Acquiring the polysemous adverb HAI in Chinese by English-speaking, Japanese-speaking, and Korean-speaking CSL learners

Miao-Ling Hsieh [謝妙玲] and Yu-Fang Wang [王萸芳]
National Taiwan Normal University [國立臺灣師範大學]
National Kaohsiung Normal University [國立高雄師範大學]

Using corpus analysis and error analysis, this study investigates English-speaking, Japanese-speaking, and Korean-speaking Chinese learners’ acquisition of various meanings of hai in Mandarin Chinese, including its temporal meaning ‘still, yet’ as well as its abundant atemporal meanings involving addition, comparison and counter-expectation. We found a preponderance of misselection errors across the three groups of learners. The next most common error type for the Japanese-speaking and Korean-speaking learners was omission, while omission and over-inclusion were equally challenging for the English-speaking learners. Further analysis of errors in misselection shows that many learners failed to distinguish the temporal hai from the atemporal haishi required in a concessive sentence for the counter-expectation meaning.

Keywords: Chinese adverb hai, Chinese as a second language, learner corpus, error analysis, second language acquisition

關鍵詞：漢語副詞「還」、對外漢語、中介語語料庫、錯誤分析、二語習得
1. Introduction

The Chinese adverb *hai* is a polysemous word. *Hai* by itself functions as an adverb, occurring in a preverbal position after the subject, giving rise to a temporal reading ‘still’ or an additive reading ‘also’. As an adverb, it can also mean ‘moderately’ typically when it occurs with *suan* ‘considered’ or mean ‘in a higher degree’ in a comparative sentence. It can also combine with the copula verb *shi* ‘be’ forming *haishi*, which is used as an adverb in concessive sentences conveying a counter-expectation meaning. *Haishi* can also occur as a conjunction in a disjunctive question and in sentences marking a choice or a suggestion. The former has an alternative reading ‘or’, while the latter an evaluative reading. In this study, we emphasize the concessive *haishi*, for it bears some similarities with *hai* and can be translated as ‘still’ in English. Finally, *hai* can also occur with the existential verb *you* ‘have’ forming *haiyou* ‘and also’, which is used as a conjunction. More discussion of the various meanings and uses of *hai* is given in Section 2.

Because of its multiple meanings, a few studies in L2 have been done on how learners of Chinese acquire the various meanings and uses of *hai* (e.g., L. Li 2009, Aosakul 2011, Y. Li 2011, HY. Liu 2011, Mao 2011, Jing 2012, Jiang 2013). L. Li (2009), for example, compared *hai* used by Oriental (Japanese and Korean) learners and Western (American and European) learners. She found that the multiple uses and meanings of *hai* caused difficulties for some learners. Specifically, the Japanese and Korean learners used the additive *hai* more often, while the American and European learners used the temporal *hai* more frequently. According to L. Li, this may be due to L1 transfer; that is, Japanese learners being influenced by *mada* (‘still’) and American learners by *still* and *yet*. But as we will show in Section 4.3, while the English *still* has an additive meaning, the Japanese *mada* and Korean *ajig* can only mean ‘still, yet’. It is thus not clear how L1 transfer was involved. In a more recent study, Jiang (2013) analyzed the uses of *hai* by Japanese learners of Chinese in the HSK Dynamic Composition Corpus and from sentences his subjects created in a designed questionnaire. Classifying the meanings of *hai* into eight types, he aimed to sort out the acquisition order of *hai* by the Japanese subjects. He found that error rates at the primary intermediate and advanced levels were not high (5.4% for the former and 4.6% for the latter). Among the eight

---

1. L. Li (2009) observed that instead of using the additive *hai*, American learners tended to use *ye* for the additive meaning, which is also translated as ‘also’ in English. However, the two differ. When the addition meaning involves a different grammatical subject, *ye* but not *hai* can be used. Please see our discussion of the two in terms of focus in Section 2.2.

2. HSK (Hanyu Shuiping Kaoshi) is a standardized Chinese proficiency test for CSL/CFL learners in China.
types of *hai* in Jiang (2013), the first four types were equal to the temporal *hai*, additive *hai*, comparative *hai*, and moderate *hai* categorized in this paper. His fifth and sixth were what he called less-than-expected *hai* and earlier-or-faster-than-expected *hai*. The former marks a small amount, while the latter earliness/faster-ness. Jiang’s eighth type corresponds to the counter-expectation *hai* in this paper. It is not clear what Jiang’s seventh type refers to, though. Jiang commented that this use was rare and gave only one example that involves *lian…dou hai mei…‘even… all hai not …’* This might again correspond to our use of the temporal *hai*. Because Jiang did not include the use of *haishi*, the error rates were found to be not high.

It should be pointed out that most previous studies, if not all, have only contained a small sample size; the research findings are thus limited. Different from the above studies, which examined CSL learners’ acquisition of *hai*, Y. Wu (2010) looked at a large-scale corpus of Chinese learners’ uses of *haishi* in terms of its different grammatical meanings and semantic contexts, from the basic function of *haishi* for time continuation to a series of other functions such as the conjunctive function of choosing the modality expression of choice and unexpectedness along with its more complicated grammatical meanings and semantic contexts. Y. Wu’s study, however, did not consider the subjects’ L1.

This paper concerns the second language acquisition of the Mandarin Chinese *hai*, including its synonym *haishi*, and aims to provide a detailed description of English, Japanese, and Korean CSL learners’ various uses of *hai* in writing. We attempt to compare the three groups of learners’ uses of *hai* by examining the errors concerning *hai* that they make. It is hoped that the results will shed light on the field of Chinese teaching and offer some pedagogical implications. The rest of the paper is organized as follows. Section 2 includes the various uses of *hai* as well as *haishi* for the reason that *haishi*’s meaning is related to *hai*. Section 3 discusses the data collection and classification. It is followed by the results and discussion in Section 4. Section 5 concludes the paper.

2. **Various uses of *hai***

The adverb *hai* has many meanings. Previous literature has dealt with its various uses and their derivations (Chao 1968, Lü 1980, Li & Thompson 1981, Yeh 1998, F. Liu 2000, to mention just a few). For example, Chao (1968) argues that the adverb *hai* has three different uses: as an adverb of time, an adverb of degree, and an adverb of evaluation. Similarly, Li & Thompson (1981) observe that *hai* has three meanings: ‘still/even’, ‘also’, and ‘moderately’ (namely, Chao’s (1968) adverb of evaluation). These three major uses/meanings, however, will be shown to be
too broad for an examination of hai as used by learners. In this section, we divide the uses of hai into five categories instead: temporal hai, additive hai, comparative hai, moderate hai, and counter-expectation hai. Both comparative hai and moderate hai involve a comparison. We discuss each of them and also include the meanings of haishi 'hai + copula' to clarify the meaning differences between hai and haishi. Finally, we will also briefly touch upon the literature on the derivations of these various uses.

2.1 Temporal hai

First, hai can act as an adverb of time, signaling that a previous state continues or does not change, as in (1) and (2):

(1) Ta hai nianqing.
   he HAI young
   ‘He is still young.’

(2) a. Women hai bu zhidao.5
    1PL HAI not know
    ‘We don’t know yet.’
    (Li & Thompson 1981: 334)

   b. Tamen hai mei lai.
      3PL HAI not come
      ‘They haven’t been here yet.’

In (1), the temporal hai modifies the stative verb nianqing ‘young.’ In (2), it occurs with the negative markers bu and mei. In both (1) and (2), it is presupposed that the state denoted by the stative verb or the negated situation has existed or happened before the reference time and it is asserted that the state or negated situation continues to be true at the reference time.

