

Chapter 5

Results



Abstract This chapter presents the major study's quantitative and qualitative findings. It starts by displaying descriptive statistics for each task in order to illustrate the overall data distributions. The chapter then presents the research findings, statistical analyses, and research questions addressed in the five subsections that follow.

Keywords Quantitative findings · Qualitative findings · Research questions

5.1 Results for Routine Production

Table 5.1 provides the descriptive statistics for the different levels of productive pragmatic competence of routines scores for the three participant groups' oral responses. According to the results, overall productive pragmatic competence should be judged satisfactory, since all score rates were significantly greater than 70.00%. Furthermore, there was a consistent trend across three levels among three groups: G3 students scored significantly higher on all three sections than G2 students, and both G3 and G2 significantly outperformed G1 ($G3 > G2 > G1$), indicating that both levels of context knowledge, including the holistic productive pragmatic competence of routines, developed hand in hand with proficiency and study-abroad experience to a large extent.

To be specific, the frequency of the ASC was also more significant than that in the PC part: $\%_{ASC} (M_{ASC} = 2.61/3) = 87.00 > \%_{PC} (M_{TPC} = 2.20/3) = 73.33$. Paired-samples t-test then reveals there indeed existed a significant difference between the two levels of context knowledge manipulated by all learners with an enormous effect size: $t(142) = 40.46, p = 0.00 < 0.01$, Cohen's $d = 1.78 > 0.8$. It is apparent that in all cases, when facing specific actual situational context, participants across the three groups could accurately perform the role of the respondent in time and identify the target speech acts with pragmatic strategies based on the contextual information.

To be more explicit, the frequency of the ASC was higher than that of the PC part: $87.00\% > 73.33\%$. The paired-samples t-test then demonstrates that there was a significant difference between the two levels of context knowledge controlled by all learners, with a massive effect size: $t(142) = 40.46, p = 0.00 < 0.01$, Cohen's $d = 1.78 > 0.8$. In all situations, when confronted with a specific actual situational

Table 5.1 Descriptive statistics for productive pragmatic competence of routines

Level	Group							
	G1 (n = 51)		G2 (n = 59)		G3 (n = 33)		Overall (n = 143)	
	M	%	M	%	M	%	M	%
ASC	2.54	84.67	2.55	85.00	2.81	93.67	2.61	87.00
PC	2.06	68.67	2.14	71.33	2.52	84.00	2.20	73.33
ProPCR	4.60	76.67	4.69	78.17	5.33	88.83	4.81	80.17

Note ASC actual situational context; PC prior context; ProPCR productive pragmatic competence of routines

context, participants from all three groups were able to properly execute the role of the respondent in time and identify the target speech actions using pragmatic strategies based on contextual knowledge. In comparison, the native-like selection of target routines has several limitations. Participants lacked or were unable to properly use sufficient PC knowledge to produce the target expression to meet the demands of the current situation. It is possible to conclude that the acquisition of ASC information influenced the retrieval of PC knowledge and the accurate production of routine expressions to some extent.

When the initiating and responding modalities were compared, significant differences were found among three paired levels with small and medium effect sizes: all $p = 0.00 < 0.05$, $0.4 < \text{all Cohen's } d < 0.6$, indicating that all respondents outperformed in responding utterances rather than initiating a conversation through routines on both sides of productive pragmatic competence, which was also included. In other words, the participants had no idea what appropriate routine expressions to use to start a conversation, yet as listeners, they were reasonably adept at responding to the speaker's conversation. In the present study, initiating a discussion was shown to be substantially more challenging than replying to it utilizing routines in terms of task difficulty. The statistical findings for the initiating and responding tasks are shown in the next two sub-sections.

