



## Do accomplishments in Cantonese allow incomplete readings? A crowdsourced experimental study

Bowen Jiang<sup>1</sup>, Alan Yu<sup>2</sup>, Ming Xiang<sup>3</sup>, & Yenan Sun<sup>1</sup>

<sup>1</sup> The Chinese University of Hong Kong, <sup>2</sup> University of California, Berkeley, <sup>3</sup> University of Chicago

**Abstract.** Accomplishment predicates in Mandarin are known to permit incomplete readings under a perfective viewpoint, and previous studies have shown that the acceptability of such readings is affected by various predicate properties. This study examines whether accomplishments in Hong Kong Cantonese<sup>1</sup> show similar behavior and whether predicate type influences the acceptability of incomplete readings in Cantonese. As part of a larger crowdsourced project on language variation and change in Hong Kong, acceptability judgment data were collected from native Cantonese speakers with diverse sociolinguistic and demographic backgrounds. The results show a clear cross-linguistic difference between Cantonese and Mandarin: unlike Mandarin speakers, Cantonese speakers generally disprefer incomplete readings for perfective accomplishments. Moreover, predicate type, which significantly affects Mandarin judgments, does not significantly influence Cantonese judgments: the Cantonese pattern is largely uniform across all predicate types tested. At the same time, the acceptability of incomplete readings in Cantonese varies with speaker background. Cantonese speakers with greater Mandarin exposure, weaker local identity orientation, and less locally rooted community backgrounds are more likely to accept incomplete readings, showing more Mandarin-like judgment patterns, though not necessarily Mandarin-like sensitivity to predicate-type differences.

**Keywords:** accomplishment; incomplete readings; language contact and change; Hong Kong Cantonese

**1. Introduction.** A well-known property of Mandarin Chinese is that some accomplishment-type predicates permit incomplete readings even under the perfective aspect. For example, the sentence in (1) containing a perfective-marked accomplishment *du-le yi ben shu* ‘read a book’ does not force a culminating interpretation. This is demonstrated by its acceptability when followed by a clause explicitly stating that the book was not finished. The perfective predicate here is therefore compatible with an incomplete reading.

(1) Zhangsan du-le yi ben shu, keshi zhe ben shu ta meiyou du-wan  
Zhangsan read-PFV one CL book but this CL book he not read-finish  
‘Zhangsan read a book, but he didn’t finish reading this book.’

Previous studies have shown that the acceptability of incomplete readings is conditioned by a number of predicate properties, which can be summarized as the following three contrasts (Koenig & Chief, 2008; Soh & Kuo, 2005; Yang, 2011; Zhang, 2020). The first contrast concerns whether the object allows a partial reading. For instance, half a letter can still plausibly count as a letter, whereas half a circle normally cannot count as a circle. This distinction is described as “Allows Partial Objects” (APO) versus “No Partial Objects” (NPO).

---

<sup>1</sup> Hereafter abbreviated as Cantonese.

Soh & Kuo (2005) and Yang (2011) noted that predicates with APO objects are more compatible with incomplete readings than those with NPO objects, which accounts for the lower acceptability of (2b) compared to (2a).

- (2) a. Zhangsan xie-le yi feng xin, keshi zhe feng xin ta meiyou xie-wan.  
 Zhangsan write-PFV one CL letter but this CL letter he not write-finish  
 ‘Zhangsan wrote **a letter**<sub>APO</sub>, but he didn’t finish writing this letter.
- b. # Zhangsan hua-le yi ge quan, keshi zhe ge quan ta meiyou hua-wan.  
 Zhangsan draw-PFV one CL circle but this CL circle he not draw-finish  
 ‘Zhangsan draw **a circle**<sub>NPO</sub>, but he didn’t finish drawing this circle.

Second, event duration matters. Yang (2011) reports that predicates describing relatively longer events, such as writing a letter, are more compatible with incomplete readings than those describing shorter events, such as writing a character, as illustrated by the contrast between (3a) and (3b).

- (3) a. Zhangsan xie-le yi feng xin, keshi zhe feng xin ta meiyou xie-wan.  
 Zhangsan write-PFV one CL letter but this CL letter he not write-finish  
 ‘Zhangsan **wrote a letter**<sub>Long</sub>, but he didn’t finish writing this letter.
- b. # Zhangsan xie-le yi ge zi, keshi zhe ge zi ta meiyou xie-wan.  
 Zhangsan write-PFV one CL character but this CL character he not write-finish  
 ‘Zhangsan **wrote a character**<sub>Short</sub>, but he didn’t finish writing this character.

Third, object referentiality also makes a difference. Soh & Kuo (2005) and Zhang (2020) proposed that non-referential objects are least compatible with incomplete readings (4a), definite objects are most compatible (4c), and indefinite objects occupy an intermediate position (4b).