2.2 Additive hai

In addition to the temporal meaning, Li & Thompson (1981) point out that hai has an additive meaning ‘also’, which is used to focus on the predicate, not the sub-

---

3. In natural language, scalar expressions such as comparitives, intensifiers, and minimizers are not only used for measuring an object or event at a semantic level, but also for expressing various kinds of pragmatic information, including politeness, the speaker’s attitude, unexpectedness, and the like. Both the comparative hai and moderate hai have a scalar use in pragmatics. Please see the discussion below.

4. For all the examples in Chinese characters, please see Appendix at the end of this article.

5. The list of abbreviations is available at the end of this article.
ject of the sentence, as in (3a). This use of hai is different from the adverb ye ‘also’, which focuses on the subject, not the predicate, as shown in (3b).

(3) a. Ta hai mai-le yi ge huaping.
   3SG HAI buy-ASP one CL vase
   ‘S/He also bought a vase (in addition to doing some other things).
   (Li & Thompson 1981:334)

   b. Ta ye mai-le yi ge huaping.
   3SG also buy-ASP one CL vase
   ‘S/He (in addition to some other people) also bought a vase.’
   (Li & Thompson 1981:334)

Moreover, the additive use of hai introduces an additional event. Crucially, with such a usage, additional quantity is involved. This use of hai can thus be distinguished from the temporal hai, which as mentioned above, is used to mark the continuance of a state or a negated situation. When the sentence that hai is in involves an event instead of a state or a negated situation, the additive meaning arises, yielding the meaning of accumulation in amount or quantity, as in (3a), or simply the addition of another event, as in (4), adapted from Hou (2004:256):

(4) Zai Hangzhou youlan-le Xihu, hai/you pinchang-le naer de xiaochi.
    in Hangzhou visit-ASP West.Lake HAI/also taste-ASP there DE snack
    ‘In Hangzhou, (we) visited West Lake and tasted the snack there.’
    (Hou 2004:256)

This use of hai is, however, restricted to a series of events that happened in the past on the same occasion. In this way, hai is different from you ‘also’. In (4), hai and you can be interchangeable. However, when a series of events occur over different time spans, you but not hai can be used, as in (5):

(5) Zhuotian ta qu-le Changcheng, jintian you/*hai qu-le Yiheyuan.
    yesterday 3SG go-ASP Great.Wall today also/HAi go-ASP Summer.Palace
    ‘Today he went to the Great Wall and also the Summer Palace.’ (Hou 2004:256)

Examples (4) and (5) show that both hai and you may denote addition when referring to past events; however, hai cannot be used to denote addition across time spans, in contrast to you, which is not subject to this restriction (Ma 1999, Jing-Schmidt & Gries 2009).

Noticeably, the additive hai is also possible in a sentence that denotes a state and it often occurs with conjunctions such as bujin/budan ‘not only’ or chule ‘besides’, as illustrated in (6).
Both *congming* ‘smart’ and *yonggong* ‘hardworking’ denote a state. Without the preceding clause, the second clause in (6) would be more likely be interpreted as involving a temporal *hai*, indicating that a state is maintained, as shown in (7):

(7) Ta ()*hai* *hen* *yonggong.*

3SG HAI very hardworking

‘He is still hardworking.’

The use of *bujin/budan* can help us identify this use of *hai* in the corpus.

Finally, the additive meaning is also available when *shenzhi* ‘even’ appears, as in (8):

(8) Ta *meitian* *lianxi, shenzhi hai* *hen* *yonggong.*

3SG every day practice even HAI very dilegent

‘He practices every day and is even very dilegent.’

(8) means the dilegent state is an additional condition. Both conditions ‘practicing everyday’ and ‘hardworking’ can, for example, make him very competitive. Moreover, the latter condition is less expected than the former one, as indicated by *shenzhi* ‘even’. This function of the additive adverb *hai* is metalinguistic and subjective in the sense that it adds information to an existing proposition that the speaker, in the context, considers not informative enough (Shen 2001).

---

6. One of the reviewers points out that the additive *hai* does not necessarily co-occur with *bujin* ‘not only’, for the latter may pair with *erqie* ‘and’. We agree with him/her, but while *hai* can be used with *bujin* in the presence of *shenzhi* ‘even’, *erqie* cannot:

(i) Ta *bujin meitian lianxi, shenzhi hai* *hen* *yonggong.*

3SG not only every day practice even HAI very hardworking

‘He practices every day. Besides, he is even very hardworking.’

(ii) *Ta bujin meitian lianxi, erqie shenzhi hai* *hen* *yonggong.*

3SG not only every day practice and even HAI very hardworking

‘He practices every day. Besides, he is even very hardworking.’

The contrast between (i) and (ii) shows that compared to *erqie, hai* is more closely related to *bujin.*
2.3 Comparative hai

As mentioned above, the temporal *hai*, which involves a state or a negated situation that continues from before the reference time to the reference time, can invite the inference of accumulation in amount or quantity, hence of addition. The atemporal additive meaning can also be extended to increase in degrees. This is indeed the case when *hai* is used in a comparative sentence. Consider the following example.

(9) Eryong bi ta gege Dayong hai zhuang.
    ‘Eryong is somewhat/much stronger than his older brother Dayong.’
    (Lü 1980: 252)

It may not be immediately clear how an increase in degrees in (9) is involved, but once we add quantity phrases like *yidian* ‘a little’ or *henduo* ‘a lot’, as shown in (10), it becomes clear that this use of *hai* is an extension of the additive *hai*. But since it involves a scalar interpretation in comparison, we will call it ‘scalar *hai’; or more specifically ‘comparative *hai’.

(10) .... hai zhuang yidian/henduo.
    HAI strong a.little/a.lot
    ‘...somewhat stronger/much stronger.’

2.4 Moderate hai

The use of *hai* in comparative sentences involves an explicit target of comparison introduced by *bi* ‘than’. Another scalar use of *hai*, however, introduces a comparison between a state and an implicit standard assumed by the speakers. For example:

(11) Zhei ge banfa hai keyi.
    this CL method HAI fine
    ‘This method is fairly good.’
    (Li & Thompson 1981: 335)

(12) A: Ni zhu de difang zenmeyang?
    2SG live DE place how
    ‘How is the place (where) you’re living?’

    B: Hai hao.
    HAI good
    ‘It’s not bad.’
    (Li & Thompson 1981: 335)
The speaker in (11) states that the method he refers to was good compared to the standard degree contributed by the scalar predicate. We will call this ‘moderate hai’. Similarly, hai hao in (12), composed of hai and the stative verb hao (‘good’), is common in daily conversation (for a discussion on hai hao, see Biq 2004). In particular, the expressions hai hao, hai bucuo, and the like convey the speaker’s subjectivity and fall into the category of scalar hai.

Crucially, for this use of hai, the stative verb does not occur with a degree adverb and in many cases suan ‘considered’ is required, which is used to indicate the speaker’s point of view. Compare these two examples:

(13) Lifen hai hen piaoliang.
    Lifen HAi very pretty
    ‘Lifen is still very pretty.’

(14) Lifen hai ?? (suan) piaoliang.
    Lifen HAi considered pretty
    ‘Lifen is fairly pretty.’

The former hai has the temporal reading, while the latter the moderate reading.

2.5 Counter-expectation (CE) hai

Counter-expectation (CE) hai is used when something takes place despite the speaker’s expectation to the contrary, as in (15):

(15) Ni shengbing le, (juran) hai gen ta chuqu?
    2SG sick PRT unexpectedly HAi with 3SG go.out
    ‘You were sick and yet you would go out with him?’

