AAAI 2026 Supplementary Material Anonymous Submission

Anonymous submission

Abstract

This document provides supplementary materials for the main paper, including the detailed descriptions about the textbooks, the grammar items, the prompt engineering for LLMs to generate instruction data, the pseudocode for the curriculum-tuning framework, the binary classification dataset, the error detection, and the HSKAgent platform.

A. Textbook checklists

To construct the level-based training data, we first collect 79 widely-used textbooks based on HSK levels 3 to 6, such as *HSK Standard Course* and *Boya Chinese Course*. These textbooks are a mixture of texts and images. We manually delete the images since the multimodal inputs are not the objective of this paper. In order to ensure the semantic compactness of the texts, we also delete all the Pinyin and English symbols used to assist learning in the textbooks through scripts. Finally, the total number of tokens in the cleaned textbooks is 6.76M, with 162,074 sentences and an average of 41.74 tokens per sentence.

In addition, we conduct a systematic analysis of these textbooks. The checklist of textbooks corresponding HSK levels 3 and 4 is presented in Table 1. The checklist of textbooks corresponding HSK levels 5 and 6 is presented in Table 2. Overall, the higher the proficiency level, the fewer textbooks are available. Note that some textbooks did not have official English titles at the time of publication. For these textbooks, we provide translated English titles for reference. These translations may not be entirely accurate. Therefore, it is recommended to use the original Chinese titles when searching for the source materials.

B. Checklists of the level-based grammar items

To create a set of instruction data covering various writing exercises, we first integrate HSK levels 3 to 6 and advanced grammar items from *Chinese Proficiency Grading Standards for International Chinese Language Education*. According to the information retrieval system¹ from Chinese Tests Service Website, six types of grammar items are selected because they appear at these levels, including word,

Algorithm 1: Pseudocode of curriculum tuning

```
1: Input: Base model \theta_0, HSK levels \mathcal{L} = \{3, 4, 5, 6\}
 2: for level l \in \mathcal{L} do
 3:
          TextbookData[l] \leftarrow list of sentences for level l
          InstructionData[l] \leftarrow list of (prompt, completion)
 4:
     pairs for level l
 5: end for
 6: \theta \leftarrow \theta_0
                                                 ▶ Initialize base model
 7: for level l \in \mathcal{L} do
          for all sentence \in TextbookData[l] do
 9:
               loss_{nt} \leftarrow PretrainingLoss(\theta, sentence)
10:
               \theta \leftarrow \text{UpdateModel}(\theta, loss_{nt})
11:
          end for
          for all (prompt, completion) \in InstructionData[l] do
12:
               loss_{it} \leftarrow InstructionLoss(\theta, prompt, completion)
13:
               \theta \leftarrow \text{UpdateModel}(\theta, loss_{it})
14:
15:
          end for
16: end for
17: return \theta \triangleright \text{Final } LLM - \theta^{(6)} after curriculum tuning
```

phrase, fixed format (FF), sentence component (SC), sentence type (ST), and emphatic usage (EU). The checklist of grammar items with HSK level 3 is shown in Table 3. The checklist of grammar items with HSK level 4 is shown in Table 4. The checklist of grammar items with HSK level 5 is shown in Table 5. The checklist of grammar items with HSK level 6 is shown in Table 6. The checklist of grammar items with HSK advanced level is shown in Table 7.

C. Prompting LLMs to generate instruction data

We leverage GPT-4.1-mini (Achiam et al. 2023), DeepSeek-Chat-V3 (Liu et al. 2024), and Gemini-2.5-Flash (Team et al. 2023) with robust Chinese capabilities to generate level-based instruction data according to these grammar items using in-context learning with two shots. The LLMs are prompted to generate 10 instruction instances for each grammar item. Each piece of generated data contains an instruction, an input, and an output, where the instruction is the requirement and background of the writing exercise, the input is the specified grammar item, and the output is the expected language production. The prompt, as shown in Figure 1, is

¹https://old.chinesetest.cn/standardsAction.do?means=standardInfo

used to instruct the three LLMs to generate instruction data based on these grammar items.

