Sentence-Final Particle and Illocutionary Acts in Mandarin Learners: A Corpus-based Study of *Ma*

Rui He[1]
Universitat Pompeu Fabra
Ying Zhang
Universitat Pompeu Fabra
Keyi Wei
Universitat Pompeu Fabra
Ning Xu
Universitat Pompeu Fabra

Abstract

Ma(嘛) is an underexploited but frequently-used particle in Mandarin Chinese. Due to its frequency in daily talk and online chatting, it is important for Chinese learners to correctly understand the pragmatic function and illocutionary meaning of ma. In this study, under the framework of speech act theory, we analyzed the illocutionary acts of sentence-final particle ma by Chinese learners based on corpus data extracted from Guangwai-Lancaster Chinese Learner Corpus. Fisher's exact tests were conducted to explore the relationship between metadata and illocutionary acts, with results showing that all differences did not have statistical significance. We first concluded that ma was an informal particle learned at the intermediate level and mastered at the advanced level. Gender differences in using ma were also discussed, and based on this discussion, it was indicated that the function of ma resembled gender identity to some extent. Then we claimed that the representative act was the general and basic act of ma, followed by the directive and that ma was highly unlikely to appear with declarations. Finally, we discussed its implications for language teaching.

1. Introduction

Sentence-final particles (SP, also called modal particles), though seem to be at the periphery of the language system, prove to have diversified pragmatic functions and illocutionary forces, especially for Asian languages like Japanese (Kakegawa, 2009), Vietnam (Mac et al., 2015), Korean (Ahn, 2016), Cantonese (Sybesma & Li, 2007) and Mandarin (Dong et al., 2018). This has attracted the attention of linguists. Studying SP in language learners' data is a heated trend in SP studies with the popularization of these languages among nonnative speakers. As for Mandarin Chinese, Badan and Romagnoli (2020) studied the acquisition of four Chinese SPs (ma, a, ba, ne) among Italian learners and found that the proficiency level was directly related to the mastery of SP among Italian learners. Yan (2018) studied the behaviors of using seven SPs (le, ne1, ma, ne2, ba1, ba2, and a) among English speakers in L2 Chinese and found the cross-CP-domain features of these SPs made it difficult for learners to master their usage. These studies have laid a solid theoretical and practical foundation for studying SP among learners' language usage. However, we observed there were two main limitations. The first one

---

[1] Correspondence: Rui He, rui.he@upf.edu

was that most studies focused on most frequently used SPs like ba, le, and ma1, while other SPs, such as ma, ei3, and ba4, le, were underexploited, though they are also widely used, especially on the Internet or in casual chatting. Another one is that most of these studies (especially quantitative studies) are based on language materials from a limited number of participants, making their results less general. Corpus-based studies, which becomes increasingly popular these years, provides a possible way to deal with this issue. Xu et al. (2019) used a learner's corpus with both written and spoken data to study the frequency and errors in using SP le by Chinese learners. Lepadat (2017) studied ma using an academic Chinese corpus and discovered the co-occurrence of it with information that has an active or accessible status in the interlocutors' minds and that it functioned like discourse markers.

Based on those previous studies, and motivated by the idea of exploring the pragmatic function of ma(嘛), an underexploited SP which is often used in daily chatting, we used a Chinese leaner corpus collaboratively built by Lancaster and Guangwai and available on Sketch Engine to obtain language data. As illocutivity is regarded as an important function of SP in Chinese (Lepadat, 2017; Qi, 2002), we took Speech Act Theory (SAT) as the theoretical framework and analyzed the illocutionary acts of ma. Based on these data and analysis, we want to answer these two questions: (1) When the speaker is using ma, which kind(s) of illocutionary acts are shown with a higher frequency within utterances with ma? (2) Are there differences between native speakers and Chinese learners, or among different groups of Chinese learners?