- (4) a. # Zhangsan he-le yi sheng shui, keshi zhe sheng shui ta meiyou he-wan.  
 Zhangsan write-PFV one LITER water but this LITER water he not drink-finish  
 ‘Zhangsan drank **a liter of water**<sub>Non-referential</sub>, but he didn’t finish drinking this liter.
- b. ? Zhangsan he-le yi ping shui, keshi zhe ping shui ta meiyou he-wan.  
 Zhangsan write-PFV one bottle water but this bottle water he not drink-finish  
 ‘Zhangsan drank **a bottle of water**<sub>Indefinite</sub>, but he didn’t finish drinking this bottle.
- c. Zhangsan he-le zhe ping shui, keshi zhe ping shui ta meiyou he-wan.  
 Zhangsan write-PFV this bottle water but this bottle water he not drink-finish  
 ‘Zhangsan drank **this bottle of water**<sub>Definite</sub>, but he didn’t finish drinking that bottle.

Although this line of work has developed a substantial picture for Mandarin, the corresponding pattern in other Chinese languages, for instance Cantonese, remains unclear. Conducting a parallel study in Cantonese is theoretically valuable, as linguistic variations in closely related languages can help more precisely characterizing the nature of incomplete readings. This study experimentally investigates incomplete readings in Cantonese via a crowdsourced approach. Our research questions are as follows: (i) Do Cantonese perfective accomplishments allow incomplete readings? In other words, is a sentence like (5) acceptable to Cantonese native speakers? (ii) Do the predicate properties that condition incomplete readings in Mandarin have the same effect in Cantonese?

- (5) Siunan se-zo fung seon, batgwo fung seon keoi zungmei se-jyun.  
 Siunan write-PFV CL letter but CL letter he not write-finish

‘Siuman wrote a letter, but he didn’t finish writing this letter.’

[Cantonese]

In addition to broadening the empirical picture for the incomplete readings of accomplishment predicates, the cross-linguistic comparison between Cantonese and Mandarin also sheds important light on how language contact might shape semantic change. Cantonese spoken in the greater China region has been under intense language contact with Mandarin. It has been widely argued that language contact has led to mutual influence between Mandarin and other Chinese languages spoken in the broader China region, thereby reducing the differences among them (Yue-Hashimoto 1997; Matthews 2010; Matthews & Yip 2011; Cao & Yu 2019). Li (2019) further notes that the outcome of language contact is shaped by speakers’ sociolinguistic backgrounds, including language attitudes and language exposure. This suggests that contact-induced influence may vary across speakers. Therefore, a comparison between Mandarin and Cantonese in the acceptability of incomplete readings, along with any potential variation among Cantonese native speakers, could provide evidence for an emerging contact-induced change. This leads to two additional research questions: (iii) If Cantonese and Mandarin differ in the acceptability of incomplete readings, does this pattern reflect contact-induced change from Mandarin? (iv) Is the acceptability of incomplete readings in Cantonese predictable from sociolinguistic variables?

**2. Methodology.** The study was conducted online and implemented on Gorilla, a web-based Experiment Builder. Participants completed the task on mobile devices and received HK\$100 gift cards as compensation. Each trial presented one stimulus sentence, and participants rated its acceptability by dragging a bar on a sliding scale ranging from “cannot accept” to “can accept,” as shown in Figure 1. The presentation order of the stimuli was randomized.

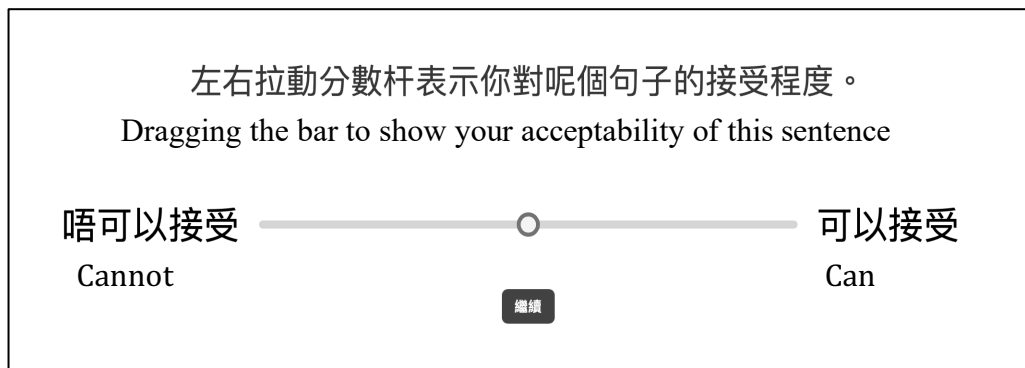


Figure 1. Experiment preview

The target stimuli were two-clause sentences in which the first clause contained an accomplishment predicate marked by the Cantonese perfective marker *zo*, and the second clause explicitly cancelled the completion of that same event, as in (6). The filler items also contained a perfective accomplishment in the first clause, but the second clause cancelled the completion of a different event rather than that of the event introduced in the first clause. For example, in (7), the cancellation targets a distinct event, “homework-writing,” rather than the first accomplishment, “letter-writing,” and therefore does not create an incomplete reading.