D. Training LLMs through the curriculum-tuning framework

Llama2-7B-Chat (Touvron et al. 2023), Mistral-7B-Instruct-v0.3 (Jiang et al. 2023), and Chinese-Alpaca-2-7B (Cui, Yang, and Yao 2024) are selected as baselines to be trained using the curriculum-tuning framework. The framework contains three process: pretraining on level-based textbooks for simulating input-based learning, instruction tuning on writing exercises for simulating output-based learning, curriculum tuning across levels. The pseudocode of this training process is abstracted in Algorithm 1.

E. The binary classification dataset for grammar detection

To enable the HSKAgent to detect grammar items from compositions, a binary classification dataset is constructed by transforming the level-based instruction data into positive data and negative data. To reduce the likelihood that the negative completion still aligns with the target grammar item, we restrict sampling to completions drawn from outside the current grammar item category. For the positive samples, the original prompt and completion are concatenated into a new positive prompt, with the corresponding answer "Yes". A positive example is shown as follow:

"instruction": "你是一个HSK语法点检测器,请你对以下句子进行判断是否符合当前语法点: No:133,级别:3级,语法项目:词类,类别:动词,细目:能愿动词,语法内容:敢。",

"**input**": "句子: 即使任务很难, 我也敢尝试完成它。",
"**output**": "是",

"output": "是" }

For the negative samples, the prompt is paired with a negative completion randomly sampled from the data pool, resulting in a new negative prompt with the answer "No". A negative example is shown as follow:

"instruction": "你是一个HSK语法点检测器,请你对以下句子进行判断是否符合当前语法点: No:133,级别:3级,语法项目:词类,类别:动词,细目:能愿动词,语法内容:敢。",

"input": "句子: 我最喜欢春节,它不仅仅是鞭炮声声的热闹,是家家户户张灯结彩的喜庆,更是它背后所承载的团圆文化是深厚的,辞旧迎新的寓意是美好的。我对春节的喜爱是如此强烈,因为它带来的,是家的温暖,是年的味道,是所有中国人共同的记忆。".

"output": "否",

}

Following this pipeline, the binary classification dataset contains an equal number of positive and negative data, which is then used to finetune the HSKAgent.

F. The error detection of HSKAgent

Error detection is one of our HSKAgent's functions for evaluating the accuracy of LLMs' language production in writing during the process of Chinese SLA. After finetuning, the HSKAgent is able to detect errors for new compositions. An example is shown as follow:

The original content of a composition:

我从小到大,对我影响最大的一个人是我的母亲。我的母亲是典形家庭妇女。但我认为伟在这个是是最伟大的人。不因为她生了我,一个人是我的母界上最伟大的人。不因为她生了我,一个人是我的母界。我的角度上说是最伟大。我母亲是一个典形为有人。我母亲是一个典别名就觉得他们看不起我的母亲是多么健强的人。对于他们看到我知道是一听这个别名就觉得他们看不起我和自己的好的母亲是多么健强的人。对于他对我们是一个动来教我。可以说是那时开始之后的好的。如此我深刻地懂'责任感'。有一天把我打得 (end)

The version for error detection and correction:

我从小到大,对我影响最大的一个人是我的母 亲。我的母亲是典型[B形]{CQ的}家庭妇女,[BC。]但 我认为在这个世界上{CQ她是}最伟大的人。不只 是 $\{CC$ 因为 $\}$ 她生了我,才说最伟大 $\{CQ$ 的 $\}[BQ,]$ 而 是{CO从}客观的角度上说是最伟大{CO的}。我母亲 是个典型[B形]{CQ的}乡下人,[BC。]已经在大城市 里住了三十年,还带着土味对待城市人。所以很 多人叫她"[BC']城市的乡下佬[B老]儿"[BC'],[BC。]她 也很喜欢听别人这样称呼她。但我呢?倒{CD也}不 喜欢听这个别名。因为一听{CQ到}这个别名就觉得 他们看不起我的母亲。我知道母亲是多么坚图健图 的人。我从小开始[BO,]她{CO就}教我知道什么是 自立和独立,什么是{CD叫}责任感。这些{CC2这}都 是她{CQ用}自己的语言和行动来教我。可以说是那 时{CQ我}开始学习{CC学了}怎么做好的人。她让我 深刻地懂{CQ得}"[BC']责任感"[BC']。有一天把我打 得{WWJ}

Based on the collected compositions written by human L2 learners, all error types and an error handling guideline are published in our anonymous repository².

