Locative Particles in Spoken Taiwan Mandarin

Yung-O Biq (畢永峨)
National Taiwan Normal University
YBIQ@ntnu.edu.tw

IN

Studies of Chinese Linguistics: Functional Approaches

Edited by Janet Zhiqun Xing

pp. 133-153

2009

Hong Kong: Hong Kong University Press
This paper examines the behavior of locative particles in the Mandarin spoken in Taiwan. Generally speaking, in order to refer to a location, the Chinese NP is followed by (1) a spatial or orientation term such as shàng, qián, lǐ, etc. (or equivalent) – (LOC1), and (2) an optional particle such as –miàn, -tóu, -biān – (LOC2). Our spoken Taiwan Mandarin data reveals that –miàn is virtually the only LOC2 going with the six basic spatial orientation LOC1s, shàng, xià, qián, hòu, lǐ, and wài. We question the difference between the simple locative form – the one without –miàn, and the complex locative form – the one with –miàn. For xià, qián, hòu, and wài sets, these two forms show a tendency to carry distinct functions: while the simple form is mostly used for non-spatial references, the complex form is mostly used for spatial references. However, such semantic/functional division of labor does not fully explain the cases of shàng(miàn) and lǐ(miàn), whose token frequencies both outnumber the other four sets in the database. Instead, a morphological/structural account seems more plausible: the simple form (i.e., shàng or lǐ) is largely a suffix or even a word-internal element, whereas the complex form (shàngmiàn or lǐmiàn) can be both a suffix and a phrasal enclitic. The two sets are further contrasted with respect to the competitive relationship between the simple form and the complex form.

1. Introduction

This paper examines the choice of simple or complex locative particles in Mandarin speech.

Except for nouns that are intrinsic place names (e.g., 台灣 taiwān ‘Taiwan’) or those that have been conventionally taken as place names (e.g., 飛機場 fēijīcháng ‘airport’) (cf. discussion in 储[Chu] 1997), Mandarin grammar requires that in order to become a place term referring to location, the noun (or NP) needs to be followed by (a) a spatial orientation term such as shàng, qián, lǐ, etc. (or equivalent), and (b) an optional suffix, i.e., -面 miàn, -頭 tóu, or -邊 biān.1 Traditionally, when (a) is present without (b), it is called a simple (or monosyllabic) locative particle (單音方位詞 dānyīn fāngwèi cí), and when both (a) and (b) are used, the form is called a complex (or disyllabic) locative particle (複合方位詞 fúhé fāngwèi cí) (e.g., 劉[Liu et al.] 1983/1996).

These items are of nominal origin. 劉[Liu] (2003) states that in Modern Chinese the simple particles are virtually no longer free morphemes. They typically follow a nominal and can be expressions with a wide range of meanings beyond those concerning spatial orientation. The particles thus become functional morphemes as a result of repeated use. The meaning of these simple particles, furthermore, is bleached and the differentiation among these spatially oriented terms has been narrowed and even neutralized. Some semantically paired-up terms (e.g., 上 shàng ‘up’ vs. 下 xià ‘down’) can be quite different in terms of use
frequency (2003: 157-158). The complex particles, on the other hand, retain their nominal characteristic to a greater extent, and can be seen as nouns about (spatial) relationships which can also be used as postpositions (2003: 158-159). Indeed, the complex forms can occupy the subject or object position in a sentence (e.g., 外面太吵了 wàimiàn tài chāo le ‘It’s too noisy outside.’), while the simple forms cannot. In this study, the free word use of the complex forms will be discussed only in contrast to the post-nominal use since the main question we ask here is the difference between the simple form and the complex form when both of them can occur after a nominal.

To native speakers, the simple form and the complex form appear to be free variations. Native speakers do not seem to have straightforward answers when asked about the difference between the two. In fact, numerous examples from our data show that the two forms do seem interchangeable since both are found to follow the same NP. Consider the following pair of examples, in which both the simple 上 shàng and the complex 上面 shàngmiàn occur after 工作 gōngzuò ‘work’:

(1a) 現在工作上會調整吧
    xiànzài gōngzuò shàng huì tiáozhěng ba
    ‘As for now (I) will probably get reappointed (promoted) at work.’

(1b) 有時候反而工作上面是完全玩不得的
    yǒushíhòu fánér gōngzuò shàngmiàn shì wánquán wánbùde de
    ‘On business trips, I sometimes am too busy and have little time for recreation.’

The two conversation units in the following were uttered a few minutes apart by the same speaker, but again both 裡 lǐ and 裡面 lǐmiàn follow the same NP, 父母的懷抱 fūmǔ de huáibào ‘parents’ bosom/protection’.

(2a) 他沒有辦法依偎在父母的懷抱裡
    tā méiyǒu fānǎ yǐwēi zài fūmǔ de huáibào lǐ suǒyí ne
    ‘He could not remain under his parents’ protection so….’

(2b) 或者是可以依偎在父母的懷抱裡
    huòzhě shì kěyǐ yǐwēi zài fūmǔ de huáibào lǐmiàn
    ‘Or (he) could remain under his parents’ protection.’

Previous studies have addressed the difference between the simple and the complex locative particles. Li and Thompson (1981:392) suggest that there is a syllabic correlation: if the preceding noun is monosyllabic, it tends to take the simple, monosyllabic locative particle; if the noun is disyllabic, it tends to take the complex, disyllabic locative particle. In 吕[Lü] (1965:209), however, this syllabic parallelism is rejected based on written data. In fact, the complex form is in general much less frequently used than the simple form in Lü’s written data, which consisted of writings mostly dated in the first half of the 20th century or earlier. Furthermore, Lü stated that while there is no functional difference between the three optional
suffix alternatives (面 miàn, 頭 tóu, 邊 biān), text types seem to matter: written language favors 面 miàn while spoken language favors -頭 tóu and -邊 biān.

Decades have passed since Lü’s corpus-based study, and the language has certainly undergone changes due to both time and regional variation. The present study attempts an investigation into how the simple and the complex locative particles are used in the Mandarin spoken in contemporary Taiwan since, differing from Lü’s findings (from written Chinese), numerous token examples of both simple and complex locative particles occur in our spoken data, which consist of about 15 hours of naturally occurring conversation. We focus on speech because we believe daily conversation most faithfully reflects the linguistic structure in the (native) speaker’s mind.

