</td>
</tr>
<tr>
<td>[+definite] (target: the)</td>
<td>83.75% the 11.25% a(n) 5% Ø</td>
<td>51.25% the 35% a(n) 13.75% Ø</td>
<td>55% the 38.75% a(n) 6.25% Ø</td>
</tr>
</tbody>
</table>

Table 7

English article substitutions by the intermediate Thai L2 learner group \((N = 20)\)

To determine the significance of the contribution of the combination of [spec] and [ESK] values to the use of the vs. \(a(n)\), a repeated-measures ANOVA was performed on the overuse of the in indefinite, and the overuse of \(a(n)\) in definite environments by context type. [spec; ESK] had a highly significant effect on (over)use of both the in [-def] and \(a(n)\) in [+def] contexts.

- Overuse of the in [-def], \(F(2, 38) = 7.924, p < .01\). Contrasts:
  - Overuse of the in [+spec; +ESK] vs. [-spec; -ESK], \(F(1, 19) = 10.341, p < .01, r = .77\)
  - Overuse of the in [+spec; +ESK] vs. [+spec; -ESK], \(F(1,19) = 8.301, p < .05, r = .74\)

- Overuse of \(a(n)\) in [-def], \(F(2, 38) = 20.675, p < .001\). Contrasts:
  - Overuse of \(a(n)\) in [+spec; +ESK] vs. [-spec; -ESK], \(F(1, 19) = 47.50, p < .001, r = .93\)
  - Overuse of \(a(n)\) in [+spec; +ESK] vs. [+spec; -ESK], \(F(1,19) = 27.092, p < .001, r = .89\)

5.1.2 The advanced Thai group

<table>
<thead>
<tr>
<th></th>
<th>[+spec; +ESK]</th>
<th>[-spec; -ESK]</th>
<th>[+spec; -ESK]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-definite] (target: a(n))</td>
<td>18.75% the 80% a(n) 1.25% Ø</td>
<td>0% the 95% a(n) 5% Ø</td>
<td>2.5% the 96.25% a(n) 1.25% Ø</td>
</tr>
<tr>
<td>[+definite] (target: the)</td>
<td>96.25% the 3.75% a(n) 0% Ø</td>
<td>68.75% the 22.5% a(n) 8.75% Ø</td>
<td>75% the 22.5% a(n) 2.5% Ø</td>
</tr>
</tbody>
</table>

Table 8

English article substitutions by the advanced Thai L2 learner group \((N = 20)\)

- Overuse of the in [-def], \(F(1.17, 22.16) = 11.779, p < .01\). Contrasts:
  - Overuse of the in [+spec; +ESK] vs. [-spec; -ESK], \(F(1, 19) = 15.545, p < .01, r = .83\)
  - Overuse of \(a(n)\) in [+spec; +ESK] vs. [+spec; -ESK], \(F(1, 19) = 9.701, p < .01, r = .76\)

- Overuse of \(a(n)\) in [+def], \(F(2, 38) = 6.151, p < .01\). Contrasts:
  - Overuse of \(a(n)\) in [+spec; +ESK] vs. [-spec; -ESK], \(F(1, 19) = 9.0, p < .01, r = .75\)
  - Overuse of \(a(n)\) in [+spec; +ESK] vs. [+spec; -ESK], \(F(1, 19) = 15.545, p < .01, r = .83\)

Although the advanced Thai group produced fewer article substitution errors, what is important to note is that the same pattern of the substitutions persisted.
In sum, the statistical results from the Thai groups show that, in the contexts directly taken over from IKW (2004), the results from IKW have been replicated. However, in the new contexts which tease apart [spec] and [ESK], the predictions of the FH were falsified: In [-def] contexts, the was overused in [+spec] contexts ONLY when [+ESK]. It was not overused in [+spec] contexts where [-ESK]. In [+def] contexts, a(n) was overused in ALL contexts where [-ESK], irrespective of whether the context was [+/-spec]. So, the results falsify the claim that overuse of the is influenced by [±spec]. They supported the prediction that, on this test, L2 speakers would make article distinctions based on [±ESK]. [±spec] seemed to play no role in their article choice.

5.2 Results from the French groups

5.2.1 The intermediate French group

<table>
<thead>
<tr>
<th></th>
<th>[+spec; +ESK]</th>
<th>[spec; -ESK]</th>
<th>[+spec; -ESK]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-definite]</td>
<td>2.5% the</td>
<td>0% the</td>
<td>1.25% the</td>
</tr>
<tr>
<td>(target: a(n))</td>
<td>95% a(n)</td>
<td>95% a(n)</td>
<td>97.5% a(n)</td>
</tr>
<tr>
<td></td>
<td>2.5% Ø</td>
<td>5% Ø</td>
<td>1.25% Ø</td>
</tr>
<tr>
<td>[+definite]</td>
<td>92.5% the</td>
<td>83.75% the</td>
<td>90% the</td>
</tr>
<tr>
<td>(target: the)</td>
<td>3.75% a(n)</td>
<td>11.25% a(n)</td>
<td>6.25% a(n)</td>
</tr>
<tr>
<td></td>
<td>3.75% Ø</td>
<td>5% Ø</td>
<td>3.75% Ø</td>
</tr>
</tbody>
</table>

Table 9

English article substitutions by the intermediate French L2 learner group (N = 20)

A repeated-measures ANOVA was performed on the overuse of the in indefinite environments, and the overuse of a(n) in definite environments by context type. [spec; ESK] had a non-significant effect on (over)use of the or a(n):

- Overuse of the in [-def] contexts, F (1.46, 27.70) = 1.000, p > .05.
- Overuse of a(n) in [+def] contexts, F (1.50, 28.33) = 1.956, p > .05