The speaker in (15) was very surprised as indicated by the use of juran ‘unexpectedly’ that the listener went out with somebody despite being sick. This use of hai is not temporal because it does not mean that a state or a negated situation con-

This use of hai may be marked in the sense that some predicates do not require suan, but others seem to, as shown in (i):

(i) Laowang de wuzi hai ganjing.
    Laowang DE room HAi clean
    ‘Laowang’s room is moderately clean.’

The structure which contains suan ‘considered’, such as the one in (14) in the text, is considered a special case of comparative construction (one in which the comparison standard is left implicit) by F. Liu (2000) and Yang (2017), who term it ‘marginal hai’, and by Donazzan (2008), who terms it ‘borderline hai’.

---

7. This use of hai may be marked in the sense that some predicates do not require suan, but others seem to, as shown in (i):
tinues. In such usage, *hai* highlights the discrepancy that the speaker sees between the actuality and his/her expectation (F. Liu 2000, G. Wu 2009).

CE *hai* can also be used for emphasis only, as seen in (16), taken from the textbook *Practical Audio-Visual Chinese* Vol. 3, Lesson 12:

(16) Ta de meili  **hai** zhen bu xiao.
    he *DE charm* **hai** truly not small
    ‘He is actually very charming.’

Crucially, for this use of *hai*, a degree adverb marking a high degree is always used (e.g., *zhen* ‘truly’ in (16)).

It should be pointed out that *hai*, when combined with the copula verb *shi*, can also indicate the speaker’s subjectivity. According to Li & Thompson (1981), when *hai* acts as an adverb of time (or anaspectual particle), it may optionally occur with the copula verb *shi*, as illustrated in (17):

(17) Ta **hai**(shi) xihuan Lisi.
    3SG **hai**(shi) like Lisi
    ‘S/he still likes Lisi.’

(Li & Thompson 1981: 334)

It is, nevertheless, not true that this use of *haishi* occurs freely with *hai*. As shown in the example below, only a clause containing *haishi*, but not *hai*, can follow a clause introduced by the concession pattern *suiran*..., *keshi*... ‘although..., yet...’, marking a contrast in meaning:

(18) Suiran  Lisi bu li ta, keshi ta **haishi/*hai** xihuan Lisi.
    although Lisi not pay.attention.to 3SG but 3SG **HAISHI/HAI** like Lisi
    ‘Although Lisi doesn’t pay attention to him, he still likes Lisi.’

Similarly, in (19) *haishi* must be used, because none of the uses of *hai* is possible.

(19) Jinguan ta bu congming, ta **haishi/*hai** hen yonggong.
    although 3SG not smart 3SG **HAISHI/HAI** very diligent
    ‘Although he is not smart, he is still very diligent.’

In contrast, when the temporal *hai* is possible, *haishi* does not have to be used.

(20) Xianzai hen wan le, keshi ta de jingshen **hai** hen hao.
    now very late PRT but he *DE spirit* **hai** very good
    ‘Now it is late, but he is still full of energy.’

---

We will take the use of *haishi* in (18) and (19) to be concessive and call it ‘concessive *haishi*’ for the sake of simplicity. This use of *haishi* is an adverb because it cannot be used before a subject, as shown in (21).

(21) ... *haishi* ta xihuan Lisi.
... *HAISHI* 3SG like Lisi
‘... still he likes Lisi.’

The concessive pattern *suiran...*, *(keshi/danshi/buguo)....* ‘although..., (yet)...’ will be important for us in identifying this use. But crucially only those instances that are not temporal belong to this category. Much research has pointed out that the notions of concession and contrast are linked to expectation. For example, Quirk et al. (1985) point out that “[T]he contrast may be in the unexpectedness of what is said in the second conjoin in view of the content of the first conjoin” (p.935) and in connection with concessive clauses they hold that they “indicate that the situation in the matrix clause is contrary to expectation in light of what is said in the concessive clause” (p.1098). Specifically, information is presented in the matrix clauses in (18) and (19) containing *haishi* as an unexpected result or consequence ensuing from the situation presented in the concessive clauses introduced by *suiran* and *jinguan*, separately. *Haishi* imparts a strong contrast, while *hai* (see (20)) indicates a weaker contrast.\(^9\) That is, the degree of contrast is stronger with *haishi*. In particular, when the temporality is downplayed or the situation is atemporal, the contrast between expectation and reality is accentuated; hence, *haishi* instead of *hai* is preferred (see (18)–(19)).

### 2.6 Derivations of various uses of *hai*

It is in no doubt from the discussion above that some uses of *hai* seem closely related and all of its uses may be argued to be derived from one source. Indeed, many studies have suggested this in one way or another.

From diachronic and synchronic perspectives, Yeh (1998) identifies two major functions of *hai*, i.e., a temporal operator meaning ‘still’ and a connective marker meaning ‘also’, which developed into other uses around the 7th–8th century. Specifically, *hai* as a connective marker provides additional confirmation of a conclusion derived from the preceding discourse. The function of *hai* which originally indicates temporal continuance is analyzed as a scalar operator, which reflects the speaker’s relative ranking of the alternatives within a set of alternatives. Hence, it relates the value of the focused expression to a set of alternatives. Yeh further argues that the scalar analysis can be extended to account for the other

---

\(^9\) Note there is a contrastive marker *keshi* in (20), which cannot be deleted.
uses of *hai*: repetitive, moderate, scalar, concessive and CE. Her repetitive is our additive.

On the other hand, F. Liu (2000) holds that the possible meanings of *hai* include 'still', 'again', 'also' and 'even'. Analyzing *hai* as a polysemous word with an additive presupposition, paired with a scalar implication, she suggests that all occurrences of *hai* have a basic meaning, i.e., 'without interruption' and that it evokes a relation between two propositions to be evaluated in a scalar model. To be more specific, as a focus particle, *hai* is scalar and additive; it is associated with a higher value, and its occurrence in a sentence entails another proposition already existing in the context, one associated with a lower value. This analysis will work for a comparative *hai*, but not a moderate *hai*.

On the other hand, both HJ. Liu (2009) and Yang (2017) treat *hai* as an additive particle, but differ in explaining how the scalar meaning is derived. HJ. Liu (2009) maintains that *hai* is related to an increment in a scale, either by adding or by comparison. This analysis, just like F. Liu’s (2000) analysis mentioned above, will only work for a comparative *hai*, not a moderate *hai*. In a looser way in relating the additive and scalar meaning, Yang classifies *hai* into (i) temporal, (ii) additive, (iii) comparative (explicit scalar), (iv) marginal (implicit scalar), (v) concessive, and (vi) counter-expectation (CE) meanings and argues that whether or not it implies a scalar sense depends on whether the context induces an order effect.

Finally, capitalizing on the use of *hai* in so-called implicit comparatives, Donazzan (2008) argues that *hai* contributes to an evaluative reading subjective to the speaker and behaves like a scalar particle. *Hai* can induce an ordering or scale into the interpretation of a sentence that contains it and is both an additive and a scalar particle. Put differently, the additive *hai* can be a scalar focus modifier that invokes alternatives on a scale but also ranks the focus of that scale, usually with other similar elements (e.g., qualities, events, persons, etc.). This ranking often signals a specific discourse evaluation, such as expectation. To be specific, it serves as a marker of mood indicating the speaker’s stance (Biq 2004), subjectivity or evaluation. Donazzan’s analysis, however, does not distinguish the concessive *haishi*, which requires *shi*, and the moderate *hai*, which does not require *shi*. Furthermore, the temporal *hai* describes a state or a negated situation, but it does not have to be scalar.