2. Methodology

Speech act theory (SAT) was first proposed by Austin John Langshaw (1962) to explain the relation of speech acts and truth. He sketched three speech act types in his framework, that is, locutionary acts, illocutionary acts, and perlocutionary acts. Austin's SAT framework was questioned by John Searle, who raised a new three-type categorization in his one book and later developed it into the currently commonly used five-type division of illocutionary acts (Searle, 1969, 1975) based on conditions illocutivity. To be specific, Searle's five types of illocutionary acts are representative, directive, commissive, expressive, and declaration. A detailed description of the conditions for each type are listed in Table 1 (Searle, 1976; XU et al., 2020).

Table 1
Searle's Classification of Illocutionary Acts

Type	Point	Direction of Fit	Sincerity conditions
representative	Commit S to truth of p	utterance → objective world	speaker believes what s/he says is true
directive	S wants H to do something	objective world → utterance	speaker motivates the addresses to act

commissive	Commits S to future action	objective world → utterance	speaker intends to act
expressive	S's attitude towards p	suppose: objective world → utterance	speaker's attitude is expressed in the utterance
declaration	S utters p to make p true	objective world → utterance	/

2.1.   Data collection and processing

In this study, linguistic data are obtained from the Guangwai-Lancaster Chinese Learner Corpus (GLCLC) in Sketch Engine (Kilgarriff et al. 2014). GLCLC consists of 1,294,714 million words, with a balanced division of written sub-corpus (GLCLC_W) (625,329 tokens, 50.38%) and spoken sub-corpus (GLCLC_S) (642,385 tokens, 49.62%) (Xu et al. 2019). It is construed in more complex language contexts (McEnery et al. 2019) and contains a full set of metadata, so useful for exploration on the correlation between demographic features and language usage. This enables us to make comparisons on language usage between natives and learners.

2.2.   Data and Metadata

In the Concordance function of Sketch Engine, we use the CQL *[word = "嘛"] and [tag = "SP"]* to extract sentences we need. As in our study, 337 out of all 365 sentences are from GLCLC_S (dialogues), with 16 from GLCLC_W and 12 from GLCLC_S (monologue). We found the ratio of sentences from oral materials was overwhelmingly high, within which the majority of data were from dialogues. We, thus, only analysed the data from dialogues.

Nearly all of the metadata of native speakers were not given, and the metadata of Test Score was not available for most non-native speakers. Error metadata, which the corpus was famous for, was neither available for the most sentence we got in this study. So in our study, we select the metadata of Country, First Language, Gender, Genre, and Proficiency. As participants are from dozens of countries and speaking dozens of languages, it's meaningless to analyze based on these data themselves. Therefore, we group the countries by region into four categories, i.e. African, American, European, and Asian countries; and we group the languages by their grammar structures into three categories, analytic languages, agglutinating languages, and fusional languages.

2.3.   Illocutionary Act Labelling

As for illocutionary act labels, two of us first gave labels to each sentence independently. If the two assign the same labels, then that label is taken as the final label. If not, the third person will look at those sentences independently and gave his/her labels, and then picked the most agreed one as the final label. If it is still not agreed, then the fourth person will engage and initiate a discussion to reach a consensus. Below we give an example of each category of the illocutionary act:

(1) Representative, from Nigeria intermediate

我　　有时候　　　来的　　　　　嘛。
Wǒ　yǒushíhòu　láide　　　　ma.
I　　sometimes　come here　PART.

(2) Directive, from Korean beginner

你，　　加油　　　嘛。
Ní,　　jiāyóu　　ma.
You,　come on　PART.

(3) Commissive, from native speaker

我　　给你　　　打电话　　　　　嘛。
wǒ　gěinǐ　　dǎdiànhuà　　ma.
I　　willyou　call　　　　　　PART.
"I will call you to go together."

(4) Expressive, from Korean advanced

喜欢　　　用　　　　那种　　　语言　　　嘛
xǐhuān　yòng　　nàzhǒng　yǔyán　　ma.
Like　　to use　that　　　language　PART.

