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

[英日韓二語學習者對於漢語多義詞「還」之習得]

Miao-Ling Hsieh [謝妙玲] and Yu-Fang Wang [王莢芳]

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

Appendix available from <https://doi.org/10.1075/consl.00017.hsi.appendix>
<https://doi.org/10.1075/consl.00017.hsi>

Concentric 46:2 (2020), pp. 173–205. issn 1810-7478 | e-issn 2589-5230

Available under the CC BY-NC 4.0 license. © John Benjamins Publishing Company

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*.¹ 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 HSK² 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/fastness. 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*.³ 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):⁴

- (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 comparatives, 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).

- (6) Ta ?? (bujin) hen congming, **hai** hen yonggong.
 3SG not.only very smart HAI very hardworking
 'He is not only smart, but also hardworking.'

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):⁶

- (8) Ta meitian lianxi, shenzhi **hai** hen yonggong.
 3SG every.day practice even HAI very diligent
 'He practices every day and is even very diligent.'

(8) means the diligent 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 than 3SG older.brother Dayong HAI strong
 ‘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 required⁷, 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-

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):

- (i) Laowang de wuzi **hai** ganjing.
 Laowang **DE** room **HAI** clean
 ‘Laowang’s room is moderately clean.’ (F. Liu 2000: 42)

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*’.

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 12⁸:

- (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 an aspectual 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.'

8. National Taiwan Normal University (ed.) 2008. *Practical Audio-Visual Chinese* (2nd edition, vol. 3). Taipei: Cheng Chung Book Co., Ltd.

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.¹⁰ 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 1. The number of the English-speaking, Japanese-speaking, and Korean-speaking learners of Chinese participating in the corpus

Proficiency levels Participants	A2 No. of texts/ learners	B1 No. of texts/ learners	B2 No. of texts/ learners	C1 No. of texts/ learners	Total
English-speaking	123	252	59	18	452
Japanese-speaking	256	402	125	13	796
Korean-speaking	74	144	57	8	283
Total	453	798	241	39	1531

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 xin de ganjue.
 think have disgusted DE feeling (Japanese, B2)
 '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]
 jue de xue Zhongwen hen youyong.¹³
 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.’

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 2. Uses of *hai* in the data of English-speaking learners of Chinese

Proficiency level	Target-like		Non-target-like		Total	
	N	%	N	%	N	%
A2	18	66.7	9	33.3	27	100
B1	144	68.6	66	31.4	210	100
B2	35	72.9	13	27.1	48	100
C1	77	87.5	11	12.5	88	100
Total	274	73.5	99	26.5	373	100

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

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

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

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

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

	A2	B1	B2	C1	Total	
	N	N	N	N	N	%
Omission	5	7	2	0	14	14.1
Misselection	3	46	9	10	68	68.7
Misordering	0	3	0	0	3	3.0
Over-inclusion	1	10	2	1	14	14.1
Total	9	66	13	11	99	100

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.’

- *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, Bi)
 ‘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
nūhai hai shoudao zhongnan qingnū 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 henduo. 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 (23), 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] gei 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
So need take two CL half hour only return.home but in car LOC
hai [Ø] liao de hao kaixin.
HAI chat DE very happy (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 zenmeyang? Tamen dou guai guai de ma? Yinggai
 new semester DE 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

	A2	B1	B2	C1	Total	
	N	N	N	N	N	%
Omission	5	19	12	0	36	29.0
Misselection	4	40	16	9	69	55.6
Misordering	0	3	3	2	8	6.5
Over-inclusion	0	7	2	2	11	8.9
Total	9	69	33	13	124	100

Like the English-speaking learners, they tended to use *hai* in a concessive sentence, which requires *haishi*, as manifested in (24), repeated in (34):

- (34) *... 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.
matter HAI participate baseball team because 1SG like baseball
(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.’

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 *rengran* ‘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 biaooshi ganxie, erqie Ø [hai] gen
 however 3SG not.only not have express appreciation moreover Ø [hai] with
 ni shuo “Ni de Zhongwen yao jiaoyou!”
 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

	A2	B1	B2	C1	Total	
	N	N	N	N	N	%
Omission	4	8	5	1	18	34.0
Misselection	0	20	7	2	29	54.6
Misordering	0	2	1	0	3	5.7
Over-inclusion	0	3	0	0	3	5.7
Total	4	33	13	3	53	100

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
1SG hope student not COP only learn language 3PL Ø [HAI] will learn
yong zhe yuyan de sixiang wenhua shemedede.
use this language DE thought culture and.so.on (Korean, B1)
'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 8. The percentages of non-target-like uses of *hai* in the English-speaking, Japanese-speaking, and Korean-speaking learners' data

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

For illustration, see Figure 1.