2. Narrowing Down the Focus of Study

Hereafter, we will call the (a) particle LOC1 and the (b) particle LOC2. Many spatial orientation terms can serve as LOC1, including 上 shàng, 下 xià, 前 qián, 後 hòu, 裡 lǐ, 外 wài, 左 zuò, 右 yòu, 東 dōng, 南 nán, 西 xī, and 北 běi. We will focus on six terms – 上 shàng ‘up’, 下 xià ‘down’, 前 qián ‘before’, 後 hòu ‘after’, 裡 lǐ ‘in’, and 外 wài ‘out’, for two reasons: (i) they designate the most basic human spatial orientations, and (ii) they also occur much more frequently in daily speech than the other LOC1 candidates. While it has been repeatedly documented that different LOC1s show different co-occurrence patterns with the three LOC2 alternatives (e.g., Chao 1968: 528; Li & Thompson 1981: 391; 呂 [Lü] 1965), the six LOC1s that we are interested in generally agree with one another in their co-occurrence likelihood with the three LOC2s in spoken Taiwan Mandarin. In fact, they mostly co-occur with 面 miàn only (if they ever take a LOC2). In our database, -頭 tóu and -邊 biān seldom go with the six LOC1s. The LOC2 –頭 tóu is found to co-occur with 裡 lǐ four times in our 15 hour conversation data, and it is not found to co-occur with any of the other 5 LOC1s. On the other hand, while the LOC2 –邊 biān is part of such spatial deictic words as 左邊 zuòbiān ‘left’, 右邊 yòubiān ‘right’, and 旁邊 pángbiān ‘side’, it is not found to co-occur with any of the 6 LOC1s that we are concerned with here at all. In this study, we will then focus on both the simple form (LOC1 only) and the complex form (LOC1 + LOC2) as they show up as particles at the post-nominal position, with the 6 primary spatial terms as LOC1 and 面 miàn as the only LOC2.

While the six LOC1s are spatial orientation terms, all of them have also demonstrated meaning extensions in their combination with nominals that do not refer to concrete objects or anything that has a physical or spatial orientation. However, the extent of this functional “versatility” varies among the six terms. Previous studies have pointed out that 上 shàng and 裡 lǐ are the two most “versatile” items (e.g., Chao 1968; 呂 [Lü] 1965; 劉 [Liu] 2003) and accounts offered are mostly cognitively oriented (e.g., 崔 [Cui] 2002; 方 [Fang] 2003; 鄒 [Zou] 2001; 鄒 王 [Zou and Wang] 2002). Our spoken data confirm the special status of 上 shàng and 裡 lǐ. For one thing, they show token frequencies much higher than the other four sets in our database. Table 1 below provides the token numbers for the six sets in the database. We can see that 上 shàng is the most frequent simple form (first column) (288 tokens). On the other hand, 裡面 lǐmiàn is the most frequent complex form, either as a particle (second column; 189 tokens), as a free word (third column; 77 tokens), or in total (rightmost column; 266 tokens). Because of their frequency and functional versatility, we will pay special attention to 上(面) shàng(miàn) and 裡(面) lǐ(miàn) in this study.
To sum up, the question we ask in this paper is whether there is any difference between the simple form and the complex form as a particle. Are there any detectable consistencies across the 6 LOC1s with regard to the simple vs. complex choice? Is it the case that the two alternatives have distinct functions? Furthermore, our attention is particularly drawn to 上(面) shàng(miàn) and 裏(面) lǐ(miàn) since both show outstanding frequency and functional versatility compared to the other four sets. In the following discussion, we will examine the other four sets first and then return to 上(面) shàng(miàn) and 裏(面) lǐ(miàn). For each LOC1, we will discuss the simple form (“without 面 miàn”) before the complex form (“with 面 miàn”).

3. 下(面) xià(miàn), 前(面) qián(miàn), 後(面) hòu(miàn), 外(面) wài(miàn)

We will first approach the data in terms of meaning extension. This is because we notice that several simple forms have shown regular, stabilized “extended” uses in which the LOC1 no longer retains the spatial sense. For example, xià is often used as part of the construction (在…) 下(zài)…xià ‘under … circumstance’, and both qián and hòu are commonly used in temporal expressions, e.g., 三年前 sān nián qián ‘three years ago’ or 下課後 xià kè hòu ‘after class’. On the other hand, xiàmiàn cannot replace xià in the construction for the conditional reading, nor can qiánmiàn or hòumiàn replace their respective simple counterparts in the temporal expressions. These cases lead us to suspect that the difference between the simple form and the complex form is a semantic/functional one.

3.1. without - 面 miàn

Indeed, the simple form has mostly been used in non-spatial senses for these four LOC1s.

3.1.1. 下 xià

In the 11 tokens of xià, 3 (27%) are for spatial reference (e.g., 放在自己的那個座位 下 fàngzài zìjǐde nàge zuòwèi xià ‘place (it) under one’s one seat’). The rest -- the majority (73%) -- of the tokens (8) denotes a conditional relationship, i.e., “under … circumstances”, and 6 of them go with the locative marker zài to form the discontinuous construction (在…) 下(zài)…xià for this conditional relationship (e.g., 在我們的教育體制下 zài wǒmen de jiàoyù tǐzhì xià ‘under our educational system’). The shift from concrete to abstract relationship is natural from a cognitive point of view. To sum up, xià has mostly been used to refer to the conditional relationship, although it does retain its spatial referring function.

3.1.2. 前 qián

In the 24 qián tokens, only 3 (12.5%) have a spatial reference (e.g., 在台前 zài tái qián ‘in front of the platform’). The rest -- the majority (87.5%) -- of the tokens (21) denotes temporal relationship. The shift from spatial to temporal relationship is natural from a cognitive point of view. The meaning/functional extension is paralleled with greater formal flexibility. The preceding element no longer needs to be a nominal, and the construction can
be further generalized as “X qián”, where X can be a verb (phrase). Thus, we have not only
temporal expressions with an NP, such as 几百年前 jǐ bǎi nián qián ‘several hundred years
ago’, but also those with a VP, such as 结束前 jiéshù qián ‘before (it) end(s)’.

3.1.3. 後 hòu

None of the 21 hòu tokens in our data retains spatial reference; all of them are used to
indicate temporal relationship. Just like the qián case, the host can be nominal, verbal, or
even clausal (e.g., 兩個禮拜後 liǎnggè lǐbài hòu ‘two weeks later’, 畢業後 bìyèhòu ‘after
graduation’, 我去淡水後 wǒ qù Dànsuǐ hòu ‘After I went to Danshui’), and the function of
hòu shows a shift from denoting spatial relationship to temporal relationship.

3.1.4. 外 wài

In the 4 wài tokens, 2 are used for spatial references (e.g., 車廂外 chēxiāng wài
‘outside of the carriage’) and the other 2 indicate a non-spatial relationship: “in addition to”
or “except for…” (e.g., 除了遊記 wài chúle yóujì ‘except for travelogues’).

3.1.5. interim summary

Except for the fourth simple form, wài, whose data are too small to draw any
significant conclusion from, the first three simple forms discussed here show a general
preference for following a host that designates non-spatial reference, although for the first
two items (xià and qián) the spatial use is still available to a certain extent. Thus, generally
speaking, in spoken Taiwan Mandarin, the likelihood for the simple form to be used in
contexts of non-spatial relationship is greater than that of its use in contexts of spatial
orientation. Table 2 summarizes the distributions.

(TABLE 2 here)

3.2. with -面 miàn

In contrast with the simple form, the complex form is mostly used in the spatial sense
for the four LOC1s. As discussed in Section 2, the complex forms can be both a post-
nominal particle and a free word. From Table 1 provided above, it is clear that the free word
use has a much higher frequency than the particle use for all four of the sets to be examined
in this section. In the following discussion of spatial vs. non-spatial references, both the
particle use and the free word use are examined.