2.4. Statistical Methods

We used SPSS 26.0 to conduct statistical analysis. We first made a general description of the metadata and illocutionary acts. Then, we conducted chi-square tests to see whether the illocutionary acts of ma differ in different demographic groups. As for all chi-square tests, the minimum expected counts were less than 5, we chose to use the results of Fisher's exact test. The Fisher's exact test has been adapted to fit table larger than 2×2.

3. Results

3.1. General Description

We obtained 337 sentences from GLCLC_S (dialogue). In those sentences, 89 (26.4%) were from native speakers of Mandarin, whose roles are labelled as "native speaker" or "examiner" and 248 (74.6%) were from non-native speakers. For learners' data, 92 (37.1%) were from male participants while 156 (62.9%) were from female participants. Most of these sentences were produced by Asian participants (189, 76.2%), and then by African participants (49, 19.8%). Only 9 were produced by participants from Europe and only 1 by participants from the Americas (Latin America). Regarding their mother tongue, people speaking analytic languages produced 104 samples (41.9%), those speaking agglutinating languages produced 89 samples (35.9%), while those speaking fusional languages produced 52 samples (21.0%). Another 3 samples came from bilingual participants. Regarding genres of the source text, 184 (74.2%) of the sentences came from the oral interview; 29 (11.7%) came from oral instruction and 35 (14.1%) came from oral exams. The majority of non-native speakers using ma has reached advanced Chinese level (217, 87.5%), followed by intermediate level (19, 7.3%) and then beginner level (13, 5.2%).

3.2. Illocutionary Acts of 'ma'

In this study, we analysed the illocutionary acts of sentence-final particle *ma* in the 337 sentences we obtained from the corpus with a comparison between

natives and non-natives. Results are shown in Table 2. In all samples, 89 (26.4%) of them were produced by native Chinese speakers, among which 74 samples (83.1%) expressing representative acts, 9 samples (10.1%) expressing directive acts, 3 samples (3.4%) expressing commissive acts, and 3 samples (3.4%) expressing expressive illocutionary acts. For learner data, we have obtained 248 samples, among which 227 samples (91.5%) expressing representative acts, 11 samples (4.4%) expressing directive acts, 4 samples (1.6%) expressing commissive acts, and 6 samples (2.4%) expressing expressive illocutionary acts. No declaration act was detected from either native speakers or Chinese learners. As is shown by the chi-square test result, there is no significant difference between native speakers and learners in the illocutionary acts when using *ma* (x^2 = 5.53, *p* = 0.12>0.05).

Table 2
Illocutionary Acts of ma between Native Speakers and Learners

Item	Group	Illocutionary Acts								χ^2	p
		Repre-sentative		Directive		Com-missive		Expre-ssive			
Langu-age	Native	74	83.1%	9	10.1%	3	3.4%	3	3.4%	5.53	0.12
	Learner	227	91.5%	11	4.4%	4	1.6%	6	2.4%		

Note: Fisher's Exact Test

3.3.   Interaction between Illocutionary Act and Metadata

We further analysed illocutionary speech acts of *ma* in samples obtained from Chinese learners. We applied factor analysis to our data. Complete results were shown in Table 3 and showed that there was no significant difference among factors, including gender, region, first language, proficiency, and genre, when we compared illocutionary speech acts in samples (*p* > 0.05).

Table 3
Illocutionary Acts of ma(嘛) among Different Groups

Item	Group	Illocutionary Acts								x2	p
		Repre-sentative		Directive		Com-missive		Expre-ssive			
		num	%	num	%	num	%	num	%		
Gender	Male	83	90.2	5	5.4	1	1.1	3	3.3	1.24	0.76
	Female	144	92.3	6	3.8	3	1.9	5	1.9		
Region	African	46	93.9	1	2.0	0	0.0	2	4.1		
	American	1	100	0	0.0	0	0.0	0	0.0	11.65	0.42
	Asian	172	91.0	10	5.3	4	2.1	3	1.6		
	European	8	88.9	0	0.0	0	0.0	1	11.1		
1st Lang	analytic	98	94.2	4	3.8	1	1.0	1	1.0		
	agglutina-ting	48	92.3	1	1.9	0	0.0	3	5.8	7.60	0.66
	fusional	78	87.6	6	6.7	3	3.4	2	2.2		
	bilingual	3	100	0	0.0	0	0.0	0	0.0		
Profi-ciency	beginner	10	76.9	2	15.4	0	0.0	1	7.7		
	inter-mediate	17	94.4	0	0.0	0	0.0	1	3.4	9.86	0.09
	advanced	200	93.0	9	4.2	4	1.9	2	0.9		