Figure 1. Error rates of the four error types made by the English-speaking, Japanese-speaking and Korean-speaking learners

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

	Learners	English-speaking learners	Japanese-speaking learners	Korean-speaking learners
Proficiency level				
A2		3	4	0
B1		34	33	16
B2		9	14	7
C1		9	9	2
Total		55	60	25

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 misselection 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’.¹⁵ 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 がくせい だった。Mary も また がくせい 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

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

References

- Aosakul, Woraphras. 2011. A Comparative Study and the Error Analysis on Adverbs *Hai, You, Zai* in Mandarin and *Ik* and Corresponding Word in Tai. MA thesis, Zhejiang University, Hangzhou.
- Berlin, Brent, and Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley, CA: University of California Press.
- Biber, Douglas. 2004. Historical patterns for the grammatical marking of stance: A cross-register comparison. *Journal of Historical Pragmatics* 5.1:107–136.
<https://doi.org/10.1075/jhp.5.1.06bib>
- Biber, Douglas, and Edward Finegan. 1988. Adverbial stance types in English. *Discourse Processes* 11.1:1–34. <https://doi.org/10.1080/01638538809544689>
- Biber, Douglas, and Edward Finegan. 1989. Style of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text & Talk: An Interdisciplinary Journal of Language, Discourse & Communication Studies* 9.1:93–124.
- Biq, Yung-O. 2004. From collocation to idiomatic expression: The grammaticalization of *hao* phrases/constructions in Mandarin Chinese. *Journal of Chinese Language and Computing* 14.2:73–95.
- Brugman, Claudia, and George Lakoff. 1988. Cognitive topology and lexical networks. *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology and Artificial Intelligence*, ed. by Steven L. Small, Garrison W. Cottrell and Michael K. Tanenhaus, 477–508. San Mateo, CA: Morgan Kaufmann Publishers.
<https://doi.org/10.1016/B978-0-08-051013-2.50022-7>
- Cai, Jinting. 2013. Lexical transfer in Chinese EFL learners' motion event expressions: Group tendency and individual variation. Paper presented at the 2nd SLA Summit Forum in China, Ningbo.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.
- Chen, Fred Jyun-gwang. 2006. Interplay between forward and backward transfer in L2 and L1 writing: The case of Chinese ESL learners in the US. *Concentric: Studies in Linguistics* 32.1:147–196.
- Corder, Stephen Pit. 1967. The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching* 5.1–4:161–170.
- Donazzan, Marta. 2008. Presupposition on times and degrees: The semantics of Mandarin *hai*. *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, vol. 2, ed. by Marjorie K. M. Chan and Hana Kang, 597–610. Columbus, OH: The Ohio State University.
- Dulay, Heidi C., Marina K. Burt, and Stephen D. Krashen. 1982. *Language Two*. Oxford, UK: Oxford University Press.
- Ellis, Rod. 1994. *The Study of Second Language Acquisition*. Oxford, UK: Oxford University Press.
- Goffman, Erving. 1967. *Interaction Ritual: Essays in Face-to-Face Behavior*. Chicago, IL: Aldine Publishing Company.
- Hou, Xuechao (ed.) 2004. *Xiandai Hanyu Xuci Cidian* [Dictionary of Modern Chinese Function Words]. Beijing: Peking University Press.