G. The platform of HSKAgent

To use this HSKAgent efficiently to assess the Chinese SLA development in LLMs, HSKAgent is deployed on a platform. The functional diagram of the platform is shown in Figure 2 to 5. In the near future, we plan to launch an online version of this platform, aiming to provide professional and accurate writing assessment services to a wide range of Chinese learners.

²https://anonymous.4open.science/r/HSKB-5E70

Textbook Checklist (HSK level 3 to 4)						
HSK 3		HSK 4				
Titles	Tokens	Titles	Tokens			
考试大纲 HSK3 Exam Outline HSK3	12798	考试大纲 HSK4 Exam Outline HSK3	14017			
词汇 HSK3 Vocabulary HSK3	9931	词汇 HSK4 Vocabulary HSK4	20844			
教师用书 HSK3 HSK3 Teacher's Book	109898	教师用书HSK4上 HSK4 Teacher's Book 01	78372			
标准教程3 HSK HSK Standard Course 3	49825	教师用书HSK4下 HSK4 Teacher's Book 02	89045			
中文4 Chinese 4	23372	HSK标准教程4上 HSK Standard Course 4 01	49702			
中文5	29701	HSK标准教程4下	60742			
Chinese 5 中文6	35980	HSK Standard Course 4 02 中文7	34967			
Chinese 6 初级综合-01	81155	Chinese 7 中文8	38897			
Elementary Comprehensive-01 初级综合-02	108997	Chinese 8 中级写作-01	46656			
Elementary Comprehensive-02 初级读写-01	39899	Intermediate Writing-01 中级写作-02	49390			
Elementary Reading and Writing-01 初级读写-02	30140	Intermediate Writing-02 中级综合-01	90602			
Elementary Reading and Writing-02 博雅汉语01	83216	Intermediate Comprehensive-01 中级综合-02	157269			
Boya Chinese 01 博雅汉语02	69563	Intermediate Comprehensive-02 中级阅读-01	82834			
Boya Chinese 02 新HSK词汇突破3级	29645	Intermediate Reading-01 中级阅读-02	83005			
New HSK Vocabulary Breakthrough HSK3 语法点速记3-4级	180917	Intermediate Reading-02 阅读写作HSK4(刘云)	61067			
Grammar Item Shorthand Level 3-4		Reading and Writing HSK4 (Liuyun) 博雅汉语03	120001			
		Boya Chinese 03 博雅汉语04	47060			
		Boya Chinese 04 四级应试指南	140688			
		Level 4 Exam Guide 新HSK词汇突破4级	27441			
		New HSK Vocabulary Breakthrough HSK4 语法点速记3-4级 Grammar Item Shorthand Level 3-4	180917			
Total	895037	Total	1473516			

Table 1: The checklist of textbooks with HSK level 3 to 4 (Appendix A).