Genre	interview	168	91.3	9	4.9	4	2.2	3	1.6		
	instruction	27	93.1	1	3.4	0	0.0	1	3.4	3.35	0.68
	exam	32	91.4	1	2.9	0	0.0	2	5.7		

4. Discussions

4.1. *Ma: An Informal SP Learned in Intermediate but Mastered in Advanced with Gendered Usage*

As is suggested in the work of Liu (2014), the most frequently occurring modal particles are ne1, "ma1, ba, a1, de, and le3. Comparing to them, ma is used less frequently and more casually. As language learning always starts from easier and formal usage, this could explain why results showed that among language learners the frequency of using ma increased with Chinese proficiency and decreased with the formality of genre. Level of proficiency is a noteworthy variable in this study ($0.05<p<0.1$), though the correlation between proficiency level and its pragmatic usage is not so significant as is suggested in studies of other SPs (e.g. le3 in (Xu et al. 2019)). It is worth attention that compares to intermediate learners (Medium: 1.7%, IQR (interquartile range): 48.9%) and advanced learners (Medium: 3.05%, IQR: 47.2%), beginners data has a higher medium (11.55%) but moderately lower IQR (42.3%). It indicates that intermediate learners begin to use ma with diversified acts and that advanced learners could use it not only in the various acts but also more balanced like what native speakers did. Besides, gender differences also deserve our attention. Though male and female behavior similarly in terms of illocutionary acts, the frequency of using ma is much higher for female participants (1.7 times the males). As is suggested by ZHAO & YANG (2016), ma is related to four types of emotions, dissatisfaction, impatience, disobedience (but with fondness, something like being cute), and carelessness, but expressed with a softer tone to make it less emotionally strong. We think it is related to the stereotypes and shackles of females to be weak, flattering, and reserved in (east) Asian cultures. Males are required to be strong, leader-acting, and straight-forward. Such stereotypes are mapped in language usage. If this argument is accepted, then it would not be surprising to find that such gender differences have also been detected in other SPs and languages, like je and jek in Cantonese (Chan 1999), ma1 and *ba* in Mandarin (Wamsley 2019), and *masu* in Japanese (SturtzSreetharan 2006). In other words, these SPs have functioned like gender identity markers in these languages.

4.2. *Representative act: a general and basic usage of ma*

One of the key findings in our study is that the speech act of *ma* as an SP does not significantly differ neither between natives and learners nor among the different groups of learners. Regarding language proficiency, the representative act has appeared and got frequently used since the beginner state, but other acts, especially commissive acts, only took place (or frequently) among advanced learners and native speakers. It indicates that representative act is not only a general pragmatic behavior of *ma* but also the most basic one. When people chose *ma* as the SP, normally they are stating or explaining something they believe is true, and this remains the major act during the ascending of language proficiency. The main pragmatic function of

ma is to state or explain what the speaker believes to be true. Though the point of representative acts is to commit the truth of the uttered proposition, this pragmatic function may also interact with the emotive function of *ma.* This interaction depends on the common ground shared by the speaker and hearer(s). Take the sentence "在下雨嘛 (with the transcript as below)" as an example. Imagined that this is a response from a lazy boy who did not want to throw the rubbish as his mother told him and tried to find excuses. As he and his mother may share a common ground that raining makes it difficult to go out, *ma* in this utterance could also convey the emotion of disobedience and the next sentence is likely to be "I don't want to go out now".

在   下雨   嘛
zài   xiàyǔ   ma
(It) is   raining   PART .