3.2.1. 下面 xiàmiàn

There are 32 xiàmiàn tokens. While 75% (24 tokens) of them are used as a free word,
25% (8 tokens) is used as a particle. Among the 32 tokens, only 6% (2 tokens) show non-
spatial references (e.g., 只會吼吼我們這些下面的人, zhǐhuì hǒuhǒu wǒmen zhèxiē xiàmiàn
de rén ‘(he) does nothing but yell at us who are below him (in rank)’) and the majority (94%,
30 tokens) indicates spatial reference (e.g., as free word, 下面還有一張紙耶 xiàmiàn hái yǒu
yīzhāng zhǐ yě ‘There is still a piece of paper underneath’; as particle, 老兵都在樹蔭下面睡
lǎobīng dōu zài shùyīn xiàmiàn shuì ‘The veterans all slept under the trees’). Among the 8
particle tokens, all are used for spatial reference and none for non-spatial reference.

3.2.2. 前面 qiánmiàn

There are 61 qiánmiàn tokens. While 82% (50 tokens) are used as a free word, 18% (11 tokens) are used as a particle. Among the 61 tokens, 25% (15 tokens) show non-spatial references – mostly for temporal reference (就因為你前面一個計畫耽擱了 jìushi yīnwěi nǐ qiánmiàn yíge jìhuà dāngē le ‘that is because your previous project was delayed’), and 75% (46 tokens) refer to spatial relationships (e.g., as free word, 剛好她就坐在前面嘛 gānghǎo tā jiù zuò zài qiánmiàn ma ‘She happened to sit in the front’; as particle, 在我姑媽家前面玩 zài wǒ gūmā jiā qiánmiàn wán ‘play(ed) in front of my aunt’s house’). Among the 11 particle tokens, 10 are used for spatial reference and 1 is for non-spatial reference.

3.2.3. 後面 hòumiàn

There are 52 hòumiàn tokens. While 83% (43 tokens) are used as a free word, 17% (9 tokens) are used as a particle. Among the 52 tokens, 19 % (10 tokens) show non-spatial references – mostly for temporal reference or sequencing relationship (such as in age, rank, and discourse) (e.g., 不給機會給後面的人去學習 búgěi jīhuí gěi hòumiàn de rèn qù xuéxí ‘(they) do not give people of younger age the opportunity to learn’). However, the majority, i.e., 81% (42 tokens), refers to spatial relationships (e.g., as free word, 後面有人舉牌說 hòumiàn yǒu rén jǔ pí shuō ‘In the back someone raised a board and said’; as particle, 我們就直接走到他們後面去看啊 wǒmen jiù zhíjié dào tāmen hòumiàn qù kàn a ‘we went directly behind them to take a look’). Among the 9 particle tokens, 8 are used for spatial reference and 1 is for non-spatial reference.

3.2.4. 外面 wàimiàn

There are 66 wàimiàn tokens. While 94% (62 tokens) are used as a free word, 6% (4 tokens) are used as a particle. Among the 66 tokens, 33% (22 tokens) show non-spatial references (e.g., 讓外面的那一些駭客也沒有辦法進來 ràng wàimiàn de nàyīxiē hàikè yě méiyǒu bànfǎ jìnle ‘so that the hackers outside will not be able to get in’). However, the majority, i.e., 67% (44 tokens), refer to spatial relationships (e.g., as free word, 要在家裡吃還是去外面吃 yào zài jiālǐ chī háishi qù wàimiàn chī ‘(you) want to eat at home or out?’; as particle, 那個擠到會場外面都是 nàge jǐdào huìchǎng wàimiàn dōushì ‘it was so crowded that (there were even people) outside of the conference site’). All of the 4 particle tokens are used for spatial reference and none is for non-spatial reference.

3.2.5. interim summary

All of the four complex forms are more likely to refer to spatial relationship, although non-spatial references are possible. Table 3 summarizes the distributions without distinguishing the free word use and the particle use. Table 4 provides the spatial vs. non-spatial reference distributions for the particle use only. A comparison of Table 3 and Table 4 reveals that the particle use is even more tilted toward spatial references.

(TABLE 3 here)
3.3. Summary

For the four LOC1s discussed here, there appears to be a tendency for the simple form and the complex form to complement each other functionally, though this functional division is not categorical. The simple form (without –màn) tends to be used for non-spatial references, while the complex form (with –màn) tends to be used for spatial references.

So far, it seems the semantic/functional account can explain these occurrences. We will proceed to examine shàng(màn) and lǐ(màn) in the next section.

4. 上(miàn) and 理(miàn)

In Section 2 we pointed out that shàng and lǐ are outstanding in their frequency compared to the other four LOC1s in both the simple and the complex forms (cf. Table 1). They are the focus of this section.

4.1. 上 shàng and 上面 shàngmiàn

4.1.1. 上 shàng

The simple form shàng stands out because of its sheer high frequency compared to the other five LOC1s. There are altogether 288 tokens of shàng in our database.9 Among all of these tokens, 32% (91 tokens) are for spatial relationships (e.g. 我在車上等你 wò zài chē shàng děng nǐ ‘I’ll wait for you in the car’), whereas 68% (197 tokens) are used for non-spatial relationships (e.g., 這事法律上規定的 zhè shì fǎlǜ shàng guīdìng de ‘This was regulated by law’). This conforms to the pattern shared by the other simple forms discussed in 3.1: the non-spatial uses generally outnumber the spatial uses.

However, there is something special about shàng which is not observed in the other four simple LOC1s examined in 3.1: there are a number of shàng expressions which recur in the data. These expressions are either terms with basic bodily or physical orientations, which are often used in daily conversation, such as 身上 shēnshàng ‘on (the) body; with (the person)’ (29 tokens), 路上 lùshàng ‘on the road; on the way’ (12 tokens), and 手上 shǒushàng ‘in (the) hand’ (6 tokens).10 Or they are limiting, modifying expressions, such as 基本上 jīběnshàng ‘basically’ (17 tokens), 事實上 shíshìshàng ‘in fact’ (16 tokens), 實際上 shíjìshàng ‘actually’ (13 tokens), 世界上 shìjièshàng ‘in the world’ (11 tokens), and 感覺上 gǎnjuéshàng ‘(it) feel(s); (it) seem(s)’ (8 tokens). It seems, then, shàng owes its outstanding (token) frequency to the fact that it is a part of some basic vocabulary whose token frequency is typically high in daily speech.

Another related observation, which is also not detectable in the previous four simple LOC1s in 3.1, is that the NP in these frequent expressions is either monosyllabic or disyllabic. In fact, these two types form the majority of the entire shàng database – 32% (91 tokens) for monosyllabic NPs (e.g., 身上 shēnshàng ‘on (the) body; with (the person)’), and 53% (152 tokens) for disyllabic NPs (e.g., 基本上 jīběnshàng ‘basically’). Setting aside the 4 cases that contain an English host word due to code-mixing (1%), cases with a three-or-more syllable...
NP constitute only 14% (41 tokens) of the data (e.g., 毀在一卷錄影带 on huí zài yǐjuàn lùyǐngdài shāng ‘(his reputation) was ruined on a videotape’). The syllable issue becomes significant when shàng and shàngmiàn are compared with each other.