Textbook Checklist (HSK level 5 to 6)							
HSK 5		HSK 6					
Titles	Tokens	Titles	Tokens				
考试大纲 HSK5		21天征服6级阅读	0=0.50				
Exam Outline HSK5	21130	21Days Reading HSK6	87960				
词汇 HSK5	=0=2	21天征服6级写作	00640				
Vocabulary HSK5	7973	21Days Writing HSK6	80642				
教师用书HSK5上	02142	考试大纲 HSK6	27.425				
HSK5 Teacher's Book 01	93142	Exam Outline HSK6	37425				
教师用书HSK5下	01246	词汇 HSK6	51511				
HSK5 Teacher's Book 02	91246	Vocabulary HSK6	54514				
HSK标准教程5上	61539	教师用书HSK6上	175899				
HSK Standard Course 5 01	01339	HSK6 Teacher's Book 01	173899				
HSK标准教程5下	63188	教师用书HSK6下	190876				
HSK Standard Course 5 02	03100	HSK6 Teacher's Book 02	190670				
HSK语法点速记速练(高级篇)	156408	HSK标准教程6上	99274				
HSK Grammar Item Quick Practice (Advanced)	130400	HSK Standard Course 6 01	JJ214				
中文10	51916	HSK标准教程6下	112459				
Chinese 10	31910	HSK Standard Course 6 02	112433				
中文9	47366	中文11	51470				
Chinese 9	47300	Chinese 11	31470				
中级写作-01	46656	中文12	59157				
Intermediate Writing-01	10050	Chinese 12	37137				
中级写作-02	49390	六级应试指南	253116				
Intermediate Writing-02	.,,,,	Level 6 Exam Guide	200110				
中级综合-01	90602	写作 HSK6 (刘云)	101735				
Intermediate Comprehensive-01		Writing HSK6 (Liuyun)					
中级综合-02	157269	阅读 HSK6 (刘云)	191351				
Intermediate Comprehensive-02		Reading HSK6 (Liuyun)					
中级阅读-01	82834	博雅汉语07	87443				
Intermediate Reading-01		Boya Chinese 07					
中级阅读-02	83005	博雅汉语08	164812				
Intermediate Reading-02		Boya Chinese 08					
五级应试指南	63365	6级语法点(外研社)	172650				
Level 5 Exam Guide		Level 6 Grammar Items (Waiyanshe)					
写作 HSK5 (刘云) Writing HSK5 (Liuyun)	40196	新HSK词汇突破6级	134159				
网读 HSK5 (刘云)		New HSK Vocabulary Breakthrough Level 6					
	147227	高级写作-01	54769				
Reading HSK5 (Liuyun) 博雅汉语05		Advanced Writing-01 高级写作-02					
Boya Chinese 05	159425	同級与11-02 Advanced Writing-02	61396				
博雅汉语06		高级综合-01					
Boya Chinese 06	139426	Advanced Comprehensive-01	130526				
新HSK词汇突破5级		高级综合-02					
別 ロSK 岬に大根 35及 New HSK Vocabulary Breakthrough Level 5	63875	미의보기 급 -02 Advanced Comprehensive-02	150498				
1.0. 11511 rocuomai y Dicaminough Level 5		高级阅读-01					
		Advanced Reading-01	106868				
		高级阅读-02					
		Advanced Reading-02	119622				
Total	1717178	Total	2678621				

Table 2: The checklist of textbooks with HSK level 5 to 6 (Appendix A).

			G	rammar Items (HSK	Level 3)
No.	Level	Item	Туре	Detail	Content
133	三级	词类	动词	能愿动词	敢
134	三级	词类	动词	能愿动词	需要
135	三级	词类	动词	离合词 (动宾式)	帮忙、点头、放假、干杯、见面、结婚、 看病、睡觉、洗澡、理发、说话
136	三级	词类	动词	离合词 (动补式)	打开、看见、离开、完成
137	三级	词类	代词	疑问代词	疑问代词的非疑问用法: (1) 任指用
					法: 疑问代词+都/疑问代词疑
138	三级	词类	代词	指示代词	问代词; #(2) 不定指用法 各、各种、各个、每、任何
139	三级	词类	量词	名量词	把、行、架、群、束、双、台、张、支、
	<i>→ 137</i>) = 1) I/.	E /-	-1 E /=	只、种品
140 141	三级	词类 词类	量词 量词	动量词 量词重叠	顿、口、眼 量词重叠:AA
141	二级	词类	副词	程度副词	里两里看:AA 比较、更、还、相当
143	三级 三级 三级	词类	副词	范围、协同副词	光、仅、仅仅、就、至少
144	三级	词类	副词	时间副词	本来、才、曾经、从来、赶紧、赶快、立
	<i>→ 1377</i>	\ = 1.44	로마크	医素 委员司马	刻、连忙、始终、已、早已
145	三级 三级	词类 词类	副词 副词	频率、重复副词 关联副词	通常、往往、总、总是 再
146 147	三级三级	词类	副词	方式副词	五相、尽量、亲自、相互 五相、尽量、亲自、相互
148	三级	词类	副词	情态副词	大概、恐怕
149	三级	词类	副词	语气副词	白、并、当然、到底、反正、根本、果
					然、简直、绝对、难道、其实、千万、确实、几亿、绝关
150	三级	词类	介词	引出时间、处所	实、只好、终于 由1
151	三级	词类	介词	引出时间、处所	自从
152	三级三级	词类	介词	引出方向、路径	朝
153	三级	词类	介词	引出对象	为
154 155	三级三级	词类 词类	介词 介词	引出对象 引出目的、原因	向 由于、因为
156	三级	词类	介词	引出目的、原因引出目的、原因	为了
157	三级	词类	介词	引出施事、受事	把、被、叫、让
158	三级	词类	介词	表示排除	除了
159	三级	词类	介词	引出凭借、依据	按、按照
160	三级	词类	连词	连接分句或句子	并且、不光、不仅、另外、要是、于是、 因此、由于、只有
161	三级	词类	拟声词		哈哈
162	三级	短语	结构类型	其他结构类型	其他结构类型2: ①介宾短语#②方位短
	<i>→ 137</i>	<i>L</i> = \∓	/		语#③兼语短语#④同位短语
163 164	三级 三级	短语 短语	结构类型 固定短语	四字格	数量重叠:数词+量词+数词+量词 不A不B
165	三级三级	短语	固定短语	其他	和ANB 看起来
166	三级	短语	固定短语	其他	看上去
167	三级	短语	固定短语	其他	有的是
168	三级	固定格式			除了(以外),还/也/都
169	三级三级	固定格式			从起
170 171	二级	固定格式 固定格式			对来说 像一样
171	三级 三级	固定格式			越越
	.,,,,	–			· = · · · · · · = · · · · · ·