---

10. In linguistics, stance is viewed as an individual speaker’s expression of his/her internal psychological state and is essentially synonymous with concepts such as affect, evaluation, subjectivity, footing, alignment, assessment, etc. (For details, see Goffman 1967, Biber & Finegan 1988, 1989, Traugott 1995, Traugott & Dasher 2002, Biber 2004.)
Overall, previous studies on *hai* have focused on its syntactic, semantic, and pragmatic uses. While some propose a core-meaning for all its uses (e.g., Yeh 1998, F. Liu 2000, Donazzan 2008, among others), others pursue the idea that *hai* has various meanings (e.g., Li & Thompson 1981, Yang 2017). But none of them seem to include all the uses of *hai* plus its synonym *haishi*. Without aiming at deriving the various meanings, we simply conclude that *hai* can express temporality, objective quantity, and subjective quantity, or even evaluation, namely, temporal *hai*, additive *hai*, comparative *hai*, moderate *hai*, and concessive *haishi*. Since *hai* has many meanings involving temporality, quantity and speaker’s subjectivity, we wonder whether learners of Chinese can distinguish between the various meanings of *hai* and use them appropriately.

3. Data collection and classification

3.1 The data

The methods used in this study include corpus analysis and error analysis (Corder 1967). Our data came from the Learner Corpus – Test of Chinese as a Foreign Language (Learner Corpus – TOCFL). This corpus was compiled from online proficiency tests administrated by the Steering Committee for the Test of Proficiency-Huayu at National Taiwan Normal University. The tests were given at four levels (A2, B1, B2 and C1) according to the Common European Framework of Reference (CEFR) – with A2 the lowest and C1 the highest level, B1 and B2 the levels in between. 452 English-speaking learners, 796 Japanese-speaking learners and 283 Korean-speaking learners of Chinese participated in this study. All produced the same number of tasks, as summarized in Table 1.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Proficiency levels</th>
<th>A2 No. of texts/learners</th>
<th>B1 No. of texts/learners</th>
<th>B2 No. of texts/learners</th>
<th>C1 No. of texts/learners</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>English-speaking</td>
<td></td>
<td>123</td>
<td>252</td>
<td>59</td>
<td>18</td>
<td>452</td>
</tr>
<tr>
<td>Japanese-speaking</td>
<td></td>
<td>256</td>
<td>402</td>
<td>125</td>
<td>13</td>
<td>796</td>
</tr>
<tr>
<td>Korean-speaking</td>
<td></td>
<td>74</td>
<td>144</td>
<td>57</td>
<td>8</td>
<td>283</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>453</strong></td>
<td><strong>798</strong></td>
<td><strong>241</strong></td>
<td><strong>39</strong></td>
<td><strong>1531</strong></td>
</tr>
</tbody>
</table>
3.2 The analytical framework

In second language acquisition, learner language, namely, “interlanguage” termed by Selinker (1972), is important for research, as learners forming their own self-contained independent linguistic systems. It can reveal insights into the acquisition process, especially regarding the errors committed by learners. Learner errors are thought of as indications of incomplete learning, and that the speaker or hearer has not yet accumulated enough linguistic knowledge to enable them to avoid linguistic misuse in using their target language (Richards & Schmidt 2002). Corder (1967) was probably the first to point out and discuss the importance of errors learners make in the course of learning a second language and established error analysis (EA). The aim of EA, according to Richards & Schmidt (2002), is (1) to identify strategies which learners use in learning their target language, in terms of the approaches and strategies used in both teaching and learning, (2) to try to identify the causes or sources of learners’ errors, i.e., exploring the reasons behind such errors as the first attempt to get rid of them, and (3) to obtain information on common difficulties in language learning, as a reference to teaching or in the preparation of the teaching materials. Due to insufficient linguistic knowledge of the target language, learners have to express themselves with the help of communication strategies, i.e., strategies learners use when using their target language. One of the most commonly used communication strategies is avoidance, learners refraining from using linguistic forms which they do not know well (James 1998). Kleinmann (1977, 1978) found that language learners may resort to avoidance strategies, when they perceived a word or structure in a target language difficult to produce.

In this study, following Norrish (1983) and Ellis (1994), we define an error as a systematic deviation from the norms of the target language, which indicates that a learner has not learned something, and consistently gets it wrong. That is, an error takes place when a deviation arises as a result of lack of linguistic knowledge (i.e., competence). However, we would like to use a neutral term “non-target-like” to replace “error” in this study, similar to Chen (2006). One way of categorizing non-target-like use is based on how the surface structure of a sentence or expression is altered by learners. Surface strategy taxonomy, proposed by Dulay et al. (1982), is based on the ways surface structures are changed. Learners may omit necessary items or add unnecessary ones; they may misform items or misorder them. To develop a comparative taxonomy, a researcher should classify error types based on “comparisons between the structure of L2 errors and certain other types of constructions” (Dulay et al. 1982: 163). Later, James (1998) showed different types of learner errors relating to omission, overinclusion, misselection (use of wrong words, not wrong forms), misordering, and blend (which arises
when two alternative grammatical forms are combined to produce an ungrammatical blend), taking into consideration not only structure but also meaning and context. James argues that this model for the classification of errors can give an understanding about cognitive processes, which might reveal the learners’ reconstruction of their target language. His taxonomy is a typical analytical model for L2 errors.

3.3 Identification and categorization of errors

In Learner Corpus – TOCFL, the process of classifying words into their parts of speech (namely, word classes) and labeling them accordingly is known as part-of-speech (POS) tagging. The POS tagging processed words in sentences, and attached a part of speech to each word. Apart from the POS tagging, there were tags indicating where the learners have produced non-target-like uses, and what kinds of non-target-like uses these are based on James (1998).

In our study, in addition to taking into consideration the tags regarding hai in the corpus, we coded the data again to double check. The researchers were the coders for the data. The second author of the paper coded the errors, and then discussed them with the first author to reach agreement. In some cases, a learner’s sentence was vague and could be corrected in different ways; more than one target language form was possible. We mainly depended on the context to decide about what seemed to be correct in the target language. First, we searched for the occurrences of hai in the corpus and categorized the instances into target-like and non-target-like (i.e., errors). Next, the errors concerning hai in the learners’ corpus were classified into five types in terms of James’s taxonomy (1998) – over-inclusion, omission, misselection, misordering and blend. Errors were classified according to their type in comparison with what was judged to be correct in the target language. The example for each type11 is given as follows:

Type I: Over-inclusion – The presence of an item that should not appear in a well-formed sentence, as illustrated in (22) below.

(22) *Xiansheng yu shi sui de erzi biaoshi shenti bu shufu. Houlai, wo ye hai
Husband and 10 age DE son express body not well later 1SG also Hai
xiang you exin de ganjue.
think have disgusted DE feeling
‘My husband and my 10-year-old son said they were not feeling well. Later, I also felt disgusted.’

11. The error type, blend, involving the combination of more than one error type, was not found in the data.
Hai in (22) is not needed. The additive ye ‘also’, which focuses on the subject, is needed and hai is redundant here. In short, an over-inclusion error is defined as a redundant item (e.g., a word or a group of words) which would not have appeared in a well-formed sentence.

Type II: Omission – The absence of an item that would appear in an otherwise well-formed sentence, as shown in (23):

(23) *Yinwei ta budan xihuan ta laoshi de jiaofa, ta Ø¹² [hai] because 3SG not.only like 3SG teacher DE teaching.method 3SG Ø [HAI] juede xue Zhongwen hen youyong.¹³ think learn Chinese very useful

‘He not only liked the way his teacher taught, but also thought that learning Chinese was very useful.’

This type of error may result from semantic and contextual errors. In this particular example, both xihuan ‘like’ and juede ‘think’ denote a state. If the additive hai were used, the sentence would be grammatical (See our discussion related to (6)).

Type III: Misselection – The use of the wrong form of a morpheme or structure. A common error of this type in our data is the confusion between hai and haishi.