4.1.2. 上面 shàngmiàn

There are 99 tokens for shàngmiàn, which is the second largest group among the six complex particles. Like the four complex forms discussed in Section 3.2, the majority of the shàngmiàn tokens is used as a free word – 60% (59 tokens); the particle use is the minority – 40% (40 tokens). Among the total 99 tokens, while 30% (30 tokens) are used for non-spatial references (e.g., as free word, 上面還在考慮嘛 shàngmiàn háizi kāolü ma ‘The boss/“people of higher rank” is/are still thinking about/considering it’; as particle, 一個人旅 xíngrén lǚ shàngmiàn yào jiǎngjiù shì bǐjiào kǔnnán de ‘It is not easy to have haute cuisine at every meal while traveling solo’), 70% (69 tokens) are used for spatial references (e.g., as free word, 行人可以站在上面 xíngrén kěyǐ zhànzài shàngmiàn kàn ‘passengers can stand on it and look around’; as particle, 就直接把 jiù zhíjié bā tāmen míngzi xiě zài nàge dānzi shàngmiàn ‘just wrote their names directly on that form’). This distribution pattern is consistent with those of the other four complex forms: they all seem to be used more often to refer to spatial relationship than to non-spatial relationship. Notice that this distribution pattern is also in complementary relationship with that of shàng (68% non-spatial vs. 32% spatial) discussed in 4.1.1. Thus, shàng and shàngmiàn appear to complement each other functionally, just like the four sets discussed earlier.

In contrast to shàng, the shàngmiàn token examples do not have any recurring host words except for those apparently arising from the same conversation. In the 40 tokens of shàngmiàn used as a particle, 3 cases have English words as the host (8%). In only 1 case (3%) is the host a monosyllabic word (一個人旅行在吃上面要講究是比較困難的 yíge rén lǚxíng zài shàngmiàn yào jiǎngjiù shì bǐjiào kǔnnán de ‘It is not easy to have haute cuisine at every meal while traveling solo’). While another 30% (12 tokens) is disyllabic (e.g., 浴室上面的空間呀 yùshì shàngmiàn de kōngjiān ya ‘the space above the bathroom’), the remaining 60% (24 tokens) of the data has a three-or-more syllable host expression (e.g., 然後就站在我的肩膀上面 ránhòu jiù zhànzài wǒ de jiānbáng shàngmiàn ‘then stood on my shoulders’).

This is in sharp contrast with the profile of shàng as discussed in 4.1.1. In the case of shàng, 84% (91 + 152 = 243 tokens) of the host expression is either monosyllabic or disyllabic, whereas only 14% (40 tokens) is three-or-more syllables. By contrast, in the case of shàngmiàn, the three-or-more syllable host expression is the majority (60%).

4.1.3. 上 shàng vs. 上面 shàngmiàn

So far we have observed two things. First, the shàng/shàngmiàn pair seems parallel to the other four pairs discussed in Section 3 in showing a (non-categorical) semantic/functional distinction. However, different from the other four pairs, shàng and shàngmiàn manifest another kind of complementary distribution with respect to the syllable structure of the host expression. Shàng tends to go with monosyllabic or disyllabic expressions (84%) while shàngmiàn tends to go with three-or-more syllable expressions (60%). Table 5 summarizes the contrast.
We recall Li and Thompson’s (1981) proposal at this point. While their hypothesis is not supported by our data if taken literally (just as it was rejected in 吕 [Lü] (1965)), there nevertheless seems to be some truth to their hunch. Li and Thompson made the distinction between monosyllable and two-syllable hosts. Our data, however, point to a dividing line at one “level” up: that is, if the monosyllable and the two-syllable hosts are taken as one group and the three-syllable and more-than-three syllable hosts are taken as another group, the contrast between the simple form and the complex form becomes more apparent. It seems shàng prefers going with the first group and shàngmiàn prefers the second. This dividing line (between two syllable units and three syllable units) is justified, given that basic words in Modern Mandarin are largely monosyllabic or disyllabic. Beyond two syllables, the more syllables there are in a unit, the more likely the unit is less of a word and more of a phrase. Indeed, a large number of the three-or-more syllable hosts preceding shàngmiàn in our data are noun phrases consisting of the core noun word plus one of the elements typically occurring in a noun phrase, e.g., determiners, pronouns, classifiers, modifying adjectivals, and the de. The following are two examples:

(3) 然後我覺得我這次錢幾乎都花在那種莫名其妙的事情上面
    Ránhòu wǒ juéde wǒ zhècì qián jìhū dōu huā zài nàzhòng mòmíngqímiào de shìqíng shàngmiàn
    ‘And I think this time my money was mostly spent on trivial things (lit. the kind of things that come from nowhere).’ (non-spatial reference)

(4) 你的身份證上面沒有英文名字？
    Nǐde shēnfēnzhèng shàngmiàn méiyǒu yīngwén míngzi
    ‘There’s no English name on your ID card?’ (spatial reference)

To sum up, shàng and shàngmiàn appear to be associated with two types of polarization: functionally, shàng is more likely to refer to non-spatial relationship while shàngmiàn is more likely to refer to spatial relationship; formally, shàng is more likely to be part of a word, whose stem is typically monosyllabic or disyllabic, whereas shàngmiàn is more likely to be an enclitic (i.e., a postposed clitic) (Hopper & Traugott 1993/2003) to a noun phrase, which typically comprises three or more syllables. The particle shàng probably owes its outstanding high frequency to the high recurring rate of some of the basic expressions of which it is a part. On the other hand, recurrence of the same host is not a feature found in the case of shàngmiàn. This is not surprising given that shàngmiàn is largely used after a mutli-syllable phrase, which is typically constructed online. The contrast between shàng and shàngmiàn, from the collocation point of view, can then be captured in terms of Sinclair’s (1991) contrast between idiom principle and open choice principle: the former is part of some prefabricated expressions – most likely words of high recurrence in daily speech, and the latter is an enclitic marking a phrase constructed in real time.

4.2. 裡 lǐ and 裡面 lǐmiàn

4.2.1. 裡 lǐ
For the 88 tokens of agrid, 48% (42 tokens) refer to spatial relationships (e.g., 擺在冰箱裡 gē zài bīngxiāng lǐ ‘place (it) in the refrigerator’) while 52% (46 tokens) refer to non-spatial relationships (e.g., 我心裡 wǒ xīnlǐ ‘I was thinking’). There does not seem to be any clear preference for either category. This near-even distribution suggests that the semantic factor does not seem to be critical here.

On the other hand, the relationship between agrid and its hosting expression is similar to that of shàng. Some expressions recur in the data. For example, 家裡 jiālǐ ‘at home’ and 屋裡 wūlǐ ‘inside the house’ together make 38 tokens, making up 43% of the total 88 agrid tokens, and 心裡 xīnlǐ ‘in (one’s) heart/mind’ has 11 tokens, taking up another 13% of the data. These highly recurrent words all have a monosyllabic host. Just like the shàng case, the 14 three-or-more-syllable host expressions are in the minority, taking up only 16% of the total 88 agrid tokens (others being 1 English host word, 3 more monosyllabic hosts, and 21 disyllabic hosts). Lumping all monosyllabic host cases (38 + 11 + 3 = 52 tokens), it is clear that agrid favors monosyllabic hosts (59%, 52 out of 88 tokens).