			Gramma	r Items (HSK leve	el 3) (continued)
No.	Level	Item	Туре	Detail	Content
173	三级	句子成分	主语		动词或动词性短语作主语 # 形容词或形容 词性短语作主语
174	三级	句子成分	宾语		动词或动词性短语作宾语 # 形容词或形容词性短语和主谓短语作宾语
175	三级	句子成分	定语		动词或动词性短语作定语 # 主谓短语作定 语
176 177	三级三级	句子成分 句子成分	补语 补语	结果补语 趋向补语	结果补语:动词+到/住/走 复合趋向补语的趋向意义用法:动词+出 来/出去/过去/过来/回来/进去/进来/起来/上 来/上去/下来/下去
178	三级	句子成分	补语	可能补语	可能补语: 动词+得/不+动词/形容词 # 动词+得/不+了
179	三级	句子成分	补语	程度补语	程度补语:形容词/心理动词+得很/极了/死了
180	三级	句子成分	补语	数量补语	数量补语(动词+数量补语):宾语和数 量补语共现
181	三级	句子成分	补语	数量补语	数量补语(动词+时量补语):表示动作 持续的时间
182	三级	句子成分	补语	数量补语	数量补语(动词+时量补语):表示动作 结束后到某个时间点的间隔时间
183 184	三级三级	句子的类型 句子的类型	句型 特殊句型	单句 "把"字句	主谓句4: 主谓谓语句 "把"字句: 表处置(1)主语+把+宾语+动 词+在/到+处所 # "把"字句: (2)主 语+把+宾语1+动词(+给)+宾语2 #"把"字
185	三级	句子的类型	特殊句型	被动句	句: (3) 主语+把+宾语+动词+结果补语/趋向补语/状态补语被动句: 主语+被/叫/让+宾语+动词+其他成分
186	三级	句子的类型	特殊句型	连动句	连动句: (1) 前一动作是后一动作的方式 #连动句: (2) 后一动作是前一动作的
187	三级	句子的类型	特殊句型	兼语句	目的 兼语句 表使令:主 语+叫/派/请/让+宾语1+动词+宾语2
188	三级	句子的类型	特殊句型	比较句	比较句: (1) A比B+动词+得+形容词 #比较句: (2) A不比B+形容词#比较句: (3) A+动词+得+比+B+形容词#比较句:
100	<i>→ 4</i> 77.	クマか来到		丢动力	(4)A比B+多少/早/晚+动词+数量短语
189 190	三级 三级	句子的类型 句子的类型	特殊句型 复句	重动句 并列复句	主语+动词+宾语+动词+补语 (也).....,也.....
191	三级	句子的类型	复句	并列复句	一会儿,一会儿
192	三级	句子的类型	复句	并列复句	一方面,另一方面
193	三级	句子的类型	复句	并列复句	又, 又
194	三级 三级	句子的类型 句子的类型	复句 复句	承接复句 递进复句	首先,然后 ,并且
195 196	三级 三级	可于的类型 句子的类型	夏可 复句	遊进复句 递进复句	,开旦 不仅/不光,还/而且
197	三级	句子的类型	复句	选择复句	不是,就是
198	三级	句子的类型	复句	转折复句	X是X,就是/不过
199	三级	句子的类型	复句	假设复句	要是,就
200 201	三级 三级	句子的类型 句子的类型	复句 复句	条件复句 因果复句	只有,才 (由于,)所以/因此
202	三级	句子的类型	复句	目的复句	为了