For example, based on the context, what is needed in (24) is a concessive haishi.

(24) *... jiu zhidao gaozhong de lianxi duo xinku, suiran wo zhidao zhe ge then know high.school DE practice how hard although 1SG know this CL shiqing, hai canjia bangqiu dui, yinwei wo xihuan bangqiu. (cf.... matter HAI participate baseball team because 1SG like baseball haishi canjia...) haishi participate (Japanese, B1)

‘...I knew how hard the practice in high school was. Although I knew this, I still joined the baseball team because I liked baseball.’

This type is related to the expression of meaning, which involves semantic, contextual and/or syntactic errors. For identifying this type, we searched for cotexts which often occur with hai and haishi.

12. ø stands for omission of hai.
13. As one reviewer points out, (23) can be analyzed as an omission of erqie ‘and’. We agree this is a possible analysis. But the fact that the speaker of (23) used neither erqie nor hai shows that he/she lacks the knowledge that either erqie or hai should be used in this case. So we can say there is an omission of erqie as well as an omission of hai in (23).
Type IV: Misordering – The incorrect placement of a morpheme or a group of morphemes in a sentence. For example, the adverb hai cannot occur higher than a subject, as shown in (25).

(25) *Na shihou, wo budan zhu zai ta de fumuqing jia, hai changchang that time 1SG not.only live at 3SG DE parent house HAI often tamen kaiche song women youming de difang, shenzhiyu youshihou ta 3PL drive send 1PL famous DE place even sometimes 3SG ma qinzi zuo cai. (cf. ... tamen haishi changchang ...) mother personally do dish 3PL HAISHI often (Korean, B1) ‘At that time, not only did I live at her parents’ house, they also drove us to famous places. Sometimes her mother even cooked for us.’

(25) involves the incorrect placement of hai before the subject tamen ‘they’ in the sentence, suggesting a syntactic error.

4. Results and discussion

4.1 The uses of hai in the data of the English-speaking, Japanese-speaking, and Korean-speaking CSL learners

In the corpus, there are 373 instances of hai produced by the English-speaking learners, 584 by the Japanese-speaking learners and 228 by the Korean-speaking learners, as summarized in Tables 2–4, respectively. Table 2 presents the overall frequencies of occurrences of hai across the four proficiency levels in the data of English-speaking learners of Chinese:

<table>
<thead>
<tr>
<th>Proficiency level</th>
<th>Target-like</th>
<th>Non-target-like</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>18</td>
<td>66.7</td>
<td>27</td>
</tr>
<tr>
<td>B1</td>
<td>144</td>
<td>68.6</td>
<td>210</td>
</tr>
<tr>
<td>B2</td>
<td>35</td>
<td>72.9</td>
<td>48</td>
</tr>
<tr>
<td>C1</td>
<td>77</td>
<td>87.5</td>
<td>88</td>
</tr>
<tr>
<td>Total</td>
<td>274</td>
<td>73.5</td>
<td>373</td>
</tr>
</tbody>
</table>

Of the 373 instances in the English-speaking learners’ writing, 99 cases are non-target-like. The error rate made by the English learners across the levels is
26.5%. Table 1 also shows that the error rate decreases from A2 to C1 (33.3%, 31.4%, 27.1% and 12.5%).

Of the 584 instances produced by the Japanese-speaking learners, 124 cases (21.2%) are non-target-like, as shown in Table 3.

Table 3. Uses of hai in the data of the Japanese-speaking learners of Chinese

<table>
<thead>
<tr>
<th>Proficiency level</th>
<th>Target-like</th>
<th></th>
<th>Non-target-like</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>A2</td>
<td>11</td>
<td>55</td>
<td>9</td>
<td>45</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>B1</td>
<td>279</td>
<td>80.2</td>
<td>69</td>
<td>19.8</td>
<td>348</td>
<td>100</td>
</tr>
<tr>
<td>B2</td>
<td>97</td>
<td>74.6</td>
<td>33</td>
<td>25.4</td>
<td>130</td>
<td>100</td>
</tr>
<tr>
<td>C1</td>
<td>73</td>
<td>84.9</td>
<td>13</td>
<td>15.1</td>
<td>86</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>460</td>
<td>78.8</td>
<td>124</td>
<td>21.2</td>
<td>584</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3 shows that the error rate decreases from A2 to B1 (45% vs. 19.8%) and from B2 to C1 (25.4% vs. 15.1%), but the error rates increase from B1 to B2 (19.8% vs. 25.4%). It seems that even learners at the high proficiency level had difficulty in mastering hai. This finding is different from Jiang’s (2013) result that error rates at the primary intermediate and advanced levels of the Japanese learners in HSK were not high (5.4% for the former and 4.6% for the latter), which will be discussed in the following section. Table 4 summarizes the instances of hai produced by the Korean-speaking learners of Chinese:

Table 4. Uses of hai in the data of the Korean-speaking learners of Chinese

<table>
<thead>
<tr>
<th>Proficiency level</th>
<th>Target-like</th>
<th></th>
<th>Non-target-like</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>A2</td>
<td>6</td>
<td>60</td>
<td>4</td>
<td>40</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>B1</td>
<td>97</td>
<td>74.6</td>
<td>33</td>
<td>25.4</td>
<td>130</td>
<td>100</td>
</tr>
<tr>
<td>B2</td>
<td>39</td>
<td>75</td>
<td>13</td>
<td>25</td>
<td>52</td>
<td>100</td>
</tr>
<tr>
<td>C1</td>
<td>24</td>
<td>88.9</td>
<td>3</td>
<td>11.1</td>
<td>27</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>166</td>
<td>75.8</td>
<td>53</td>
<td>24.2</td>
<td>228</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4 shows that the error rate made by the Korean learners across the levels is 24.2%. The error rate decreases from the low to high proficiency levels (40%, 25.4%, 25% and 11.1%).
4.2 Errors made by the CSL learners

The error types made by the three groups of learners were calculated, as summarized in Tables 5–7, separately.

4.2.1 Errors made by the English-speaking CSL learners

Table 5 summarizes the error types made by the English-speaking learners in our data.

Table 5. Non-target-like uses of *hai* across the proficiency levels in the English-speaking learners’ data

<table>
<thead>
<tr>
<th></th>
<th>A2</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Omission</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Misselection</td>
<td>3</td>
<td>46</td>
<td>9</td>
<td>10</td>
<td>68</td>
</tr>
<tr>
<td>Misordering</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Over-inclusion</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>9</td>
<td>66</td>
<td>13</td>
<td>11</td>
<td>99</td>
</tr>
</tbody>
</table>

The most common error type that the English-speaking learners made was misselection (68.7%, or 68 out of 99). There were three instances of misselection in A2; all concerned the replacement of *haishi* by *hai*. There were 46 instances of misselection in B1: 34 cases about the replacement of *haishi* by *hai*, as seen in (26) and (27), where the concessive *haishi* was required:

(26) *Suiran women dou zai hai li, baozheng hen anquan, gen women lai although 1PL all in sea inside guarantee very safe with 1PL come de Taiwan ren hai bu gan peng-dao shui, tamen shuo shi yinwei hai DE Taiwan people HAI not dare touch-arrive water 3SG say COP because sea limian you gui. inside have ghost. (English, B1)

‘Although we were all in the sea, it was guaranteed to be safe. The Taiwanese people who came with us still didn’t dare to touch the water. They said it was because there were ghosts in the sea.’
(27) *Wo Zhongwen shuiping tigao-le   hen duo. Huiqu Meiguo yihou, 1SG Chinese   level   improve-ASP   very   much   back   America   afterwards wo juede wo hai xiang qu zhu zai waiguo. 1SG feel 1SG HAI   want   go   live   in   foreign.country  (English, B1) ‘My Chinese has improved a lot. I think I still want to live in a foreign country when I go back to the US.’