Shàng and agrid share the same pattern in that they gain high frequency because they are part of some frequently used expressions. Both forms also show a small likelihood of pairing up with three-or-more syllable hosts.

4.2.2. 裡面 lǐmiàn

Lǐmiàn is the most frequently used item in all of the six complex particles. Its total, 266 tokens, is 46% of the total 576 miàn tokens working with the six LOC1s in the database. Lǐmiàn is distinct from the other five complex forms in two more aspects: First, while all of the other five complex forms have more free word uses than particle uses, lǐmiàn demonstrates a reverse pattern: 29% (77 tokens) are used as free words while 71% (189 tokens) are used as particles (cf. Table 1).

Second, while all of the other five complex particles show a tendency toward spatial use rather than non-spatial use, lǐmiàn shows an almost balanced distribution: 53% for spatial use (e.g., as free word, 裡面都是泥巴 lǐmiàn dōushì níbā ‘(It is) all mud inside’; as particle, 她現在衣櫃裡面有多少件灰色的啊? tā xiànzài yīzhān lǐmiàn yǒu duōshǎo jiàn huīsè de a? ‘How many grey clothes do you have in your wardrobe now?’), and 47% for non-spatial use (e.g., as free word, 應該可以看看裡面 yīnggāi kānkàn lǐmiàn ‘(we) should be able to take a look at what’s in there (lǐmiàn here refers to an online hyperlink)’; as particle, 電影裡面的蘇俄軍人 diànyǐng lǐmiàn de Sū-è jūnrén ‘The Soviet soldiers in the movie’).

In other words, compared to other complex forms, lǐmiàn (1) is more likely to be used as a particle than as a free word, and (2) shows no preference in distribution for either spatial or non-spatial references.

On the other hand, for lǐmiàn, the syllable structure of the preceding host seems skewed. In the 189 tokens of lǐmiàn as particle, 12 cases have an English host (6%). Monosyllabic hosts are a minority (14%, 27 tokens) (e.g., 裡面裡一定有兩個人嘛 lǐmiàn lǐmén yídìng yǒu liǎnggè rén ma ‘There must be two people in the store’). Then, disyllabic hosts take up another 34% (64 tokens) (e.g., 電影裡面的蘇俄軍人 diànyǐng lǐmiàn de Sū-è jūnrén ‘The Soviet soldiers in the movie’). Finally, three-or-more syllable hosts constitute
the largest percentage – 46% (86 tokens) (e.g., 在很現代的房子裡面 zài hěn xiàn dài de fángzi lǐ miàn ‘in very modern houses’). Together, multi-syllabic hosts constitute 79% of the data (64 + 86 = 150 tokens).

4.2.3. 裡 lǐ vs. 裡面 lǐ miàn

When we compare lǐ and lǐ miàn, there does not seem to be anything semantic or functional that distinguishes the two terms. Both forms show near-even distributions for spatial and non-spatial references. This is different from the case of shàng and shàng miàn since the latter two forms are still functionally complementary.

However, just like the shàng/shàng miàn pair, lǐ and lǐ miàn form a sharp contrast in the syllable structure of their host NP. As particle, lǐ prefers monosyllabic hosts, while lǐ miàn prefers multi-syllabic hosts – if not particularly three-or-more syllable hosts. Table 6 provides the summary.

(TABLE 6 here)

The preference for three-or-more syllable hosts suggests that lǐ miàn is a particle that is flexible in combination with not only a word (typically mono-syllabic or disyllabic) but also with a phrase (three-or-more-syllabic). In other words, the distinction between lǐ and lǐ miàn seems to manifest the “idiom principle vs. open choice principle” contrast even better than shàng and shàng miàn do.

4.3. Parallels between 上 shàng/上面 shàng miàn and 裡 lǐ/裡面 lǐ miàn

If we put the statistics together, we find a parallel between the two sets, shàng/shàng miàn and lǐ/lǐ miàn, in terms of their co-occurrence patterns with their host. For both shàng and lǐ, the three-or-more syllable hosts are the least likely to co-occur. On the other hand, for both shàng miàn and lǐ miàn, the three-or-more syllable host is their favorite host. Thus, in both pairs, the complex form is more likely than the simple form to be used as an enclitic to a phrasal host, and the simple form is more likely than the complex form to be used as an enclitic/suffix to a word host or a morpheme host. In many cases, the simple form is virtually a word-internal element (詞內成分 cí nèi chéngfèn) (cf. 董[Dong] 2004). Such further grammaticalization – from a clitic to an affix to a word-internal element -- is especially conspicuous in cases where the host is a monosyllabic morpheme likely to be used repeatedly in daily speech, such as in 身上 shēnshàng ‘on (the) body; with (the person)’, and 家裡 jiālǐ ‘at home’, etc.

Having identified the preference patterns, it is nevertheless important to be reminded again that these are just tendencies after all and the distinction is far from categorical. The morpheme/word vs. phrase polarization is not neatly reflected in the syllable structure. Many three-syllable hosts are in fact compound words but not phrases (e.g., 運動會上面 yùndòng huì shàng miàn ‘at the athletic game’; 圖書館裡面 tūshūguǎn lǐ miàn ‘in the library’), and some disyllabic hosts are phrases (e.g., 他班上 tā bān shàng ‘in his class’; 他家裡面 tā jiā lǐ miàn ‘in his house/family’). Nevertheless, our data show that for shàng(miàn) and lǐ(miàn), the disyllabic hosts are still overwhelmingly words rather than phrases, and at least 65% of the three-or-more syllable hosts are phrases. See Table 7 for summary.
5. General Discussion

5.1. 下面 xià(miàn), 前面 qián(miàn), 後面 hòu(miàn), 外面 wài(miàn) revisited

In Section 4 a morphological/structural account is offered to explain the simple and complex choices with regard to 上面 shàng(miàn) and 里面 lǐ(miàn), which is quite different from the semantic/functional account offered to explain the other four sets in Section 3. A reasonable question at this point is whether it is the case that the morphological account can actually work for the earlier four sets as well? In other words, among the earlier four pairs, does the simple form prefer monosyllabic hosts and the complex form prefer three-or-more syllabic hosts, like the shàng(miàn) and lǐ(miàn) pairs? Our data, however, do not support this proposal. Table 8 shows that for the eight forms there is an across-the-board preference for disyllabic or three-or-more syllable hosts. No distribution patterns can be found to differentiate the simple and the complex alternatives in any of the four pairs. Thus, we are quite confident that the structural account does not affect the earlier four sets.

5.2. 上面 shàng(miàn) and 里面 lǐ(miàn) revisited

So far in our discussion we have been lumping shàng(miàn) and lǐ(miàn) together to make a contrast with the other four sets. Nonetheless, despite that they share the same morphological/structural distributional patterns, these two sets are quite different in terms of the paradigmatic relationship between the simple and the complex forms.

The “competition” between lǐ and limiàn is relatively straightforward. As discussed before, in contrast with the other complex forms, limiàn is used as a particle more often than as a free word (cf. Table 1). From the perspective of grammaticalization, limiàn seems to be leading its progression on the cline of “word → enclitic → suffix” (Hopper and Traugott 1993/2003) at a speedier pace than the other five complex forms.