	Grammar Items (HSK level 3) (continued)							
No.	Level	Item	Type	Detail	Content			
203 205 206	三级 三级 三级	句子的类型 强调的方法 强调的方法	复句	紧缩复句	了(就) 用"一点儿也不"表示强调 用反问句表示强调 反问句1:不 是吗? /难道吗?			
207	三级	强调的方法			用"是"表示强调			

Table 3. The checklist of grammar items with HSK level 3 (Appendix B).

			Gr	ammar Items (HSK	K level 4)
No.	Level	Item	Туре	Detail	Content
212	四级	词类	动词	能愿动词	得
213	四级	词类	代词	人称代词	人家
214	四级	词类	量词	名量词	打、袋、根、卷、裸、批
215	四级	词类	量词	借用量词	(1) 名量词: 碗、脸、手、屋子、桌子; #(2) 动量词: 刀、针
216	四级	词类	副词	程度副词	格外、极、极其
217	四级	词类	副词	范围、协同副词	共
218	四级	词类	副词	时间副词	按时、即将、急忙、渐渐、尽快
219	四级	词类	副词	频率、重复副词	一再、再三
220	四级	词类	副词	关联副词	却
221	四级	词类	副词	否定副词	未必
222	四级	词类	副词	情态副词	几乎、似乎
223	四级	词类	副词	语气副词	的确、反而、还、竟然、究竟
224	四级	词类	介词	引出时间、处所	自
225	四级	词类	介词	引出对象	对于
226	四级	词类	介词	引出对象	关于
227	四级	词类	介词	引出对象	替
228	四级	词类	介词	引出凭借、依据	根据
229	四级	词类	介词	引出凭借、依据	作为
230	四级	词类	连词	连接词或词组	并、以及
231	四级	词类	连词	连接分句或句子	此外、而、而是、既然、可见、甚至、假
					如、总之
232	四级	词类	助词	其他助词	似的
233	四级	词类	叹词		啊
234	四级	短语	固定短语	四字格	大A大B
235	四级	短语	固定短语	四字格	-A-B
236	四级	短语	固定短语	其他	看来
237	四级	短语	固定短语	其他	来得及/来不及
238	四级	短语	固定短语	其他	说不定
239	四级	_短语	固定短语	其他	一般来说
240	四级	固定格式			一+量词+比+一+量词
241	四级	固定格式			(自)以来
242	四级	固定格式			由组成
243	四级	固定格式			在方面
244	四级	固定格式			在上/下/中
245	四级	句子成分	主语	主谓短语作主语	主谓短语作主语
246	四级	句子成分	主语	受事主语	受事主语
247	四级	句子成分	定语	多项定语	多项定语