Directive act was used relatively frequently regardless of representative act in most groups, and nearly all of the examples of directive acts are imperative sentences, indicating that using *ma* may potentially convey an implication of motivating the hearer to do something. This tendency is also detected when comparing with a similar SP *bei* (Guo 2012). As is suggested by Guo, *ma* is more likely to be used with simple and plainest facts with an emphatic tone implying the wishes that the hearer could understand (or could have understood) these facts and then perform certain behaviors. Another thing to note is that declaration act is not detected in these data. We think it could be explained by its pragmatic function of expressing politeness (Lepadat 2017) and softening tones (Yang 2017). When adding *ma* at the final of the sentence, the speaker would like to make the imperative sound more like a request allowing discussion rather than a determination or command, to make it more likely to be accepted and performed. This is not what declarations usually do. The typical performance verbs in declaration acts are "nominate", "announce", "declare", etc. They are more likely to be used in a more formal situation to declare determined decisions from those in a higher position (e.g. the leader, the manager, the authority). *Ma* is not very likely to be used in such situations.

4.3.   Implications for Teaching Chinese to Foreign Learners
In teaching Chinese sentence-final particles to foreign learners, it's more common to start in the intermediate stage. It would be useful to specify that ma is an informal oral SP and mostly used to state what you believe is true. As its pragmatic functions are related to politeness and tone-softening, it would help to let students know that they could use *ma* when they want to sound politer and more acceptable. As the function of the gender identity marker of ma derives from gender stereotypes, it is also important for teachers to avoid showing such a tendency in the teaching process. We need to erase these gender stereotypes and inequality in language use and bring it to a neutral one. Another thing to note is that the homophone might be a major challenge in the acquisition of *ma* and the main reason for misusages. There is another SP in Chinese also pronounced *ma* but in the first tone (*ma1*, 吗) while the *ma*(嘛) we discussed in this study is pronounced in the second tone (*ma2*) or neutral tone (*ma*). *Ma1* is a question marker used more with

commissive or expressive acts, while *ma* is used to state what is believed true and more with representative acts. For learners from non-tone language backgrounds, it is challenged to distinguish two characters with only tonal differences in speech. Mixed usage could lead to pragmatic anomie in speech and cause misunderstandings. Read the examples (1) and (2) as below. Using *ma1* makes the sentence a question waiting for confirmation and using *ma* makes it a promise. Teachers need to help learners distinguish these two SPs.

(1) I read it?   (2) I'll read it.
   我   读   吗   我   读   嘛
   Wǒ   dú   mā   Wǒ   dú   ma
   I   read it PART ?   I   read it PART .

5. Conclusions

In this study, we discussed the usage and illocutionary acts of *ma* in Chinese learners based on data from Guangwai-Lancaster Chinese Learner Corpus. We think *ma* is an informal sentence-final particle that learners learn in the intermediate stage and become more familiar with and capable of at the advanced level. The frequency of using *ma* differs between females and males, indicating that it has been used like gender identity markers due to traditional stereotypes. Based on these, we presented some suggestions for teaching, with emphasis on the pragmatic functions of *ma* and distinguishing it from its homophones. As our study is based on a single learner corpus and human labeling, the size of samples is not that large, some metadata is missing, and human efforts are needed. Future studies could focus on building a larger learner corpus (or use multiple corpora) with complete metadata and the automatic recognition of illocutionary acts in sentences.

References

Ahn, M. (2016). Surprise in Discourse: The mirative meanings of ta(ha)-derived sentence final particles in Korean. 언어와 언어학 71:95–113.

Austin, J. L. (1962). How to Do Things With Words. Oxford university press.

Badan, L., & Chiara, R. (2020). The Acquisition of Mandarin Sentence Final Particles by Italian Learners. *International Review of Applied Linguistics in Language Teaching*, 58(4), 475–94. doi: 10.1515/iral-2017-0090.

Chan, M. K. M. (1999). Sentence-Final Particles in Cantonese: A Gender-Linked Survey and Study. Corpus ID: 199424230

Dong, F., Xiangyue, C., & Rui, H. (2018). The Pragmatic Functions and Prosodic Features of the Sentence-Final Particle Ba in the Imperatives. *Communication and Linguistics Studies*, 4(3), 80–87. doi: 10.11648/j.cls.20180403.13.