If we put all the statistics for lǐ and limiàn (as particle) together, as in Table 9, the dominance of limiàn becomes apparent. As limiàn is well developed into an enclitic, it dominates the particle slot after both disyllabic and three-or-more syllable hosts, no matter whether these hosts are words or phrases. The only slot that lǐ dominates over limiàn is the monosyllabic host. Recall that some of the monosyllabic hosts for lǐ recur, such as in 家裡 jiālǐ ‘at home’, 房裡 fánglǐ ‘inside the house’ and 心裡 xīnlǐ ‘in (one’s) mind’ (cf. 4.2.1). In fact, the token total of these three expressions is 49, which is 94% of the entire monosyllabic host data (52 tokens in total). There are only three other different hosts. In other words, in the one-syllable host category, while the token frequency is quite high, the type frequency -- varieties of hosts -- is quite low. This fact contributes to the dominance of lǐ (over limiàn) since (i) lǐ is a word-internal element in these three words and is not, therefore, easy to replace, and (ii) these three words happen to occur with high frequency in daily speech.
The relationship between *shàng* and *shàngmiàn* is more complicated. *Shàng* and *shàngmiàn* maintain a rough semantic/functional divide: *shàng* prefers non-spatial references while *shàngmiàn* prefers spatial references.

On the other hand, the morphological/structural account seems equally plausible. Different from *limiàn*, *shàngmiàn* is still more often used as a free word than as a particle (cf. Table 1). More significantly, as a particle, *shàngmiàn* is in no competition to the simple *shàng* in terms of sheer frequency. Furthermore, when we put all the statistics for *shàng* and *shàngmiàn* (as particle) together, as in Table 10, it is apparent that *shàngmiàn* is not in a dominant position in any but the three-or-more syllable word category. It appears, then, that *shàngmiàn*, compared to *limiàn*, is still at an earlier stage on its progression on the cline of “word → enclitic → suffix”. It has begun the journey all right, but it has not attained a stable status as an enclitic (let alone as a suffix) in competition with *shàng*.

(TABLE 10 here)

The simple form, *shàng*, is still dominant in most of the categories, including as an enclitic for three-or-more syllable phrasal hosts. In terms of type frequency, *shàng* also demonstrates more variety than *lì*. For example, in the monosyllabic host category, despite the fact that there are recurring expressions such as *shàng* in (the) body; with (the person), *lì* on the road; on the way, and *shàng* in (the) hand, there are 19 other hosts in the total 91 tokens. In the disyllabic word category, in addition to the recurring *shàng* ‘basically’, *shàng* ‘in fact’, *shàng* ‘actually’, *lì* in the world’, and *shàng* ‘(it) feel(s); (it) seem(s)’, there are also 51 other hosts in the total 150 tokens. It appears that *shàng* is strong and well as an enclitic and a (suffix turned) word-internal element.

Both the semantic/functional and the morphological/structural accounts, therefore, work for the *shàng(miàn)* set. However, it seems the morphological/structural account is the key to the differentiation while the semantic/functional account is less convincing. The reason is that, unlike the other four simple forms, *shàng* (and also *lì*) does not have stabilized, non-spatial senses that cannot be expressed by the complex form, *shàngmiàn* (or *limiàn*). As discussed in Section 3, the other four simple forms can be used as temporal or conditional particles in a frame with a structurally flexible host (X) (e.g., 在 X, 前 X, 後 X, (除了) X 外 (chúle) X). Their respective complex counterparts cannot replace them in these frames. In the case of *shàng* (and of *lì*), no such meaning frames (e.g., temporal or conditional, etc.) have been developed. However, we certainly need to check more data before any definitive conclusion can be drawn.

6. Conclusion

6.1. Summary

For the differentiation between simple locative particle and complex locative particle, we have identified two types of development along two different tracks. First, there is a semantic/functional distinction for the *xià(miàn), qián(miàn), hòu(miàn), vài(miàn)* sets, in which the simple form is largely for non-spatial, extended references, and the complex form is largely for spatial references. However, meaning differentiation does not explain the *lì(miàn)* set. Instead, there is a morphological/structural divide, in which the simple *lì* largely
works not only as a word enclitic/suffix but even as a word-internal element, and the complex *lǐmiàn* works as an enclitic attached to words or phrases. Finally, the *shàng(miàn)* set seems to be open to both accounts but we believe the morphological/structural account better explains the relationship between *shàng* and *shàngmiàn*, which is similar to that of *lǐ* and *lǐmiàn*. However, the *shàng(miàn)* set and the *lǐ(miàn)* set are not exactly the same: *shàng* holds strongly as a particle (enclitic, suffix, or even word-internal element), and *shàngmiàn* shows positive but only incipient signs of becoming a particle. By contrast, *lǐ* holds strongly as a word-internal element in some high frequency words but is losing ground to *lǐmiàn*, whose particle status has stabilized with its sheer occurrence frequency and its adaptability in working with hosts of both a word form and a phrasal form, and of all types of syllable structure.

It is important to note again that both tracks of differentiation are not stabilized but still evolving, so these are just distribution patterns rather than absolute categorical distinctions. From the perspective of grammaticalization, the fact that the two types of particles are found to have overlapping functions and occur after the same types (semantic and/or morphological) of host expression is by no means surprising, as language change takes time and during the process the newly emerged layer of uses coexists with the older layer of uses, forming what Hopper (1991) calls the “layering” phenomenon. Again from the perspective of grammaticalization, it is also reasonable to find that, in general, the simple form is more “advanced”, both functionally (for non-spatial, extended references) and structurally (as a word-internal element), than the complex form, as historically the simple forms started their grammaticalization much earlier than the complex forms (劉 [Liu] 2003; Sun ms.).

6.2. Directions for future research

The data size needs to be expanded for future research. In the present study, not much can be concluded for the profile of the *wài(miàn)* set due to inadequate data. Also, whether *shàng* and *shàngmiàn* maintain a functional distinction is undetermined. A more thorough understanding would require a database much larger than the one employed here.

We limited ourselves to the discussion of only the simple form and the complex form in this study, but of course the reality of how spatial-oriented terms behave in daily speech is more complicated than the small sketch drawn here. For one thing, other items are interchangeable with some of the forms discussed here. For example, technically, *nèi* ‘inside’, 之內 *zhīnèi* ‘inside’, 之中 *zhīzhōng* ‘in (the middle of)’, and 之中 *zhīzhōng* ‘in (the middle of)’, are equivalents of *lǐ/lǐmiàn* (also cf. footnote 5). A thorough investigation would have to include these items.

In Taiwan Mandarin, the LOC2 for the six spatial terms has been narrowed down to *miàn*. However, in other varieties of Modern Chinese the LOC2 may be *tóu* or *biàn* or multiple alternatives. A cross-dialectal investigation would be necessary to further clarify the issue.