			Gramma	r Items (HSK leve	14) (continued)
No.	Level	Item	Type	Detail	Content
248	四级	句子成分	补语	趋向补语	趋向补语3 表示结果意义(引申用法): 动词+上/出/起/下
249	四级	句子的类型	特殊句型	"把"字句	"把"字句: 表处置 (1) 主语+把+宾语+动词 (+一/了) +动词 #"把"字句 (2) 主语+把+宾语 (+给) +动词+了/着 #"把"字句 (3) 主语+把+宾语+动词+动量补语/时量补语
250	四级	句子的类型	特殊句型	被动句	被动句2: 主语+被+动词+其他成分
251	四级	句子的类型	特殊句型	存现句	存现句2: (1)表示出现:处所词+动词+趋向补语/结果补语+动态助词(了)+数量短语+人物 #存现句2: (2)表示消失:处所词+动词+结果补语+动态助词(了)+数量短语+人物
252	四级	句子的类型	特殊句型	兼语句	兼语句2(1)表爱惜义:主语+表扬/批评+宾语1+动词+宾语2 #兼语句2(2)表称谓或认定义:主语+叫/称 (呼)/说/认/选+宾语1+做/为/当/是+宾语2
253	四级	句子的类型	特殊句型	"是的"句	"是的"句2: 强调说话人的看法或态度
254	四级	句子的类型	复句	并列复句	不是 而是
255	四级	句子的类型	复句	并列复句	既, 又/也
256	四级	句子的类型	复句	承接复句	首先, 其次
257	四级	句子的类型 句子的类型	复句 复句	承接复句 递进复句	,于是 ,甚至
258259	四级 四级	句子的类型	复印复句	选择复句 选择复句	,甚至 或者,或者
260	四级	句子的类型	复句	转折复句	·····,
261	四级	句子的类型	复句	假设复句	,否则
262	四级	句子的类型	复句	假设复句	假如, (就)
263	四级	句子的类型	复句	假设复句	万一, (就)
264	四级	句子的类型	复句	条件复句	不管 都/也
265	四级	句子的类型	复句	条件复句	无论, 都/也
266	四级	句子的类型	复句	因果复句	既然, 就
267	四级	句子的类型 句子的类型	复句	因果复句	,可见 哪怕,也/还
268269	四级 四级	句子的类型 句子的类型	复句 复句	让步复句 目的复句	哪怕,也/处 ,好
270	四级	句子的类型	复印复句	写的复页 紧缩复句	,好 无标记
271	四级	句子的类型	复句	紧缩复句	不也
274	四级	强调的方法	~: 3	ANIH X	用反问句表示强调 反问句2:由疑问代词 构成的反问句
275	四级	强调的方法			用双重否定表示强调
276	四级	强调的方法			用"一+量词(+名词)+也(都)/也没 (不)"表示强调
277	四级	强调的方法			用"连也/都"表示强调

Table 4. The checklist of grammar items with HSK level 4 (Appendix B).

	Grammar Items (HSK level 5)							
No.	Level	Item	Туре	Detail	Content			
288	五级	词类	代词	指示代词	彼此、如此			
289	五级	词类	量词	名量词	册、朵、幅、届、颗、匹、扇			
290	五级	词类	副词	程度副词	过于、可、稍、稍微、尤其			
291	五级	词类	副词	范围、协同副词	大都			
292	五级	词类	副词	时间副词	不时、将、将要、仍旧、时常、时刻、依			
		. 150	H1,1	.11.111.1	旧、一向			
293	五级	词类	副词	频率、重复副词	偶尔、再次			
294	五级	词类	副词	方式副词	偷偷			
295	五级	词类	副词	语气副词	毕竟、不免、差(一)点儿、倒是、干 脆、就、居然、可、明明、总算			
296	五级	词类	介词	引出时间、处所	随着			
297	五级	词类	介词	引出目的、原因	将			
298	五级	词类	介词	引出施事、受事	由			
	五级	词类		引出凭借、依据	凭			
299			介词					
300	五级	词类	介词	引出凭借、依据	依据			
301	五级	词类	介词	引出凭借、依据	依照			
302	五级	词类	介词	引出凭借、依据	依照			
303	五级	词类	连词	连接分句或句子	从而、加上、完了、一旦			
304	五级	词类	助词	其他助词	也好			
305	五级	短语	固定短语	四字格	A来A去			
306	五级	短语	固定短语	四字格	A着A着			
307	五级	短语	固定短语	四字格	没A没B			
308	五级	短语	固定短语	四字格	有A有B			
309	五级	短语	固定短语	其他	不得了			
310	五级	短语	固定短语	其他	不敢当			
311	五级	短语	固定短语	其他	得了			
312	五级	短语	固定短语	其他	用不着			
313	五级	固定格式			从来看			
314	五级	固定格式			到为止			
315	五级	固定格式			够的			
316	五级	固定格式			拿来说			
317	五级	固定格式			A的A,B的B			
318	五级	固定格式			在看来			
319	五级	句子成分	宾语	宾语	宾语的语义类型1: (1) 施事宾语; # 宾 语的语义类型1: (2) 受事宾语			
320	五级	句子成分	状语	多项状语	多项状语			
321	五级	句子成分	补语	趋向补语	趋向补语4表示时间意义(引申用法)			
321	11:100	H)]]]X()]	小垣	旭川作品	(1) 表示动作行为的开始: 动词+上/起来			
					#趋向补语(2)表示动作行为的持续:动词+下去/下来			
322	五级	句子成分	补语	可能补语	可能补语2: 动词+得/不得			
323	五级	句子成分	补语	程度补语	程度补语2: (1) 形容词/心理动词+得+不			
525	-14-1/X	1 14V/1	11 74		得了/慌/厉害;#程度补语2: (2) 动词/形容词+坏/透+了			
324	五级	句子成分	补语	状态补语	状态补语2: 动词/形容词+得+短语(1)动			
					词/形容词+得+动词短语 #状态补语2: (2)动词/形容词+主谓短语 #状态补			
					语2: (3) 动词/形容词+得+固定短语			
325	五级	句子的类型	特殊句型	"有"字句	"有"字句3: (1) 表示存在、具有: 主			
525	-1-7A	577月7人土	147/14/31五	11 1 1.7	语+有+着+宾语; #"有"字句3: (2) 表示			
					附着: 主语+动词+有+宾语			