Other instances of the misselection error type include the following: one about the replacement of qie (‘and’) by hai in a contrastive context; another about the replacement of you ‘also’ by hai in an additive context (see our discussion in Section 2.2); another about the replacement of dou ‘all’ by hai in a distributive context; another about the replacement of rengran ‘still’ by hai in a context indicating the meaning of ‘as before’. Recall from our discussion in Section 2.1, hai can obtain its temporal interpretation if it modifies a state or a negated situation. As shidai de nühai shoudao zhongnan qingnü na zhong kanfa de yingxiang in (28) does not involve a state or a negated situation, the temporal hai is not possible. Instead, rengran is allowed.

(28) *Nanren bu yinggai jin   chufang jiashi shi nüren zuo de. Shidai de man not should get.in kitchen housework COP woman do DE time DE nühai hai shoudao zhongnan qingnü na zhong kanfa de girl HAI get look.up.to.men look.down.on.women that kind notion DE yingxiang. influence  (English, B1) ‘Men shouldn’t go to the kitchen. Housework is women’s work. Modern girls are still affected by the notion of ‘looking up to men and down on women’.

In the B2 group, there were nine cases of misselection, all of which were about the replacement of haishi by hai; similarly, in C1, nine out of ten cases of misselections were about the replacement of haishi by hai, the others being about the replacement of que (‘but’) by hai. The English-speaking CSL learners from low to high proficiency levels all encountered such difficulty.

In addition, there were five cases where the learners used hai to replace the additive haiyou (‘and also’), as in (29):

(29) *Chule youyong zhi wai, hai bian de huodong hai hendrao. Wo xiwang besides swimming ZHI outside sea side DE activity HAI a.lot 1SG hope jianglai you shijian keyi zai qu Taiwan youming de hai bian. future have time can again go Taiwan famous DE sea side.  (English, B1) ‘There are still a lot of beach activities besides swimming. I hope I can have time to go to the famous beach in Taiwan again.’
In this example, *hai* should be followed by the verb *you* ‘have’ forming *haiyou* ‘and also’.

Omission was the second most common error type (14.1%, or 14 out of 99), most of which involved the additive *hai*, as seen in (30), repeated in (30), and (31) below:

(30) *Yinwei ta budan xihuan ta laoshi de jiaofa, ta Ø [hai] because 3SG not.only like 3SG teacher DE teaching.method 3SG Ø [HAI] juede xue zhongwen hen youyong. think learn Chinese very useful (English, A2)

‘He not only liked the way his teacher taught, but also thought that learning Chinese was very useful.’

(31) *Jieguo budan xiang bu chulai da’an, erqie Ø [hai] nei wo consequently not.only think not come.out answer moreover Ø [HAI] give 1SG ziji geng duo yali. self more much pressure (English, B2)

‘It turned out that I not only couldn’t figure out the answer, but also gave myself more pressure.’

In both (30) and (31), because *budan* ‘not only’ appears in the first clause, it would be well-formed if the second clause contained the additive *hai*. The omission type of error occurred at the A2, B1 and B2 levels and most were related to the additive *hai*, meaning ‘and also’.

Like the omission error type, over-inclusion was the second most common error type, as shown in (32):

(32) *Suoyi yao hua liang ge ban zhongtou cai huijia, dan zai che shang hai [Ø] liao de hao kaixin. So need take two CL half hour only return.home but in car LOC hai [Ø] liao de hao kaixin. (English, B1)

‘So it will take us two and half hours to get home. But we had a happy chat.’

(32) reveals that the English-speaking learners might know that like *still*, *hai* is related to the notion of contrast, but they overuse it in contrastive contexts because of their insufficient understanding of *hai* and *haishi*.

Misordering only occurs in B1, as shown in (33).
(33) *Xin xueqi de xuesheng zemenyang? Tamen dou guai guai de ma? Yinggai new semester student how 3PL all good good DE PRT should hai bi wo de xuesheng [hai] guai ba! HAI compare 1SG DE student [HAI] good PRT

(English, B1)

‘How are the students in the new semester? Are they all good? They should still be better than my students.’

The English-speaking learner put the adverb *hai before the preposition *bi (‘than’), instead of putting it before the stative verb (adjective). But most of the English-speaking learners (the Japanese-speaking and Korean-speaking learners as well) did not have problems with the scalar *hai meaning ‘moderately’, such as *hai bucuo ‘not bad’. This could be because they may be learned as formulaic expressions in daily conversation.

4.2.2 Errors made by the Japanese-speaking CSL learners

Similar to the English-speaking learners, the highest error rate made by the Japanese-speaking learners came from the type of misselection (55.6%, or 69 out of 124 in Table 6), but the error rate of omission (29%, or 36 out of 124 in Table 6) was higher than that made by the English-speaking learners (14.1%, or 14 out of 99 in Table 5).

Table 6. Non-target-like uses of *hai across the proficiency levels in the Japanese-speaking learners’ data

<table>
<thead>
<tr>
<th></th>
<th>A2</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Omission</td>
<td>5</td>
<td>19</td>
<td>12</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>Misselection</td>
<td>4</td>
<td>40</td>
<td>16</td>
<td>9</td>
<td>69</td>
</tr>
<tr>
<td>Misordering</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Over-inclusion</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>9</td>
<td>69</td>
<td>33</td>
<td>13</td>
<td>124</td>
</tr>
</tbody>
</table>

Like the English-speaking learners, they tended to use *hai in a concessive sentence, which requires *haishi, as manifested in (24), repeated in (34):
‘...I knew how hard the practice in high school was. Although I knew this, I still joined the baseball team because I liked baseball.’

Of the misselection errors in A2, all were instances where the learners used *hai in place of haishi. In B1, there were 33 cases concerning the replacement of haishi by hai; three cases concerning the replacement of rengr {‘still’} by hai; two cases concerning the replacement of the additive haiyou by hai; one concerning the replacement of ceng ‘ever’ by hai; and one concerning the replacement of dou ‘all’ by hai. In B2, there were 14 cases concerning the replacement of haishi by hai and two concerning about the replacement of the additive haiyou by hai. Even in C1, all of the misselection errors were due to the replacement of haishi by hai: There were nine cases in C1.

The second highest error type that the Japanese-speaking learners made was omission; they tended to omit the temporal hai, as shown in (35):

(35) Ta hui-le jia, jiu kaishi nianshu, dao banye Ø [hai] kan shu, 3SG return-ASP home then start study till midnight Ø [HAI] read book hen wan cai shuijiao.14 very late only sleep (Japanese, A2)

‘He went home, then started studying. Till midnight, he was still studying. He went to bed very late.’

Similar to the English-speaking learners, they also avoided using the additive hai, as seen in (36):

(36) Tamen chule tiaowu yiwai, Ø [hai] he hen duo jiu Tamen 3PL besides dancing outside Ø [HAI] drink very much wine 3PL wan-le shijian guo le. forget-ASP time pass ASP (Japanese, A2)

‘Besides dancing, they also drank a lot of wine. They forgot about the time.’

In the more advanced B2 level, the omission of the additive hai could be found, as shown in (37):

14. In addition to hai, the sentence involves an omission of the progressive marker zai right after hai.
(37) "Buguo ta budan mei you biaoshi ganxie, erqie Ø [hai] gen however 3SG not.only not have express appreciation moreover Ø [HAI] with ni shuo ‘Ni de Zhongwen yao jiayou!’ 2SG say 2SG de Chinese need cheer.up (Japanese, B2) ‘However, he didn’t appreciate it. Moreover, he said, “You need to improve your Chinese.”’