- (6) Siuman se-*zo*            fung   seon,   batgwo fung   seon   keoi            zungmei se-*jyun*.  
Siuman write-PFV   CL   letter   but   CL   letter   he            not   write-finish  
‘Siuman wrote a letter, but he didn’t finish writing this letter.’
- (7) Siuman se-*zo*            fung   seon,   batgwo go   fung   gungfo            keoi   zungmei se-*jyun*.

Siuman write-PFV CL letter but that CL. homework he not write-finish  
 ‘Siuman wrote a letter, but he didn’t finish writing that homework.’

Since previous studies have identified three predicate contrasts that affect the acceptability of incomplete readings in Mandarin, the present experiment included seven corresponding predicate categories. Specifically, these consisted of two categories for APO vs. NPO, two categories for long vs. short duration, and three categories for object referentiality: non-referential, indefinite, and definite. Each category contained four target stimuli and one filler, yielding a total of thirty-five stimuli.

In addition to the judgment task, participants completed a sociolinguistic questionnaire that collected their demographic information. The demographic variables included family Mandarin background, school Mandarin background, age, income, language attitude, identity, educational level, and region that the participant mainly live before age twelve<sup>2</sup>. The details of these variables are given in Table 1.

Demographic Variables	Values
Family Mandarin Background	1 ~ 4 (More and more influenced by Mandarin in family)
School Mandarin Background	1 ~ 3 (More and more influenced by Mandarin in school)
Age	< 18, 18 ~ 20, 21 ~ 30, 31 ~ 40, 41 ~ 50, 51 ~ 60, > 60
Income	≤ 30,000, 30,001 ~ 40,000, 40,001 ~ 50,000, 50,001 ~ 60,000, 60,001 ~ 70,000, 70,001 ~ 80,000, > 80,000
Language Attitude <sup>3</sup>	Cantonese Only vs. With Mandarin
Identity	Hongkonger vs. Others (Chinese, Chinese Hongkonger, English Hongkonger)
Educational Level	High school and below vs. College
Region B12 <sup>4</sup>	New Territories vs. Kowloon & Hong Kong Island

Table 1. Demographic factors

The Cantonese experiment involved 291 participants, all of whom self-identified as Hong Kong Cantonese native speakers and had lived in Hong Kong before the age of 12. The sample included 131 male participants and 160 female participants distributed across seven age groups (<18, 18-20, 21-30, 31-40, 41-50, 51-60, and >60), with the 18-20 and 21-30 groups being the largest. Participants were drawn from Hong Kong’s 18 districts (their distribution is shown in Table 2), allowing the study to investigate the potential regional variations based on a geographically representative sample.

Region B12	N	Region B12	N
Islands	5	Kowloon City	25
Kwai Tsing	12	Kwun Tong	25

<sup>2</sup> Hereafter abbreviated as Region B12.

<sup>3</sup> The question asked is “In which language should Chinese classes be taught in Hong Kong?”

<sup>4</sup> We distinguish the New Territories from Hong Kong Island/Kowloon because there is a broader sociolinguistic contrast between these areas, with the former appearing more locally oriented and the latter relatively more exposed to Mandarin- or English-medium interaction.

North	21	Sham Shui Po	19
Sai Kung	11	Wong Tai Sin	12
Sha Tin	17	Yau Tsim Mong	3
Tai Po	21	Central & Western	4
Tsuen Wan	25	Eastern	8
Tuen Mun	32	Southern	5
Yuen Long	41	Wan Chai	4

Table 2. Geographic distribution of the participants

A Mandarin baseline experiment was also conducted via *Qualtrics*. It used the same experimental design as the Cantonese study, except that the materials were Mandarin translations of the Cantonese test items. The sample included 197 participants, including 137 females and 60 males. Most participants were between 18 and 30 years old. Specifically, one participant was aged 17 or below, 30 were aged 18–20, 91 were aged 21–30, 54 were aged 31–40, 14 were aged 41–50, and 7 were aged 51–60. The majority of the participants (84.3%) are from the Mandarin-speaking (*guanhua*) regions of China.<sup>5</sup>

Before turning to the results, it is useful to lay out the main predictions. First, if Cantonese generally disfavors incomplete readings, target stimuli should receive lower ratings than fillers. Second, if Cantonese patterns like Mandarin, predicate type should affect acceptability, with APO, long-duration, and referential predicates receiving higher ratings than NPO, short-duration, and non-referential predicates. If Cantonese differs from Mandarin, however, such predicate-type effects may be absent. Third, in the Mandarin baseline, because APO, long-duration, and referential predicates are more compatible with incomplete readings, their target stimuli should not differ significantly from the fillers, whereas target items with NPO, short-duration, and non-referential predicates should receive significantly lower ratings. Finally, given previous claims that language contact may reduce differences between Mandarin and other Chinese languages, and that such effects may vary with speakers' sociolinguistic backgrounds (Yue-Hashimoto 1997; Matthews 2010; Matthews & Yip 2011; Cao & Yu 2019; Li 2019), we expect the acceptability of incomplete readings in Cantonese to vary across speakers with different demographic and sociolinguistic profiles.