5.1.1 Results for Initiating Utterances

As shown in Table 5.2, the competence of initiating utterances was deemed rather good, with an average score of 4.64 and a scoring rate of 77.33%, which was significantly higher than 70.00% overall. Furthermore, the average mean score in the ASC section was 2.54 (means above an exceptional level of 80.00%), demonstrating that they were able to successfully capture contextual information. The total PC score, on the other hand, was somewhat higher than the threshold regarding "acceptable routines without mentioning the lexical core" ($M = 2.09 > 2.00$), indicating that learners' retrieval of PC knowledge was not substantially impeded. As a result, they were unable to start a conversation in the target language appropriately. The

Table 5.2 Descriptive statistics for competence of initiating utterances

Level	Group							
	G1 (n = 51)		G2 (n = 59)		G3 (n = 33)		Overall	
	Mean	%	Mean	%	Mean	%	Mean	%
ASC	2.48	82.67	2.44	81.33	2.84	94.67	2.54	84.67
PC	1.92	64.00	2.00	66.67	2.52	84.00	2.09	69.67
CIU	4.40	73.33	4.43	73.83	5.36	89.33	4.64	77.33

Note ASC actual situational context; PC prior context; CIU competence of initiating utterances

group scores demonstrate that learners' reported initiating competence of routines increased with proficiency levels and residence duration. The sole exception to this variance trend was discovered among non-residence learners, with lower-level counterparts reporting ACS section scores that outperformed those of higher-level peers, demonstrating that the mastery degree of ASC information does not always rise with proficiency level.

Based on cross-sectional performance across three groups, the reported initiating ASC scores for the routine formulae range from 2.44 (G2, indicating that each respondent in a group achieved at least two levels of contextual information provided in the ASC section) to 2.84 (G3, indicating that not all high-level, abroad-residence learners can fully comprehend the requirements of ASC). In comparison, the reported initiating PC scores of two non-residence groups were either "without lexical core level" or fell short of this level on average. However, respondents with study-abroad experience might reach the highest "alternative wording sequence" level, emphasizing the contributive role of study-abroad experience in initiating utterances.

According to the independent sample t-test results, proficiency level had no effect on both levels of initiating competence, including itself, for all $p > 0.05$. On the contrary, study-abroad experience (together with both factors' interaction) had a significant impact on learners' initiating competence with very large effect sizes: all $p < 0.01$, all Cohen's $d > 0.8$.

When comparing ASC and PC in the CIU task, a significant difference was observed with a significant effect size: $p = 0.00 < 0.01$, Cohen's $d = 1.36 > 0.8$, indicating that learners were much more likely to comprehend the contextual reminders embedded in the ASC when compared to fulfilling the preferred way of routine selection by their PC knowledge.

5.1.2 Results for Responding to Utterances

As shown in Table 5.3, the competence of responding to utterances was rated substantially higher than scores acquired in the section, with the average score arriving at 4.92 and its scoring rate attaining 82.00%, which was higher than the 80.00% overall.

Table 5.3 Descriptive statistics for competence of responding to utterances

Level	Group							
	G1 (n = 51)		G2 (n = 59)		G3 (n = 33)		Overall (n = 143)	
	M	F	M	F	M	F	M	F
ASC	2.58	86.00	2.62	87.33	2.79	93.00	2.65	88.33
PC	2.16	72.00	2.23	74.33	2.52	84.00	2.27	75.67
CRU	4.74	79.00	4.85	80.83	5.32	88.67	4.92	82.00

Note ASC actual situational context; PC prior context; CRU competence of responding to utterances

Furthermore, the average mean score in the ASC part was 2.65, exceeding an exceptional level of 85.00%, indicating that they were able to capture more contextual information than they did in the initial task. In contrast, the total PC score was greater than Level 3, but fell short of the “alternative appropriate wording” level ($2.27 > 2.00$), as evidenced by learners’ retrieval of PC knowledge being less restricted in the CRU section than in the initiating section ($2.27 > 2.09$). As a result, respondents were far more likely to correctly respond to the speaker’s utterances using the target routines. With practically no exceptions, group scores demonstrate that learners’ reported responding to an utterance by the routine expressions improved with proficiency and study-abroad experience from the leftmost to the terminal column.

The reported responding ASC scores for the routine formulae vary from 2.58 (G1) to 2.79 (G3), indicating that each respondent in a group reached at least two levels of contextual knowledge supplied in the ASC section. In contrast, the reported responding PC scores of two non-residence groups totaled the level “without lexical core”. However, respondents with abroad experience might reach the top level of “alternative wording sequence”, emphasizing the significance of abroad residence in competence of responding to utterances.