From the context, it seems that the speaker of (37) means to convey that the situation denoted by the second sentence is much less expected compared to that denoted by the first sentence, although shenzhi ‘even’ is not used in this sentence. For this reason, we consider that hai is needed, rather than erqie ‘and’ (See our discussion of (8)).

4.2.3 Errors made by the Korean-speaking CSL learners

Similar to the errors made by the English-speaking and the Japanese-speaking learners, two common errors made by the Korean-speaking learners were misselection and omission, i.e., 54.6% and 34% respectively, as shown in Table 7.

Table 7. Non-target-like uses of hai across the proficiency levels in the Korean-speaking learners’ data

<table>
<thead>
<tr>
<th></th>
<th>A2</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Omission</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>Misselection</td>
<td>0</td>
<td>20</td>
<td>7</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>Misordering</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Over-inclusion</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>33</td>
<td>13</td>
<td>3</td>
<td>53</td>
</tr>
</tbody>
</table>

The errors made by A2 were all about omission; such errors also appeared in B1, B2 and C1, which suggests that hai might be difficult for them to master. Interestingly, both the Korean and Japanese learners made a higher portion of errors in omission (29% and 34%, respectively) than the English learners (14.1%). In the A2 group, omission occurred in the temporal hai context, as shown in (38):

(38) ‘Yinwei wo houtian yao kaoshi, keshi wo Ø [hai] mei xue because 1SG the.day.after.tomorrow will take.exam but 1SG Ø [HAI] not learn wan zhe ke. Wo dasuan zai tushuguan nianshu. finish this lesson 1SG plan in library study (Korean, A2) ‘I have an exam the day after tomorrow, but I haven’t finished this lesson. I plan to study in the library.’
Even when they were at a more advanced level and had possibly learned the additive *hai*, they sometimes did not use it, as in (39):

(39) *Wo xiwang xuesheng bu shi zhiyou xue yuyan, tamen Ø [hai] hui xue yong zhe yuyan de sixinang wenhua shemede. ‘I hope that my students will learn not only the language, but also the thoughts, the culture, and so on.’

As for the most frequently made error type – misselection, there were 16 cases concerning the replacement of *haishi* by *hai*, three concerning the replacement of *haiyou* (‘and also’) by *hai*, and one concerning the replacement of *ye* (‘also’) by *hai* in B1; there were seven cases concerning the replacement of *haishi* by *hai* in B2; two cases concerning the replacement of *haishi* by *hai* in C1. In addition, avoidance tended to occur at the beginning level. When they were more advanced, they sometimes used it in the wrong order or overused it.

4.3 Overall discussions

The error rates of the four error types made by the three groups of learners are repeated in Table 8.

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th></th>
<th></th>
<th>Japanese</th>
<th></th>
<th></th>
<th>Korean</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Omission</td>
<td>14</td>
<td>14.1</td>
<td></td>
<td>36</td>
<td>29.0</td>
<td></td>
<td>18</td>
<td>34.0</td>
</tr>
<tr>
<td>Misselection</td>
<td>68</td>
<td>68.7</td>
<td></td>
<td>69</td>
<td>55.6</td>
<td></td>
<td>29</td>
<td>54.6</td>
</tr>
<tr>
<td>Misordering</td>
<td>3</td>
<td>3.0</td>
<td></td>
<td>8</td>
<td>6.5</td>
<td></td>
<td>3</td>
<td>5.7</td>
</tr>
<tr>
<td>Over-inclusion</td>
<td>14</td>
<td>14.1</td>
<td></td>
<td>11</td>
<td>8.9</td>
<td></td>
<td>3</td>
<td>5.7</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>100.0</td>
<td></td>
<td>124</td>
<td>100.0</td>
<td></td>
<td>53</td>
<td>100.0</td>
</tr>
</tbody>
</table>

For illustration, see Figure 1.
From Figure 1, it is clear that the most common error type for all three groups of learners was misselection. Table 9 below shows the number of errors that involved the replacement of *haishi* by *hai*.

Table 9. The numbers of the replacement of *haishi* by *hai* by all three groups of learners at all the levels

<table>
<thead>
<tr>
<th>Proficiency level</th>
<th>Learners</th>
<th>English-speaking learners</th>
<th>Japanese-speaking learners</th>
<th>Korean-speaking learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td></td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>B1</td>
<td></td>
<td>34</td>
<td>33</td>
<td>16</td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td>9</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>C1</td>
<td></td>
<td>9</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>55</td>
<td>60</td>
<td>25</td>
</tr>
</tbody>
</table>

In our data, more than 80% cases of misselection involved *hai* and *haishi*, particularly the concessive *haishi*. Specifically, the percentages of the replacement of *haishi* by *hai* by the English-speaking, Japanese-speaking and Korean-speaking learners were 80.9% (55 out of 68), 87% (60 out of 68) and 86.7% (25 out of 29), respectively. One possible reason for the misselection could be because both *hai* and *haishi* are often translated into ‘still, yet’. Chinese words such as *rengran,*
yiran, etc. are glossed as ‘still’ and words such as ye, you, etc., glossed as ‘also’ in the learners’ teaching materials. Errors of misselection are evidence of gaps in knowledge of the target language rather than merely being L1 interference or transference. Empirical studies show that if a word has two or more near translation equivalents in the L2, a learner often feels confused and tends to choose the most familiar one (see Cai 2013).

Table 8 and Figure 1 also reveal that the two most common error types committed by the Japanese-speaking learners and English-speaking learners were mis-selection and omission. The error rates of omission by the Japanese learners and the Korean-speaking learners were higher than those of the English learners (29.0% and 34% vs. 14.1%), whereas the error rate of misselection by the English learners was higher than those of the Japanese-speaking learners and Korean-speaking learners (68.7% vs. 55.6% and 54.6%). A plausible reason for this difference may be L1 transfer. For the Japanese learners, they may take mada as equivalent to Mandarin hai. But note that mada does not have the meaning ‘also’.

Therefore, the Japanese learners might not use the additive hai as an adverb occurring before the main verb and might make some errors using the additive hai. Similarly, the counterpart of hai in Korean is ajig (아직), which means ‘still, yet’. Ajig does not have an additive meaning, which has to be expressed by do (도) ‘also,’ ‘too’ or ‘as well’. Thinking that ajig in their first language parallels to Mandarin hai, Korean learners might not use hai in a context meaning ‘and also.’ The difference between Japanese mada and Chinese hai, as well as the difference between Korean ajig and Chinese hai, belongs to Level 5 – Split in Prator’s (1971) six categories of hierarchy of difficulty in ascending order: One item in L1 becomes two or more in the target language, which requires learners to make a new distinction. Since the counterparts of hai in Japanese and Korean, i.e., mada and ajig, have no ‘and also’ meaning, the learners probably tended to avoid the additive hai. On the other hand, still in English has an additive meaning, as in There are many comic books in my room. One is on my desk. Another is on the floor. Still another is on my bed. This may explain why compared to the Japanese-

15. Although L. Li (2009) argues that mada has the meaning of ‘in addition’, we consulted with a native speaker of Japanese and confirmed that it is mata (又) with a low-high tone, not mada (まだ) with a high-low tone, that has an additive meaning. When mata occurs with a nominal marked by mo ‘also’, it is emphatic. Take (i) for example:

(i) John は かくせい だった. Mary も また かくせい だった.
John wa  gakusei  da-tta Mary mo  mata  gakusei da-tta
John top student be-PAST Mary also again student be-PAST
‘John was a student. Mary was also a student.’