**3. Results and discussion.** The first major comparison concerns target stimuli versus fillers. In the Cantonese data, target stimuli received a mean acceptability of 43.74, whereas fillers received a mean of 70.24. A paired t-test showed that this difference was statistically significant ( $t = -11.264, p < .05$ ). A beta regression model yielded the same conclusion: stimulus type had a significant effect, and filler items were rated significantly higher than target items ( $\beta = 2.7064, SE = 0.2240, z = 12.08, p < .001$ ). The beta regression was fitted with the *glmmTMB* package in R using a beta family. The model included random intercepts for participants and stimulus sentences, but no random slopes. Fixed effects consisted of stimulus type, the demographic predictors, and their interactions with stimulus type. Even when each predicate type is considered separately, target stimuli still receive significantly different ratings from filler stimuli (Table 3). These results suggest that Cantonese speakers generally reject sentences in which a

<sup>5</sup> Our results showed that, among the Mandarin native speakers, dialectal background did not have a significant effect on their responses, despite the fact that the participants came from different dialect regions. We therefore do not examine potential dialectal variations in the following analysis of the Mandarin baseline experiment.

perfective accomplishment is followed by cancellation of the completion of the same event. In other words, perfective accomplishments in Cantonese generally disallow incomplete readings.

Predicate category	Filler Intercept	Effect of Target	p-value
APO	68.698	-24.199	< .001 ***
NPO	70.701	-27.027	0.001 **
Long	70.790	-28.014	< .001 ***
Short	71.381	-28.401	< .001 ***
Non-referential	70.347	-26.755	< .001 ***
Indefinite	70.251	-26.168	< .001 ***
Definite	69.526	-24.971	< .001 ***

Table 3. Target Stimuli vs Corresponding Fillers in Cantonese experiment

This stands in clear contrast to the results of the Mandarin baseline experiment. In line with predictions from the existing literature, predicates with APO objects, long duration, and referential objects (indefinite or definite) were not rated significantly lower than fillers, while NPO predicates, short predicates, and non-referential predicates were rated significantly lower (Table 4). These findings replicate earlier studies (e.g., Liu, 2018) showing that at least certain Mandarin accomplishments allow incomplete readings under a perfective viewpoint, thereby confirming a difference between Mandarin and Cantonese in the acceptability of incomplete readings for perfective accomplishments.

Predicate category	Filler Intercept	Effect of Target	p-value
APO	67.019	-0.752	0.692
NPO	67.236	-4.064	0.048 *
Long	60.503	-0.091	0.965
Short	66.083	-17.196	< .001 ***
Non-referential	59.567	-5.417	0.008 **
Indefinite	64.987	-2.384	0.246
Definite	65.089	-1.833	0.361

Table 4. Target Stimuli vs Corresponding Fillers in Mandarin baseline experiment

To directly compare the Cantonese and Mandarin results, we fitted a single beta regression model to the combined target stimuli data from both languages, with Language and Predicate Category as fixed effects, together with random intercepts for participants and items. The model revealed a significant main effect of Language: Mandarin target items were rated significantly higher than Cantonese target items ( $\beta = 1.27058$ ,  $SE = 0.15722$ ,  $z = 8.081$ ,  $p < .001$ ). Predicate Category shows no significant effect. This result provides further evidence that Cantonese and Mandarin differ with respect to the licensing of incomplete readings.

Second, the results show that none of the predicate-type contrasts significantly affected the acceptability ratings for the Cantonese target stimuli (Table 5; all  $p$ -values  $> .1$ ). That is, the acceptability of incomplete readings for Cantonese perfective accomplishments does not vary

significantly across different predicate categories. In contrast, the Mandarin baseline experiment revealed significant effects for APO vs. NPO, short vs. long duration, and non-referential vs. both indefinite and definite objects, as shown in Table 6. These contrasts demonstrate that predicate type significantly affects the acceptability of incomplete readings for Mandarin perfective accomplishments, thereby revealing a variation between Chinese languages.