According to the independent sample t-test results, proficiency level has no influence on both levels of competence in responding to utterances (both $p > 0.05$), with the exception of the PC section, which is significant at $p = 0.04 < 0.05$ level with a small effect size: Cohen’s $d = 0.4 < 0.5$. On the contrary, study-abroad experience (together with both factors’ combination) repeatedly had a significant impact on learners’ capabilities to respond to utterances, with large effect sizes: all $p < 0.01$, all Cohen’s $d > 0.8$.

When the disparity between ASC and PC in the responding modality was compared, a statistical significance was found with a significant effect size: $p = 0.00 < 0.01$, Cohen’s $d = 1.72 > 0.8$. As a result, in addition to fulfilling their preferred methods of routine selection through their PC knowledge, learners were also more likely to grasp the contextual reminders embedded in the individual ASC, which was comparable to the initiating competence.

5.2 Results for Routine Recognition

Routine recognition was primarily represented by participants' mastery of their prior context knowledge under computer-animated contextual reminders, that is, their ability to select the exclusive target expression from other semantically/syntactically-approximate, functionally-deviated options. The percentage of correct answers for each routine at each proficiency level with and without abroad experience is shown in Table 5.4. It appears that participants' recognition was nearly unacceptable, with a wide range of scores ranging from 26.57 to 98.60%.

For example, Item 7 (*Can I leave a message?*) was the most well-known to all groups (all score rates over 98.00%), including both groups with and without international experience. In contrast, Item 3 (*Can I get you anything else?*) was least recognized at a high level by the non-residence group (11.86%), and the same was true for proficient G3 with abroad experience: 45.45%. It implies that learners might use their PC knowledge to distinguish native-like routines from those with similar syntactic structures, semantic meanings, or even pragmatic functions under specific contextual reminders. Their favorite choices, however, were subject to some limitations.

Furthermore, G3 learners surpassed G2, while G1 learners scored the lowest overall: 78.79% > 65.14% > 63.28%, revealing that study-abroad experience (when combined with both elements' interaction) might result in a better effect, but not the same as proficiency alone. However, it should be noted that this changing pattern may not be observed across all scenarios, such as when a G3 > G2 > G1 trend emerged in Item 1, 6, 7, and 9, but Item 4 and 5 had a G2 > G3 > G1 tendency.

The results of the independent sample t-tests show that there is no significant correlation between proficiency and learners' PC knowledge required to recognize routines: $p > 0.05$; however, study-abroad experience (with both variables combined) was positively correlated with routine recognition with very large effect sizes: both $p < 0.01$, both Cohen's $d > 0.8$.

Likewise, there were no significant differences between Items 1, 4, 6, and 7. Moreover, Items 2 and 8 had a relatively similar trend: participants' performance considerably increased in conjunction with study-abroad experience alone, with small and moderate effect sizes: both $p < 0.05$, Cohen's $d = 0.44$ approximately 0.5, and 0.71 < 0.8. Study-abroad experience was also important in recognizing *Can I get you anything else?*, *Say that again, please*, and *That's okay*, but proficiency was only significantly effective in recognizing *You're welcome*, while the integration of the two factors only produced a significant influence on recognition of *That's okay*.

The outcomes of the decontextualized routine comprehension will be shown in full in the following section.

Table 5.4 Descriptive statistics for routine recognition

ASC (Set.)	Target selections	G1 (N = 51)		G2 (N = 59)		G3 (N = 33)		Overall (n = 143)	
		N	%	N	%	N	%	N	%
1.First encounter	Glad to see you	7	78.43	10	81.36	6	81.82	23	80.42
	Good to run into you	2		1		0		3	
	Happy to find you	2		0		0		2	
	Nice to meet you <input checked="" type="checkbox"/>	40		48		27		115	
2.Cashier	All yours	2	62.75	7	47.46	0	81.82	9	60.84
	Here you go <input checked="" type="checkbox"/>	32		28		27		87	
	Please	4		7		6		17	
	There they are	13		17		0		30	
3.More food	Would you like anything extra?	26	31.37	38	11.86	18	45.45	82	26.57
	Is there more for you?	5		5		0		10	
	What can I do for you?	4		9		0		13	
	Can I get you anything else? <input checked="" type="checkbox"/>	16		7		15		38	
4.Rejection of more food	No, thanks, I'm full <input checked="" type="checkbox"/>	36	70.59	44	74.58	24	72.73	104	72.73
	No, thanks, I've done it	9		9		7		25	
	No, thanks, I've finished it	6		6		2		14	
	No, thanks, I've eaten	0		0		0		0	
5.Response to thanks	Don't bother	5	86.27	2	96.61	3	90.91	10	91.61
	Thank you	2		0		0		2	
	You're welcome <input checked="" type="checkbox"/>	44		57		30		131	
	Please	0		0		0		0	
6.Phone	Hello <input checked="" type="checkbox"/>	32	62.75	38	64.41	24	72.73	94	65.73
	Hi	1		3		6		10	