Guo, H. (2012). Convinced Particles 'Ma' and 'Bei'. *Journal of Capital Normal University (Social Sciences Edition)* (5), 79–86.

Kakegawa, T. (2009). Development of the Use of Japanese Sentence-Final Particles through Email Correspondence. In Naoko Taguchi, *Pragmatic competence*. New York: De Gruyter Mouton, Pp. 301–304.

Kilgarriff, A., Vít, B., Jan, B., Miloš, J., Vojtěch, K., Jan, M., Pavel, R., & Vít, S. (2014). The Sketch Engine: Ten Years On. *Lexicography* 1(1), 7–36. doi: 10.1007/s40607-014-0009-9.

Lepadat, C. (2017). The Modal Particle Ma 嘛: Theoretical Frames, Analysis and Interpretive Perspectives. *Quaderni Di Linguistica e Studi Orientali,* 3, 243–70. doi: 10.13128/QULSO-2421-7220-21347.

Liu, M. (2014). The Functional Analysis of Chinese Modal 'Ne' 'Ma' 'Ba' 'a' and in the Study of Teaching Chinese as a Foreign Language. Unpublished Thesis, Jilin University.

Mac, D-K., Thi-Lan, N., Alexis, M., & Do-Dat, T. (2015). Influences of Speaker Attitudes on Glottalized Tones: A Study of Two Vietnamese Sentence-Final Particles. Paper presented at *ICPhS Conference,* Scotland, United Kingdom.

McEnery, T., Vaclav, B., Dana, G., & Jayanti, B. (2019). Corpus Linguistics, Learner Corpora, and SLA: Employing Technology to Analyze Language Use. *Annual Review of Applied Linguistics* 39, 74–92.
doi: 10.1017/S0267190519000096.

Searle, J. R. (1969). *Expression and Meaning: Studies in the Theory of Speech Acts.* Reprinted. Cambridge: Cambridge Univ. Press.

Searle, J. R. (1975). A Taxonomy of Illocutionary Acts. *Language, mind, and knowledge.* 7, 344-369.

Searle, J R. 1976. A Classification of Illocutionary Acts. *Language in Society.* 5(1), 1-23.

SturtzSreetharan, C. L. (2006). Gentlemanly Gender? Japanese Men's Use of Clause-final Politeness in Casual Conversations. *Journal of Sociolinguistics* 10(1), 70–92. doi: 10.1111/j.1360-6441.2006.00318.x.

Sybesma, R., & Boya L. (2007). The Dissection and Structural Mapping of Cantonese Sentence Final Particles. *Lingua* 117(10), 1739–83.
doi: 10.1016/j.lingua.2006.10.003.

Wamsley, J. C. (2019). Gendered Usage of Sentence-Final Particles in Mandarin Chinese. *IULC Working Papers,* 19(1), 124-142.

Xu, H., Lu, X., & Brezina, V. (2019). Acquisition of the Chinese Particle le by L2 Learners: A Corpus-Based Approach. In *Chinese Language Learning Sciences* (pp. 197-216). doi:10.1007/978-981-13-3570-9_10.

Xu W., LI X., & Wu M. (2020). A Discourse Analysis of Xi Jinping's speeches from the Perspective of Speech Act Theory. *Journal of Hanjiang Normal University*, 40(6), 32–37.
doi: 10.19575/j.cnki.cn42-4892/g4.2020.06.007.

Yan, S. (2018). *Chinese Sentence-Final Particles and Their Behaviours in English Speakers' L2 Chinese.* Thesis, University of Cambridge.

Yang, Q. (2017). A Study of Pragmatic Functions of 'Ma.' *Modern Communication*, 19, 69–70.

Zhao, C., & Yang, C. (2016). On the Relevance Between Cognition and Emotion of the Sentence-final Particle 'MA.' *Journal of Foreign Languages* 39(5), 32–45.