We thank one of the reviewers for raising this point.
speaking and Korean-speaking learners, the English-speaking learners did not commit as many errors related to the omission of the additive hai.

Many studies on second language acquisition have demonstrated that L1 lexical transfer plays an important role. The results from our study conform to some earlier studies in this research field which have concluded that the majority of errors beginning L2 learners make might stem from interlingual interference. The learners at the A2 level in our study tended to rely more on their native language when dealing with aspects of the target language not yet mastered.

In addition to L1 transfer, at least three other factors affecting the acquisition of the different meanings of hai can be identified: (1) prototypicality of the meanings, (2) teaching materials, and (3) the frequency of the different meanings within the language. Firstly, according to prototype theory (Berlin & Kay 1969, Rosch 1973, 1975, 1978, Brugman & Lakoff 1988), the most central and prototypical meanings of a polysemous network will be the easiest to learn. In the present study, we identified five meanings of hai: temporal hai, additive hai, comparative hai, moderate hai, counter-expectation hai and concessive haishi. The polysemous hai is challenging for learners. Secondly, teaching materials may be a problem. Most learners cannot make a distinction between hai and haishi, both of which are glossed as ‘still’ in the learners’ teaching materials. For example, Practical Audio-visual Chinese, one of the most used textbooks in Taiwan, glosses hai as still, yet, also (p.124) in Book 1, Lesson 7, as shown in (40) and (41):

(40) Tade Yingwen hai shuo de buhao.
3SG English hai say de not.good
‘He still can’t speak English very well.’

(41) Ta hui change, hai hui tiaowu.
3SG can sing hai can dance
‘She can sing and also dance.’

Crucially, this textbook does not explain how the temporal hai in (40) is different from the additive hai in (41).

As far as the concessive haishi is concerned, it first appears in Book 1, Lesson 12 (p.261) and is simply translated into still without further explanation.

(42) Wo shuo-le bantian, ta haishi bu dong.
1SG say-asp half.day 3SG HAISHI not understand
‘I talked to him for a long time, but he still didn’t understand.’

For learners using the textbook, they may acquire most easily the central meaning, i.e., the temporal hai, because they can find an equivalent in their L1. Apart from prototypicality and teaching materials, words with greater frequency will be
learned faster; the more learners are exposed to the meaning of a word, the more they are likely to learn it. Since the temporal hai is the basic meaning of hai and is often heard in daily conversation, it is easier to acquire than the other meanings. The moderate hai collocating with bucuo (literally meaning ‘not wrong’) or hao (‘good’) is also often heard in daily life, it is easy for the learners to use the expressions haibucuo and haihao. However, the learners would overgeneralize it in contexts where the concessive haishi is required or omit it in contexts where the additive hai or haiyou is required.

As we have seen above, the data show that the most difficult area seems to be related to the distinction between hai and haishi; in particular, the learners we studied commonly used hai to replace haishi in contexts where haishi was required. Besides, they often omitted hai or haiyou when it was required. They tried to avoid using the difficult words hai, haiyou and haishi, and used simpler words instead. Avoidance in second language acquisition is defined as the choice to use one language feature over another, in order to avoid producing an error (Schachter 1974, Kleinmann 1977, 1978). That is, if a learner finds some particular construction in the target language difficult to produce it is likely that he/she will try to avoid it. This phenomenon in L2 acquisition was termed ‘avoidance behavior’ by Schachter (1974). As mentioned in Section 3, avoidance behavior was a communication strategy that L2 learners may resort to when, with insufficient knowledge of a target language word or structure, they perceive that it is difficult to produce (Kleinmann 1977, 1978).

The research synthesized in this paper has four implications for classroom practice. First, it should be emphasized that both the temporal hai and concessive haishi are adverbs and occur only in a post-subject position. Second, a concessive haishi is required when a temporal hai is not available in the contrastive context. Third, learners also need to know that the temporal hai has to do with a state or a negated situation and that the additive hai has to be associated with an event that involves quantity. For an additive hai to be used with a state, conjunctions such as bujin/budan ‘not only’ or chule/erqie ‘besides/but also’ have to be used (see (6) above). Fourth, because of the polysemy of hai, some learners might use an avoidance strategy, even at the intermediate proficiency level, especially Japanese-speaking and Korean-speaking learners.

5. Conclusion

The findings of our study are as follows. The most frequent error type made by our subjects was misselection. Of the English-speaking, Japanese-speaking and Korean-speaking learners, the English-speaking learners made most misselection
errors. Specifically, the subjects tended to use *hai* in place of *haishi* when they made concessive sentences and used the concessive *haishi* when a temporal *hai* was needed. Also they had problems in making a distinction between *hai*, *ye* ‘also’, *you* ‘also’, and *zai* ‘again’, in contexts where an additive *hai* was required.

The second most frequent error type was omission. The Japanese-speaking and Korean-speaking learners made more omission errors than the English-speaking learners. Because the additive *hai* does not correspond to the counterpart in their L1, the Japanese and Korean speakers tended not to use it, especially those at the low proficiency level.

Finally, omission and over-inclusion were equally challenging for the English-speaking learners. Compared to the Japanese-speaking and Korean-speaking learners, the English-speaking learners tended to make more errors in over-inclusion, adding a redundant item. Since *hai* is related to the notion of contrast, they might overuse it in contrastive contexts.

Four factors affecting learner acquisition of the different meanings of *hai* have been identified. They are L1 transfer, prototypicality of the meanings, teaching materials, and the frequency of the different meanings within the language. It is hoped that the suggestions for classroom practice in Section 4 can help CSL teachers and educators in making objective decisions about how to go about adopting appropriate teaching strategies to help different groups of CSL students in learning the Mandarin *hai*.

**Acknowledgements**

This study benefited from the funding of National Taiwan Normal University’s ‘Aim for the Top University Project’. We were grateful for it and would also like to thank the two reviewers and the editor Hui-shan Lin for their insightful comments. All remaining errors are ours.

**Abbreviations**

<table>
<thead>
<tr>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
<th>ASP</th>
<th>CL</th>
<th>COP</th>
<th>DE</th>
<th>HAI</th>
<th>HAISHI</th>
<th>LOC</th>
<th>PAST</th>
<th>PRT</th>
<th>TOP</th>
<th>ZHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person singular pronoun</td>
<td>2nd person singular pronoun</td>
<td>3rd person singular pronoun</td>
<td>1st person plural pronoun</td>
<td>2nd person plural pronoun</td>
<td>3rd person plural pronoun</td>
<td>aspect marker</td>
<td>classifier</td>
<td>copula</td>
<td>associative/complementizer <em>de</em></td>
<td><em>hai</em></td>
<td><em>haishi</em></td>
<td>localizer</td>
<td>past marker</td>
<td>clause final particle</td>
<td>topic marker</td>
<td>associative <em>zhi</em></td>
</tr>
</tbody>
</table>
References


Acquiring the polysemous adverb HAI in Chinese
https://doi.org/10.1016/0010-0285(73)90017-0


https://doi.org/10.1111/j.1467-1770.1974.tb00502.x


https://doi.org/10.1186/s40655-016-0018-8


### Appendix

The appendix can be found online at: https://doi.org/10.1075/consl.00017.hsi.appendix

### Address for correspondence

Miao-Ling Hsieh  
Department of English  
National Taiwan Normal University  
Taipei, TAIWAN  
mlhsieh@ntnu.edu.tw
Co-author information

Yu-Fang Wang
Graduate Institute of Teaching Chinese as a Second/Foreign Language
National Kaohsiung Normal University
Kaohsiung, TAIWAN
yfwang@nk nucc.nknu.edu.tw

Publication history

Date received: 14 August 2019
Date revised: 7 November 2019
Date accepted: 18 December 2019