<b>Contrast</b>	<b>Baseline</b>	<b>Comparison</b>	<b>Estimate</b>	<b>p-value</b>
<b>APO vs NPO</b>	APO	NPO	-0.002	0.990
<b>Short vs Long</b>	Long	Short	-0.154	0.371
<b>Referentiality</b>	Non-referential	Indefinite	0.267	0.151
	Non-referential	Definite	0.123	0.489
	Indefinite	Definite	-0.134	0.422

Table 5. Main Effect of Predicate type in a Beta Regression Model

<b>Contrast</b>	<b>Baseline</b>	<b>Comparison</b>	<b>Estimate</b>	<b>p-value</b>
<b>APO vs NPO</b>	APO	NPO	-0.002	0.011 *
<b>Short vs Long</b>	Long	Short	-0.154	< .001 ***
<b>Referentiality</b>	Non-referential	Indefinite	0.267	< .001 ***
	Non-referential	Definite	0.123	< .001 ***
	Indefinite	Definite	-0.134	0.581

Table 6. Main Effect of Predicate type in Mandarin baseline experiment

Although we did not observe predicate type had an effect on Cantonese incomplete readings, the beta regression with certain sociolinguistic and demographic factors as fixed effect did reveal some language contact effects on the overall acceptability ratings of incomplete readings in Cantonese (Table 7). The overall acceptability ratings for the Cantonese target stimuli were significantly positively correlated with family Mandarin background and school Mandarin background, suggesting that greater exposure to Mandarin is associated with higher acceptability ratings. Moreover, participants who self-identified as “Hongkonger only” tended to give lower ratings, as did those residing in the New Territories. Higher income was also correlated with lower ratings. Language attitudes of the participants played a role as well. Participants who thought that Chinese classes in Hong Kong should be taught only in Cantonese gave significantly lower ratings than those who thought that Mandarin should also be included. Age and educational level were not significant predictors. To strengthen the interpretation of these findings, the study also examined the correlations between these sociolinguistic and demographic factors and the acceptability ratings for the filler stimuli. This was done to ensure that the observed effects were not due to a general tendency among certain sociolinguistic and demographic groups to rate all sentences more strictly or leniently. As shown in Table 8, the effects of these factors on the target stimuli were either absent or reversed for the filler stimuli. This strengthens the interpretation that language-contact oriented sociolinguistic variables specifically shape the acceptability of incomplete readings for Cantonese perfective accomplishments, rather than reflecting a broader response bias.

Variable	Coefficient	p-value
Family Mandarin Background	0.296	0.003 **
School Mandarin Background	0.431	0.003 **
Identity: Chinese/Chinese Hongkonger <sup>6</sup>	0.369	0.032 *
Identity: English Hongkonger	0.693	0.002 **
Region B12: New Territories	-0.354	0.011 *
Attitude: Mandarin	1.367	<0.001 ***
Income	-0.305	<0.001 ***
Age	0.098	0.073
Educational Level: High School and below	0.015	0.916

Table 7. Coefficients of sociolinguistic and demographic variables for the target stimuli

Variable	Coefficient	p-value
Family Mandarin Background	-0.199	0.02 *
School Mandarin Background	-0.016	0.892
Identity: Chinese/Chinese Hongkonger	-0.074	0.094
Identity: English Hongkonger	-0.398	0.004 **
Region B12: New Territories	-0.099	0.389
Attitude: Mandarin	-1.125	<0.001 ***
Income	0.193	<0.001 ***
Age	-0.199	0.02 *
Educational Level: High School and below	-0.016	0.892

Table 8. Coefficient of significant sociolinguistic and demographic variables for the filler stimuli

To investigate the relationships among the relevant sociolinguistic and demographic factors, the study further conducted a Factor Analysis of Mixed Data (FAMD; Pagès, 2004), implemented with the *FactoMineR* package in R (Lê et al., 2008). The results, shown in Table 9, revealed two principal components (PCs). PC1 was interpreted as a combination of higher income and stronger local identity orientation, characterized by self-identifying as Hongkonger only, rather than Chinese, Chinese Hongkonger, and English Hongkonger, and supporting that Chinese classes in Hong Kong should be taught only in Cantonese, rather than that Mandarin should also be involved. PC2 was interpreted as Mandarin exposure, with high positive loadings from both family and school Mandarin background. These two components had opposite effects on acceptability ratings. PC1 was negatively correlated with acceptability (coefficient = -15.1448,  $p < .001$ ). Given that Mandarin native speakers tended to assign higher ratings to the corresponding stimuli, this pattern suggests that a stronger local Cantonese identity orientation is

<sup>6</sup> The original identity variable had four categories: *Chinese*, *Chinese Hongkonger*, *English Hongkonger*, and *Hongkonger*. In the analysis, *Hongkonger* served as the baseline. *Chinese* and *Chinese Hongkonger* were combined because both indicate some degree of Chinese identification and did not differ significantly in the results.

associated with less Mandarin-like judgments on the target stimuli. In contrast, PC2 was positively correlated with acceptability (coefficient = 4.2055,  $p < .001$ ), indicating that greater Mandarin exposure is linked to more Mandarin-like judgments. The regional factor did not load strongly on either principal component, suggesting that its influence should be examined independently.