- James, Carl. 1998. *Errors in Language and Use: Exploring Error Analysis*. Harlow, Essex: Longman.
- Jiang, Xiezhong. 2013. Riben liuxuesheng Hanyu fuci *hai* de xide kaocha [A study of the Japanese students' acquisition of Chinese adverb *hai*]. *Haiwai Huawen Jiaoyu* 1:71–78.
- Jing-Schmidt, Zhuo, and Stefan Th. Gries. 2009. Schematic meaning and pragmatic inference: The Mandarin adverbs *hai*, *zai* and *you*. *Corpora* 4.1:33–70. <https://doi.org/10.3366/E1749503209000215>
- Jing, Xuanjiao. 2012. Hanguo Liuxuesheng Xide Chongfu Yi *Hai*, *Zai*, *You* de Pianwu Fenxi [The Error Analyses of Korean Students' Acquisition of Repeated Meaning: *Hai*, *Zai*, *You*]. MA thesis, Jilin University, Changchun.
- Kleinmann, Howard H. 1977. Avoidance behavior in adult second language acquisition. *Language Learning* 27.1:93–107. <https://doi.org/10.1111/j.1467-1770.1977.tb00294.x>
- Kleinmann, Howard H. 1978. The strategy of avoidance in adult second language acquisition. *Second Language Acquisition Research: Issues and Implications*, ed. by William C. Ritchie, 157–174. New York: Academic Press.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA: University of California Press.
- Li, Lin. 2009. A study of the factors in Relation to errors incurred from the acquisition of Chinese modal adverbs. *Journal of Yunnan Normal University: Teaching and Research on Chinese as a Foreign Language Edition* 6:31–35.
- Li, Yanze. 2011. Waiguo Liuxuesheng Chongfuyi Fuci *Zai*, *Hai*, *Ye*, *You* de Pianwu Fenxi [An Analysis on the Repeated-meaning Adverbs *Zai*, *Hai*, *Ye*, *You*: Errors Made by Foreign Students in Learning Chinese]. MA thesis, Liaoning Normal University, Dalian.
- Liu, Feng-hsi. 2000. The scalar particle *hai* in Chinese. *Cahiers de Linguistique Asie Orientale* 29.1:41–84. <https://doi.org/10.3406/clao.2000.1565>
- Liu, Huijuan. 2009. Additive Particles in Adult and Child Chinese. Doctoral dissertation, City University of Hong Kong, Hong Kong.
- Liu, Huiying. 2011. Chuji Jieduan Riben Liuxuesheng Xuexi Hanyu Biao Chongfuyi Fuci *Zai*, *You*, *Hai* de Pianwu Fenxi [An Error Analysis of Japanese Students at Primary Level in Learning Chinese Adverbs *Zai*, *You*, *Hai*]. MA thesis, Central China Normal University, Wuhan.
- Lü, Shuxiang. 1980. *Xiandai Hanyu Babai Ci* [800 Words in Modern Chinese]. Hong Kong: Commercial Press.
- Ma, Zhen. 1999. Guanyu biao chongfu de fuci *you*, *zai*, *hai* [Concerning the adverbs *you*, *zai*, *hai* expressing repetition]. *Xiandai Hanyu Xuci Sanlun*, 128–142. Beijing: Yuwen Chubanshe.
- Mandarin Teaching Center at NTNU. 2008. *Practical Audio-Visual Chinese* (2nd edition). Taipei: Zhengzhong Publishing Company.
- Mao, Jingya. 2011. The Comparison of *Hai*, *Zai*, *You*, *Ye* and Analysis on Errors Made by L2 Learners. MA thesis, Central China Normal University, Wuhan.
- Norrish, John. 1983. *Language Learners and Their Errors*. London: Macmillan Press.
- Prator, Clifford H. 1971. Phonetics vs. phonemics in the ESL classroom: When is allophonic accuracy important? *TESOL Quarterly* 5.1:61–72. <https://doi.org/10.2307/3586112>
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Richards, Jack C., and Richard Schmidt. 2002. *Longman Dictionary of Language Teaching and Applied Linguistics* (3rd edition). London: Longman.

- Rosch, Eleanor. 1973. Natural categories. *Cognitive Psychology* 4.3:328–350.
[https://doi.org/10.1016/0010-0285\(73\)90017-0](https://doi.org/10.1016/0010-0285(73)90017-0)
- Rosch, Eleanor. 1975. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104.3:192–233. <https://doi.org/10.1037/0096-3445.104.3.192>
- Rosch, Eleanor. 1978. Principles of categorization. *Cognition and Categorization*, ed. by Eleanor Rosch and Barbara B. Lloyd, 27–48. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schachter, Jacqueline. 1974. An error in error analysis. *Language Learning* 24.2:205–214.
<https://doi.org/10.1111/j.1467-1770.1974.tb00502.x>
- Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics in Language Learning* 10.3:209–231.
- Shen, Jiaxuan. 2001. Gen fuci hai youguan de liangge jushi [Two constructions with hai in Mandarin Chinese]. *Zhongguo Yuwen* 6:483–493.
- Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalization. *Subjectivity and Subjectivisation: Linguistic Perspectives*, ed. by Dieter Stein and Susan Wright, 31–54. Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511554469.003>
- Traugott, Elizabeth Closs, and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge, UK: Cambridge University Press.
- Wu, Guo. 2009. Fuci hai de zhuguanxing yongfa [The subjective uses of the adverb hai]. *Shijie Hanyu Jiaoxue* 3:322–333.
- Wu, Ying. 2010. Haishi de duoyixing yu xide nandu [On the polysemy of haishi and its difficulty in learning]. *Huawen Jiaoxue yu Yanjiu* 4:41–48.
- Yang, Ching-Yu Helen. 2017. On the syntax-semantics interface of focus particles: The additive particle hai ‘HA!’ in Mandarin Chinese. *Lingua Sinica* 3, article no. 3.
<https://doi.org/10.1186/s40655-016-0018-8>
- Yeh, Meng. 1998. On hai in Mandarin. *Journal of Chinese Linguistics* 26.2:236–280.

Appendix

The appendix can be found online at: <https://doi.org/10.1075/cons.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@nknuc.nknu.edu.tw

Publication history

Date received: 14 August 2019

Date revised: 7 November 2019

Date accepted: 18 December 2019