Finally, a comparison between speech and writing would add another dimension to the investigation. 吕 [Lü] (1965) states that the complex forms are not frequently found in his written data. Is this still the case in present-day Chinese writing? If writing is indeed different from speech, what is the reason? Is there a different “grammar” to account for how LOC1 and LOC2 are used in Chinese writing?
References


Sun, Chaofen. Ms. Two conditions and grammaticalization of the Chinese locative.

Endnotes

∗ I would like to thank the editor of the volume, Professor Janet Xing, for her coordination, encouragement, and input. I would also like to thank the anonymous reviewer for his valuable feedback on the draft of this paper. An earlier version of this paper was presented at
the International Symposium on Corpus and Discourse Approaches to Chinese Linguistics, held at UCLA in June 2007. I am grateful to the audience, and especially to Professor Sun Chaofen, for their comments and suggestions, which helped me to revise the paper. I am, of course, responsible for all the errors and inadequacies remaining in the article.

1 A small number of spatial orientation terms, such as 中間 zhōngjiān ‘in the middle’ and 底下 dǐxià ‘below, beneath’, would not go along with any of the optional particles. These cases do not concern us and are excluded from our discussion.

2 The complex form can also be used as a free word serving as a modifying element in a noun phrase (e.g., 門外的門 wàimén de mén ‘the door outside’). In this study, however, we do not make further distinctions but lump all of the free word uses together into one group – the free word category -- as opposed to the other category – the particle category.

3 Both the simple and the complex locative expressions can be preceded by locative prepositions, e.g. 往下看 wǎng xià kān ‘look down’, 從外面進來 cóng wàimén jīnlái ‘come in from outside’. This type of use is also excluded from our discussion.

4 Whether locative particles are postpositions or enclitics has been controversial in Chinese (historical) syntax (e.g., Li & Thompson 1981; 劉[Liu] 2003; Sun ms.) but the issue is beyond the scope of this paper. Theoretically neutral terms, e.g., “particles” or “expressions”, are thus used to refer to the forms in this paper.

5 For easier reading, the intonation details in the spoken data are not provided here. The gloss conventions used in this paper are: CLFT—shì . . . de cleft construction, NEG—negation, POSS – possessive, PRT – sentence-final particle.

6 The suffix –邊 biān also co-occurs frequently with the determiners, 這 zhè and 那 nà, and the question word 哪 nǎ, forming spatial deictic nominals, 這邊 zhèbiān ‘here’, 那邊 nàbiān ‘there’ and 哪邊 nǎibiān ‘where’. See 魏[Bi] (2007) for the incipient idiomatization of the distal locative phrase 在那邊 zài nàbiān ‘over there’ in spoken Taiwan Mandarin.

7 In the rest of the paper (except for section headings), Chinese characters are not provided for the six sets of key terms when they are mentioned in the text. However, Chinese characters are provided along with Pinyin for example phrases or sentences.

8 Each of the six LOC1s participates as a morpheme in some other disyllabic spatial terms. For example, the following expressions all involve xià: 之下 zhīxià (6 tokens), 以下 yǐxià (6 tokens), and 底下 dǐxià (5 tokens). All three expressions can be used as a particle and each has its own distinct meaning: 之下 zhīxià (just like xià) for conditional relationship (e.g., 在相同平等的條件之下 zài xiāngtóng píngděng de tíáojiànxīxià ‘under the same, equal conditions’), 以下 yǐxià for quantity and degree (e.g., 六歲以下 liùsuìyǐxià ‘under six years old’), and 底下 dǐxià for spatial relationship (e.g., 桌子底下 zhuōzǐdǐxià ‘under the table’). While these disyllabic expressions can be particles as the examples show, they cannot work further with the optional LOC2s such as -miàn. Therefore, we do not discuss these cases (across the six LOC1s) in our study.
The particle use of *shàng* here refers strictly to those cases in which it follows a nominal host. Therefore, its other uses, e.g., as verbal complement, are excluded from consideration here.

Notice that while these expressions have spatial reference as their basic meaning, they can all refer to abstract, non-spatial relationship, too. In our analysis, we identify the spatial vs. non-spatial categorization according to the interpretation most appropriate for the context, not according to the meaning of the host expression.

### TABLES

<table>
<thead>
<tr>
<th></th>
<th>Simple TOTAL (without –miàn)</th>
<th>Complex- particle (with –miàn)</th>
<th>Complex-free word (with –miàn)</th>
<th>Complex TOTAL (with –miàn)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shàng</td>
<td>288</td>
<td>40</td>
<td>59</td>
<td>99</td>
</tr>
<tr>
<td>xià</td>
<td>11</td>
<td>8</td>
<td>24</td>
<td>32</td>
</tr>
<tr>
<td>qián</td>
<td>24</td>
<td>11</td>
<td>50</td>
<td>61</td>
</tr>
<tr>
<td>hòu</td>
<td>21</td>
<td>9</td>
<td>43</td>
<td>52</td>
</tr>
<tr>
<td>lǐ</td>
<td>88</td>
<td>189</td>
<td>77</td>
<td>266</td>
</tr>
<tr>
<td>wài</td>
<td>4</td>
<td>4</td>
<td>62</td>
<td>66</td>
</tr>
</tbody>
</table>

**TABLE 1.** token totals of the six LOC1s (without –miàn and with –miàn)

<table>
<thead>
<tr>
<th></th>
<th>Spatial (a)</th>
<th>non-spatial (b)</th>
<th>TOTAL (a+b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xià</td>
<td>3</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>qián</td>
<td>3</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>hòu</td>
<td>0</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>wài</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

**TABLE 2.** spatial vs. non-spatial references for *xià, qián, hòu,* and *wài*

<table>
<thead>
<tr>
<th></th>
<th>Spatial (a)</th>
<th>non-spatial (b)</th>
<th>TOTAL (a+b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xiàmiàn</td>
<td>30</td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>qiánmiàn</td>
<td>46</td>
<td>15</td>
<td>61</td>
</tr>
<tr>
<td>hòumiàn</td>
<td>42</td>
<td>10</td>
<td>52</td>
</tr>
<tr>
<td>wàimiàn</td>
<td>44</td>
<td>22</td>
<td>66</td>
</tr>
</tbody>
</table>

**TABLE 3.** spatial vs. non-spatial references for *xiàmiàn, qiánmiàn, hòumiàn,* and *wàimiàn* as both free word and particle
### Table 4. Spatial vs. Non-spatial References for xiàmiàn, qiánmiàn, hòumiàn, and wàimiàn as Particle Only

<table>
<thead>
<tr>
<th></th>
<th>Spatial (a)</th>
<th>Non-spatial (b)</th>
<th>TOTAL (a+b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xiàmiàn</td>
<td>8 100%</td>
<td>0 0%</td>
<td>8</td>
</tr>
<tr>
<td>qiánmiàn</td>
<td>10 91%</td>
<td>1 9%</td>
<td>11</td>
</tr>
<tr>
<td>hòumiàn</td>
<td>8 89%</td>
<td>1 11%</td>
<td>9</td>
</tr>
<tr>
<td>wàimiàn</td>
<td>4 100%</td>
<td>0 0%</td>
<td>4</td>
</tr>
</tbody>
</table>

**Total:** (a+b+c)

Table 4 shows the distribution of spatial and non-spatial references for the particles xiàmiàn, qiánmiàn, hòumiàn, and wàimiàn. The table highlights the percentage of 1-syllable, 2-syllable, and 3-or-more-syllable references for each particle.