<b>Variable</b>	<b>PC1</b>	<b>PC2</b>
<b>Family Mandarin Background</b>	-0.071	<b>0.738</b>
<b>School Mandarin Background</b>	0.015	<b>0.774</b>
<b>Income</b>	<b>0.774</b>	0.149
<b>Identity: Hongkonger</b>	<b>1.157</b>	0.039
<b>Identity: Chinese or Chinese Hongkonger</b>	<b>-1.253</b>	0.031
<b>Identity: English Hongkonger</b>	<b>-1.122</b>	-0.251
<b>Attitude: Cantonese</b>	<b>1.535</b>	0.110
<b>Attitude: Mandarin</b>	<b>-0.995</b>	-0.071
Region B12: NT	0.175	-0.358
Region B12: HL & KL	-0.323	0.710

Table 9. FAMD results for sociolinguistic and demographic variables

We conducted a more detailed analysis of the region where participants lived before the age of twelve. Hong Kong is administratively divided into 18 districts, which are grouped into three broader areas: the New Territories, Hong Kong Island, and Kowloon. The New Territories comprise North, Tai Po, Sha Tin, Tsuen Wan, Kwai Tsing, Yuen Long, Tuen Mun, Islands, and Sai Kung. Hong Kong Island includes Central and Western, Wan Chai, Eastern, and Southern. Kowloon consists of Kowloon City, Kwun Tong, Sham Shui Po, Wong Tai Sin, and Yau Tsim Mong. Our sample includes data from 18 districts and the results indicate that participants from

the New Territories gave lower acceptability ratings, while those from Kowloon and Hong Kong Island gave higher ratings. This contrast is illustrated in Figure 2.<sup>7</sup>

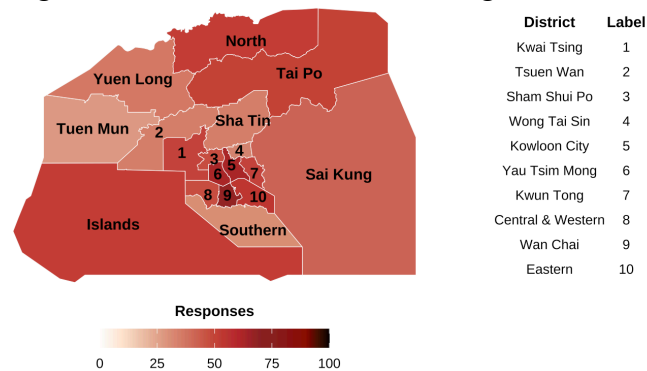


Figure 2. Rating means in 18 Districts of Hong Kong

To further examine the source of this regional contrast, we conducted a path analysis (Wright, 1921, 1934), implemented in R using the *lavaan* package (Rosseel, 2012), which showed that Region B12 affected acceptability both directly and indirectly, as shown in Table 10. The total indirect effect of Region B12 was significant (coefficient = 5.259,  $p < .001$ ), as was the total direct effect (coefficient = 6.406,  $p < .001$ ). Major mediators included identity, family Mandarin background, school Mandarin background, and income. These findings can be interpreted in sociolinguistic terms: the New Territories can be characterized as more locally oriented, with more locally born residents and more Cantonese-medium interaction, whereas Hong Kong Island and Kowloon can be characterized as less locally oriented and as having relatively more Mandarin- or English-medium interaction (Bacon-Shone et al., 2015; Hong Kong Home Affairs Department, 2025). From this perspective, the regional contrast appears to arise from broader sociolinguistic ecologies, in which more locally oriented communities are associated with less Mandarin exposure and stronger local identification, which in turn are associated with lower acceptability of incomplete readings.

Path	Coefficient	p-value
Region B12 → Attitude → Acceptability	0.043	0.841
Region B12 → Identity → Acceptability	2.092	<0.001 ***
Region B12 → Income → Acceptability	1.142	<0.001 ***
Region B12 → Family Mandarin Background → Acceptability	1.397	<0.001 ***
Region B12 → School Mandarin Background → Acceptability	0.585	<0.001 ***
Region B12 (Total indirect)	5.259	<0.001 ***
Region B12 (Total direct)	6.406	<0.001 ***

<sup>7</sup> Participants from the North and Tai Po districts assigned relatively high ratings to incomplete readings, deviating from the general tendency in the wider New Territories region. This finding is not entirely unexpected, given that the New Territories is a large and potentially heterogeneous area. This point will be revisited at the end of this section.