5.2.2 The advanced French group

<table>
<thead>
<tr>
<th></th>
<th>[+spec; +ESK]</th>
<th>[spec; -ESK]</th>
<th>[+spec; -ESK]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-definite]</td>
<td>1.25% the</td>
<td>0% the</td>
<td>0% the</td>
</tr>
<tr>
<td>(target: a(n))</td>
<td>97.5% a(n)</td>
<td>100% a(n)</td>
<td>97.5% a(n)</td>
</tr>
<tr>
<td></td>
<td>1.25% Ø</td>
<td>0% Ø</td>
<td>2.5% Ø</td>
</tr>
<tr>
<td>[+definite]</td>
<td>96.25% the</td>
<td>92.5% the</td>
<td>97.5% the</td>
</tr>
<tr>
<td>(target: the)</td>
<td>2.5% a(n)</td>
<td>6.25% a(n)</td>
<td>2.5% a(n)</td>
</tr>
<tr>
<td></td>
<td>1.25% Ø</td>
<td>1.25% Ø</td>
<td>0% Ø</td>
</tr>
</tbody>
</table>

Table 10

English article substitutions by the advanced French L2 learner group (N = 20)

- Overuse of the in [-def] contexts, F (1.55, 29.46) = .487, p > .05.
- Overuse of a(n) in [+def] contexts, F (2, 38) = 1.132, p > .05.

The results from the French groups indicate that there were no differences in the overuse of a(n) in any [+def] contexts and the in any [-def] contexts. So, the French speakers’ article choices were not influenced by either the value of [spec] or [ESK].
6 DISCUSSION

The results were shown to falsify the FH. There was no evidence indicating that specificity plays a part in L2 article choices. Overuse of the and a(n) was not found to be tied to [+spec] and [-spec], respectively. The same pattern of results was found in Trenkic (2007b) for L1 Mandarin/L2 English speakers. The results can therefore be said to undermine the validity of the construct of the ACP, and definiteness as a UG-based semantic feature. The outcome is in line with an alternative account – The Syntactic Misanalysis Hypothesis (SMH) - assuming that L2 articles are misanalysed as nominal modifiers (c.f. Trenkic 2007a).

According to the SMH, languages without articles do not have a syntactic category determiner in their grammar (cf. Lyons 1999). Determiner-like elements are argued to behave syntactically as adjectives. Due to the adjectival nature of determiner-like elements in such languages, L2 speakers from these languages are assumed to incorrectly treat L2 determiners, including articles, as adjectives (cf. Kuribara 1999; Trenkic 2007a). The fact that L2 articles are misanalysed as nominal modifiers leads to certain implications for L2 article production:

- An article will only be produced if the learner sees a communicative reason for its meaning to be expressed in the given context, resulting in frequent omissions.
- If ‘correct meanings’ have been assigned to articles, when articles are supplied, there will be no substitution errors. If not, principled substitution errors will occur, reflecting the meanings the learners have assigned to the article forms.

So, what meanings might L2 speakers assign to L2 article forms? The proposal is that, as articles are assumed to be syntactically misanalysed as adjectives, then referential, common-sense meanings of definite (‘that can be identified’) and indefinite (‘that cannot be identified’) would be attributed to the and a(n), respectively.

A discourse referent is definite if the speaker intends to refer to it, and expects it to be identifiable to the hearer. But discourse-related identifiability does not depend on either the speaker’s or the hearer’s ability to determine the ‘real-world identity’ of discourse referents (cf. Trenkic 2007b). For example,

(3) (a). Macbeth was written by a famous English playwright.
   (b). We are looking for the vandals who broke into the office last night. (Trenkic 2007b)

Although the real-world identity of a famous English playwright may be known to both the speaker and the hearer, the cannot be used. In contrast, even though neither the speaker nor the hearer might be able to identify the vandals who broke into the office last night on an identity parade, the can be employed. This is because definiteness as a discourse-related identifiability is linked to the referents’ ‘existence’ and ‘uniqueness’ in discourse-determined pragmatically delimited contexts (cf. Hawkins 1991), and does not depend on the ability of the participants in discourse to determine the real-world identity of these referents.

The adjectival (lexical) meaning of definite and indefinite, however, concerns more readily the ‘objective identifiability’ of referents, which is a broader concept than ‘discourse identifiability’. Objective identifiability can be determined against a wider range of criteria. One criterion which is salient in the test is ‘the speaker’s familiarity with the referent’ via identifying attributes. When the criteria converge, i.e. [-def; -ESK] and [+def; +ESK], correct article choices occurred. However, when there is a conflict between different criteria, i.e. [-def; +ESK] and [+def; -ESK], fluctuations in article choice were found.

The fluctuation was found only in L1 Thai/L2 English speakers’ production, but not in L1 French/L2 English speakers’ choices, suggesting that the former, but not the latter misanalyse L2 articles as adjectives. Given the presence of the syntactic category determiner
in L1 French, French learners of English are assumed to analyse English determiners, including articles, appropriately as determiners.

7 Conclusion

The results from the Thai groups patterned along the dimension of [ESK], not [spec] whereas those from the French groups consistently patterned along the line of [def]. I have argued that L2 speakers from (-article) languages analyse English articles as adjectives, attributing them the lexical, common-sense meanings of definite and indefinite, and that this leads them occasionally to link article choices with [±ESK], i.e. the speaker’s explicit claim or denial of his familiarity with the person or object being talked about. The FH and the ACP were falsified, suggesting that L2 article choices are not UG-based.

References


Nattama Pongpairoj

Department of Language and Linguistic Science
University of York
Heslington Road
York
YO10 5DD
United Kingdom

np500@york.ac.uk