### Table 5. Syllable Structure of the Host for shàng and shàngmiàn (as Particle)

<table>
<thead>
<tr>
<th></th>
<th>1-syll. (a)</th>
<th>2-syll. (b)</th>
<th>3-or-more syll. (c)</th>
<th>TOTAL (a+b+c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>shàng</td>
<td>91 32%</td>
<td>152 53%</td>
<td>41 14%</td>
<td>284 (288 - Engl. 4)</td>
</tr>
<tr>
<td>shàngmiàn</td>
<td>1 3%</td>
<td>12 30%</td>
<td>24 60%</td>
<td>37 (40- Engl. 3)</td>
</tr>
</tbody>
</table>

Table 5 presents the syllable structure of the host for shàng and shàngmiàn. The table indicates the distribution of 1-syllable, 2-syllable, and 3-or-more-syllable words and phrases for each particle.

### Table 6. Syllable Structure of the Host for lǐ and lǐmiàn (as Particle)

<table>
<thead>
<tr>
<th></th>
<th>2-syll. word (a)</th>
<th>2-syll. phrase (b)</th>
<th>2-syll. TOTAL (a+b)</th>
<th>3-or-more syll. word (c)</th>
<th>3-or-more syll. phrase (d)</th>
<th>3-or-more syll. TOTAL (c+d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>shàng</td>
<td>150 99%</td>
<td>2 1%</td>
<td>152</td>
<td>3 7%</td>
<td>38 93%</td>
<td>41</td>
</tr>
<tr>
<td>shàngmiàn</td>
<td>12 100%</td>
<td>0 0%</td>
<td>12</td>
<td>8 33%</td>
<td>16 67%</td>
<td>24</td>
</tr>
<tr>
<td>lǐ</td>
<td>20 95%</td>
<td>1 5%</td>
<td>21</td>
<td>4 29%</td>
<td>10 71%</td>
<td>14</td>
</tr>
<tr>
<td>lǐmiàn</td>
<td>52 81%</td>
<td>12 19%</td>
<td>64</td>
<td>30 35%</td>
<td>56 65%</td>
<td>86</td>
</tr>
</tbody>
</table>

Table 6 shows the syllable structure of the host for lǐ and lǐmiàn. The table details the distribution of 2-syllable words, phrases, and total syllables for each particle.

### Table 7. Word Hosts and Phrasal Hosts for shàng(miàn) and lǐ(miàn)

<table>
<thead>
<tr>
<th></th>
<th>1-syll. (a)</th>
<th>2-syll. (b)</th>
<th>3-or-more-syll. (c)</th>
<th>TOTAL (a+b+c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xià</td>
<td>1 0</td>
<td>10</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>xiàmiàn</td>
<td>1 5</td>
<td>2</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>qián</td>
<td>2 13</td>
<td>9</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>qiánmiàn</td>
<td>2 3</td>
<td>5</td>
<td></td>
<td>10 (11 - Engl. 1)</td>
</tr>
<tr>
<td>hòu</td>
<td>1 8</td>
<td>12</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>hòumiàn</td>
<td>1 3</td>
<td>4</td>
<td></td>
<td>8 (9 - Engl. 1)</td>
</tr>
<tr>
<td>wài</td>
<td>0 2</td>
<td>2</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>wàimiàn</td>
<td>0 2</td>
<td>1</td>
<td></td>
<td>3 (4 - Engl. 1)</td>
</tr>
</tbody>
</table>

Table 7 lists the word hosts and phrasal hosts for shàng(miàn) and lǐ(miàn). The table provides the number of 1-syllable, 2-syllable, and 3-or-more-syllable hosts for each particle.

### Table 8. Syllable Structure of the Hosts Preceding xià(miàn), qián(miàn), hòu(miàn), and wài(miàn)

<table>
<thead>
<tr>
<th></th>
<th>1-syll. (a)</th>
<th>2-syll. (b)</th>
<th>3-or-more-syll. (c)</th>
<th>TOTAL (a+b+c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xià</td>
<td>1 0</td>
<td>10</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>xiàmiàn</td>
<td>1 5</td>
<td>2</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>qián</td>
<td>2 13</td>
<td>9</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>qiánmiàn</td>
<td>2 3</td>
<td>5</td>
<td></td>
<td>10 (11 - Engl. 1)</td>
</tr>
<tr>
<td>hòu</td>
<td>1 8</td>
<td>12</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>hòumiàn</td>
<td>1 3</td>
<td>4</td>
<td></td>
<td>8 (9 - Engl. 1)</td>
</tr>
<tr>
<td>wài</td>
<td>0 2</td>
<td>2</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>wàimiàn</td>
<td>0 2</td>
<td>1</td>
<td></td>
<td>3 (4 - Engl. 1)</td>
</tr>
</tbody>
</table>

Table 8 details the syllable structure of the hosts preceding xià(miàn), qián(miàn), hòu(miàn), and wài(miàn). The table shows the distribution of 1-syllable, 2-syllable, and 3-or-more-syllable hosts for each particle.
<table>
<thead>
<tr>
<th></th>
<th>1-syll. (a)</th>
<th>2-syll. word (b)</th>
<th>2-syll. phrase (c)</th>
<th>3-or-more syll. word (d)</th>
<th>3-or-more syll. phrase (e)</th>
<th>TOTAL (a+b+c+d+e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ropolis</td>
<td>52 66%</td>
<td>20 28%</td>
<td>1 8%</td>
<td>4 12%</td>
<td>10 15%</td>
<td>87 (88 – Engl. 1)</td>
</tr>
<tr>
<td>limiàn</td>
<td>27 34%</td>
<td>52 72%</td>
<td>12 92%</td>
<td>30 88%</td>
<td>56 85%</td>
<td>177 (189 – Engl. 12)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>79</td>
<td>72</td>
<td>13</td>
<td>34</td>
<td>66</td>
<td>264</td>
</tr>
</tbody>
</table>

TABLE 9. ropolis vs. limiàn (as particle) in relation to the syllable and morphological structure of the host

<table>
<thead>
<tr>
<th></th>
<th>1-syll. (a)</th>
<th>2-syll. word (b)</th>
<th>2-syll. phrase (c)</th>
<th>3-or-more syll. word (d)</th>
<th>3-or-more syll. phrase (e)</th>
<th>TOTAL (a+b+c+d+e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>shàng</td>
<td>91 99%</td>
<td>150 93%</td>
<td>2 100%</td>
<td>3 27%</td>
<td>38 70%</td>
<td>284 (288 - Engl. 4)</td>
</tr>
<tr>
<td>shàngmiàn</td>
<td>1 1%</td>
<td>12 7%</td>
<td>0 0%</td>
<td>8 73%</td>
<td>16 30%</td>
<td>37 (40 - Engl. 3)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>92</td>
<td>162</td>
<td>2</td>
<td>11</td>
<td>54</td>
<td>321</td>
</tr>
</tbody>
</table>

TABLE 10. shàng vs. shàngmiàn (as particle) in relation to the syllable and morphological structure of the host