Table 10. Path analysis for sociolinguistic and demographic variables<sup>8</sup>

To make the regional pattern more explicit, we plotted the geographic distributions of four sociolinguistic or demographic variables across Hong Kong’s 18 districts (Figure 3): (a) shows family Mandarin background, (b) school Mandarin background, (c) lower-income proportion, and (d) the proportion of participants not identifying as “Hongkonger only.” These four variables were selected because they were all positively associated with acceptability, though in different ways: family Mandarin background and school Mandarin background were positively correlated with ratings; lower-income proportion was included because higher income was associated with lower ratings in the earlier analysis; and the proportion of participants not self-identifying as “Hongkonger only” was included because self-identifying as “Hongkonger only” was also associated with lower ratings. In each map, darker shading indicates a higher value of the corresponding variable in that district. Since these variables are positively associated with acceptability, darker areas are expected to correspond to districts where participants are more likely to give higher ratings. The geographic distributions of these four variables all align with the contrast between the New Territories and Kowloon/Hong Kong Island. This consistent pattern further supports the view that the region where participants lived before age twelve is closely associated with other sociolinguistic or demographic factors, and that its effect on native speakers’ judgments is partly mediated through these correlated sociolinguistic and demographic variables.

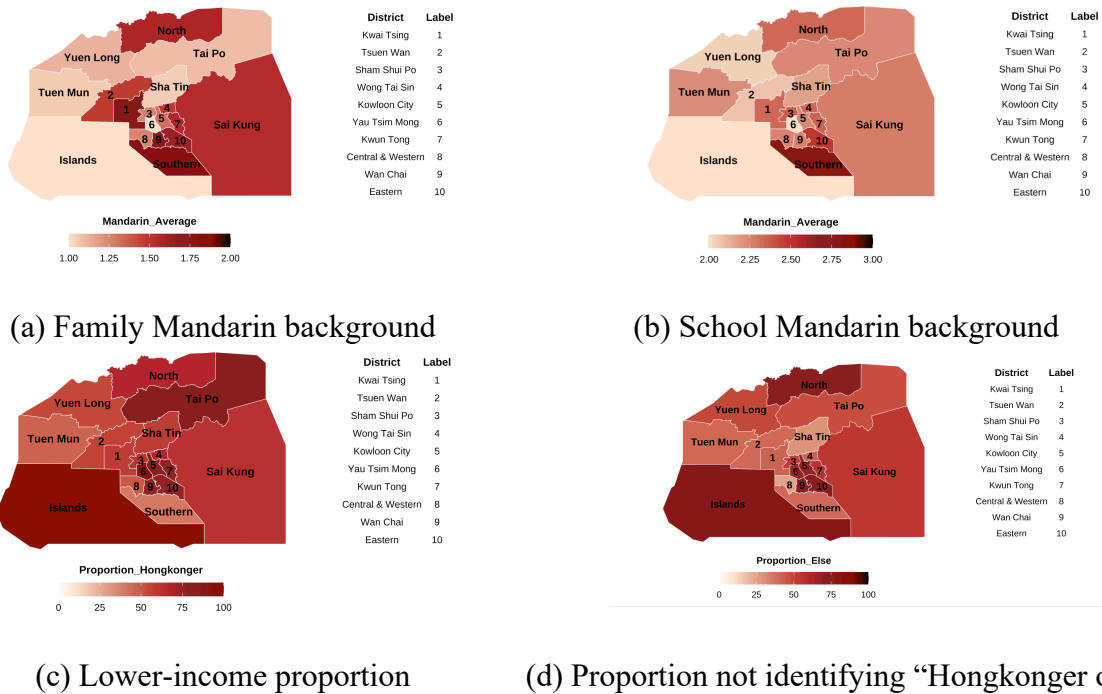


Figure 3. Geographic distributions of four sociolinguistic/demographic variables across Hong Kong’s 18 districts

A closer look at the pattern within the New Territories further shows that participants from North District and Tai Po gave significantly higher ratings than those from the other New

<sup>8</sup> The rows labeled with arrows represent indirect paths from Region to acceptability through individual sociolinguistic and demographic mediators. “Total indirect” is the sum of all mediated effects, and “Total direct” is the direct effect of Region after controlling for those mediators.

Territories districts (t-test,  $t = -3.0702$ ,  $p < 0.005$ ). According to Figure 3, this finer-grained pattern is also consistent with the geographic distribution of the relevant sociolinguistic and demographic factors. Within the New Territories, North District and Tai Po also show relatively higher values on these four variables. Since these variables are positively associated with acceptability, this pattern suggests that participants from these two districts are likewise more likely to give higher ratings. This provides further support for the view that the regional difference is not purely geographic, but arises from the sociolinguistic factors unevenly distributed across regions.

We interpret the correlation between acceptability ratings and sociolinguistic factors as evidence of incipient contact-induced influence from Mandarin on Cantonese. In short, Mandarin native speakers are more likely to accept incompletive readings, and Cantonese speakers with greater Mandarin exposure, weaker local identity orientation, and less locally rooted community backgrounds likewise tend to show more Mandarin-like judgment patterns. These parallel patterns suggest that although Cantonese and Mandarin still differ in the acceptability of incompletive readings for perfective accomplishments, contact with Mandarin appears to be reshaping the judgments of some Cantonese speakers.