(continued)

Table 5.4 (continued)

ASC (Set.)	Target selections	G1 (N = 51)		G2 (N = 59)		G3 (N = 33)		Overall (n = 143)	
		N	%	N	%	N	%	N	%
	How are you?	3		3		2		8	
	It's me	15		15		1		31	
7.Message	Can I give you information?	1	98.04	0	98.31	0	100	1	98.60
	Can I leave a message? <input checked="" type="checkbox"/>	50		58		33		141	
	Can you take a note?	0		1		0		1	
	Can you write something?	0		0		0		0	
8.Repeating	Repeat yourself, please	10	60.78	20	50.85	9	72.73	39	59.44
	Say that again, please <input checked="" type="checkbox"/>	31		30		24		85	
	Say that another time, please	7		4		0		11	
	Restate what you said, please	3		5		0		8	
9.Response to apology	Don't mention it	9	35.29	7	44.07	1	90.91	17	51.75
	It's nothing	10		14		1		25	
	No bother	14		12		1		27	
	That's okay <input checked="" type="checkbox"/>	18		26		30		74	
Total		299	65.14	336	63.28	234	78.79	869	67.52

5.3 Results for Routine Comprehension

With respect to the general trend of routine comprehension (Table 5.5), G3 students scored much higher on all three sections than G1 students, and both G3 and G1 markedly surpassed G2 ($G3 > G1 > G2$), thus substantiating that on the whole, study-abroad experience (both factors' interaction) developed hand in hand with learners' pragmatic comprehension of routines. However, the impact of proficiency alone on routine comprehension revealed an opposite trend, namely, comprehensive competence decreased with the improvement of proficiency.

As to the independent sample t-test results, study-abroad experience and its interaction with high proficiency level played influential roles at all phases of routine

Table 5.5 Descriptive statistics for routine comprehension

Group (G)	N	M	SD	Frequency (%)
G1	51	5.24	2.98	37.43
G2	59	4.05	2.76	28.93
G3	33	8.64	3.30	61.71

comprehension without exception—both $p < 0.01$ —and with large effect sizes: both Cohen' $d > 0.8$. In contrast, the impact of proficiency alone revealed a somewhat opposite pattern (a negative pattern), further confirming the more decisive role of study-abroad experience but the weaker impact of proficiency alone in the decontextualized process of routine comprehension.

As to examining the plausibility of the definitions and examples within each group, a McNemar frequency table (Table 5.6) was constructed featuring the number of responses for no PC or ASC, plausible PC, plausible ASC, and the plausible interplay of PC and ASC. McNemar chi-square tests were then used to process these data. When plausible PC and ASC were compared within each group holistically, it was found that participants within each group were more likely to provide plausible meanings based on PC than to make up specific examples in an ASC (46.78% vs. 28.01%, 38.74% vs. 19.13%, and 77.49% vs. 45.89%, all $p < 0.05$).

Table 5.7 summarizes the results of pairwise comparisons as follows. Mann–Whitney U tests were employed to examine the impact of three factors on the quadripartite levels. No significant differences at any level were detected between Groups 1 and 2 (all $p > 0.05$), signifying that proficiency had no meaningful effect on any level of context knowledge required in routine comprehension. Study-abroad experience (G2 vs. G3), by contrast, appeared to be the major influence on each level of context knowledge during the decontextualized comprehension (all $p < 0.05$). The interaction of proficiency and study-abroad experience, as an integral variable, affected the first two sections, i.e., no PC & ASC and plausible PC (both $p < 0.05$), exclusive of the other two levels (both $p > 0.05$).