However, this increased acceptability does not interact with predicate type, as no significant interaction was found between sociolinguistic and demographic factors and predicate type. In other words, although Cantonese speakers with certain sociolinguistic traits are more likely to accept incompletive readings overall, they may still be less sensitive than Mandarin native speakers to predicate-type differences. This finding is consistent with Silva-Corvalán's (1994) view of contact-induced change as a process that often involves simplification and reduction of contrasts, rather than the direct transfer of a fully preserved grammatical system. A closer parallel can be drawn with her discussion of the weakening of the Spanish preterite-imperfect opposition under language contact. In her account, one form may be extended beyond the contexts in which it is originally licensed, so that preterite forms come to be used in environments previously associated with the imperfect, thereby weakening the original contrast. The pattern hypothesized here is similar: once the incompletive reading is acquired, speakers may extend it to predicate types that do not originally favor it, resulting in a weakening or failure to maintain the original predicate contrast.

**4. Conclusion.** Our findings show that Cantonese differs significantly from Mandarin in the acceptability of incompletive readings for perfective accomplishments. While Cantonese speakers generally disfavor such readings, Mandarin speakers accept them for certain predicate types. A key finding is that predicate type, a significant factor in Mandarin, does not influence Cantonese speakers' judgments. Instead, the Cantonese pattern is uniform across all predicate types tested.

Additionally, our study reveals that the acceptability of incompletive readings in Cantonese is modulated by individual speaker background. Cantonese speakers with greater Mandarin exposure, weaker local identity orientation, and less locally rooted community backgrounds are more likely to accept incompletive readings of accomplishments, thus patterning more closely with Mandarin speakers. However, they may still be less sensitive than Mandarin native speakers to predicate-type differences. We tentatively hypothesize a process of simplification and overgeneralization: once the incompletive reading is acquired, it is extended across predicate types, resulting in a weakening or failure to maintain the original predicate contrast. Future research, with more fine-grained experimental designs, is needed to examine this

tentative hypothesis more closely, especially the mechanisms by which simplification takes place and how the incomplete reading extends to predicate types that do not originally favor it.

**Acknowledgement.** This project is funded by the Strategic Partnership Award for Research Collaboration 2024 (CUHK) and the Provost's Global Faculty Awards for Hong Kong and Southeast Asia (UChicago). We also thank the feedback from anonymous reviewers of the LSA Annual Meeting and all the comments during the conference presentation.

## References

- Bacon-Shone, John, Kingsley Bolton & Kang Kwong Luke. 2015. *Language use, proficiency and attitudes in Hong Kong*. Hong Kong: Social Sciences Research Centre, The University of Hong Kong.
- Cao, Guangshun & Hsiao-jung Yu (eds.). 2019. *Language contact and change in Chinese* (Vol. 1). Berlin: Walter de Gruyter.
- Hong Kong Home Affairs Department. 2025. *Surveys on new arrivals from the Mainland*. Hong Kong: Hong Kong SAR Government.
- Koenig, Jean-Pierre & Liangcheng Chief. 2008. Scalarity and state changes in Mandarin (and other languages). In CSSP.
- Lê, Sébastien, Julie Josse & François Husson. 2008. FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software* 25(1). 1–18.
- Liu, Jinhong. 2018. *Non-culminating causative accomplishments in Mandarin: Experimental evidence*. Nantes: Laboratoire de Linguistique de Nantes dissertation.
- Matthews, Stephen. 2010. Language contact and Chinese. In Raymond Hickey (ed.), *The handbook of language contact*, 757–769. Oxford: Wiley-Blackwell.
- Matthews, Stephen & Virginia Yip. 2011. *Cantonese: A comprehensive grammar*. London: Routledge.
- Pagès, Jérôme. 2004. Analyse factorielle de données mixtes. *Revue de Statistique Appliquée* 52(4). 93–111.
- Rosseel, Yves. 2012. lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48(2). 1–36.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Soh, Hooi Ling & Jenny Yi-Chun Kuo. 2005. Perfective aspect and accomplishment situations in Mandarin Chinese. In Paula Kempchinsky & Roumyana Slabakova (eds.), *Perspectives on aspect*, 199–216. Dordrecht: Springer.
- Wright, Sewall. 1921. Correlation and causation. *Journal of Agricultural Research* 20. 557–585.
- Wright, Sewall. 1934. The method of path coefficients. *The Annals of Mathematical Statistics* 5(3). 161–215.
- Yang, Suying. 2011. The parameter of temporal endpoint and the basic function of -le. *Journal of East Asian Linguistics* 20. 383–415.
- Yue-Hashimoto, Anne O. 1997. Syntactic change in progress—Part I: The comparative construction in Hong Kong Cantonese. In Anne O. Yue & Mantaro J. Hashimoto (eds.), *In memory of Mantaro J. Hashimoto*, 329–375.
- Zhang, Anqi. 2020. Referentiality, individuation and incomplete readings. *Journal of East Asian Linguistics* 29(4). 435–468.