More precisely, the percent chi-squared of test-takers providing plausible definitions based on PC was significantly different from that of test-takers providing plausible examples in an ASC for the expressions *All yours* (50.98% vs. 17.65% and

Table 5.6 Distribution of No PC or ASC, plausible PC and ASC, and mutual mappings

Level	G1 (Total responses: 357)		G2 (Total responses: 413)		G3 (Total responses: 231)	
	N	%	N	%	N	%
NO PC & ASC	182	50.98	251	60.77	51	22.08
Plausible PC	167	46.78	160	38.74	179	77.49
Plausible ASC	100	28.01	79	19.13	106	45.89
Plausible PC & ASC	92	25.77	77	18.64	105	45.45

Table 5.7 Summary of pairwise comparisons

Variable	Level			
	NO PC & ASC	Plausible PC	Plausible ASC	Plausible PC & ASC
Proficiency	$z = -0.70$	$z = -0.83$	$z = -1.73$	$z = -1.22$
Study-abroad experience	$z = -2.12^*$	$z = -2.38^*$	$z = -2.51^*$	$z = -2.52^*$
Two factors' interaction	$z = -2.12^*$	$z = -2.12^*$	$z = -1.61$	$z = -1.61$

Notes *, $p < 0.05$

40.68% vs. 10.17%, both $p < 0.05$) and *Thanks for having me* (45.10% vs. 17.65% and 50.85% vs. 22.03%, both $p < 0.05$). Additionally, G2 subjects were far more likely to provide plausible definitions than examples for *Here you go!* (27.12% vs. 8.47%, $p < 0.05$). The same was true in G3 for *Excuse the mess* (72.73% vs. 27.27%, $p < 0.05$). By comparison, the remaining items did not present a significant difference between plausible PC and ASC within each group (all $p > 0.05$).

5.4 Results for Routine Perception

Routine perception was divided into two sections: pragmatic awareness and the identification of ASC traits using PC knowledge between the two paired expressions. As shown in Table 5.8, the following routine tasks are the most unsatisfying in contrast to the others: Mean = 5.79, 2.89 for overall routine perception and their level of PC knowledge mastery, frequency = 38.60, 28.9%. Furthermore, pragmatic awareness was also not optimistic: mean = 2.90, frequency = 58%. More crucially, learners' pragmatic distinctive awareness was substantially lower than overall accuracy in identifying functional meanings through their PC knowledge.

Table 5.8 Descriptive statistics for routine perception

Paired routine expressions	Total (n = 143)		
	Mean awareness	Mean distinction	Mean overall
Nice to meet you versus Nice to see you	0.52	0.78	1.30
Hello versus Hi	0.50	0.37	0.87
Watch out versus Be careful	0.43	0.54	0.97
No problem versus You're welcome	0.79	0.80	1.59
Do you have the time versus Do you have a minute	0.66	0.39	1.05
Overall	2.90	2.89	5.79

For the perception task, pragmatic awareness does not always guarantee the realization of differentiating paired routines, implying that even if the subtle differences were perceived in the beginning, learners may still fail to implement the decontextualized distinction of their pragmatic features due to a lack of precise PC knowledge. In terms of specific perception types, students performed best in *No problem* versus *You're welcome* with the score rate just above 50%. However, it is unexpected that they had the weakest response to *Hello* versus *Hi* with the lowest rate of 29.00%.

The independent sample t-test results show that proficiency has no effect on overall routine perception, pragmatic awareness, or the abilities to discriminate ASC traits based on learners' PC knowledge: all $p > 0.05$. Study-abroad experience, on the other hand, had a strong positive impact that was significant at $p < 0.01$, and the between-group difference, represented as Cohen's d , was quite considerable, with all values significantly higher than 0.8.

5.5 Summary

The quantitative results have already been testified and answered based on the data given in this chapter. In general, English proficiency had significant influence on the mastery degree of PC knowledge in contextualized routine production and decontextualized routine comprehension. Almost every routine task demonstrated a high vulnerability to study-abroad experiences with large effect sizes. The results of qualitative data analysis will be discussed in depth in the following chapter, and so will the findings.