			Grammar	Items (HSK lev	el 5) (continued)
No.	Level	Item	Type	Detail	Content
326	五级	句子的类型	特殊句型	"把"字句	"把"字句3: 表处置(1)主语+把+宾 语+状语+动词 #"把"字句3: (2)主 语+把+宾语+一+动词 #"把"字句3: (3) 主语+把+宾语+动词+了 #"把"字句3: (4)主语+把+宾语1+动词+宾语2
327	五级	句子的类型	特殊句型	被动句	被动句3: 意念被动句
328	五级	句子的类型	特殊句型	连动句	连动句3: 前后两个动词性词语具有因果、转折、条件关系
329	五级	句子的类型	特殊句型	兼语句	兼语句3 表致使: 主语+叫/令/使/让+人称 代词+动词短语
330	五级	句子的类型	特殊句型	比较句	比较句5: (1) 跟相比 #比较句5: (2) A+形容词+B+数量补语
331	五级	句子的类型	复句	选择复句	或是 或是
332	五级	句子的类型	复句	转折复句	尽管, 但是/可是
333	五级	句子的类型	复句	假设复句	一旦,就
334	五级	句子的类型	复句	假设复句	要是(就),否则
335	五级	句子的类型	复句	条件复句	除非,才
336	五级	句子的类型	复句	条件复句	除非,否则/不然
337	五级	句子的类型	复句	因果复句	因而
338	五级	句子的类型	复句	让步复句	即使,也
339	五级	句子的类型	复句	目的复句	,为的是
340	五级	句子的类型	复句	目的复句	,以便
341	五级	句子的类型	复句	紧缩复句	没有就没有
342	五级	句子的类型	复句	紧缩复句	再也
343	五级	句子的类型	复句	多重复句	二重复句1:单句+复句;复句+单句
344	五级	强调的方法			用"再也不/没"表示强调
345	五级	强调的方法			用副词"可"表示强调
346	五级	强调的方法			用"怎么都/也+不/没"表示强调

Table 5. The checklist of grammar items with HSK level 5 (Appendix B).

	Grammar Items (HSK level 6)							
No.	Level	Item	Type	Detail	Content			
361	六级	词类	代词	指示代词	本、此			
362	六级	词类	量词	名量词	餐、串、滴、副、股、集、枝			
363	六级	词类	量词	动量词	番、声、趟			
364	六级	词类	副词	程度副词	特、异常			
365	六级	词类	副词	范围、协同副词	尽、净、一齐、一同			
366	六级	词类	副词	时间副词	时时、一时、早晚			
367	六级	词类	副词	关联副词	便			
368	六级	词类	副词	方式副词	不禁、赶忙、亲眼、特地、特意			
369	六级	词类	副词	情态副词	仿佛			
370	六级	词类	副词	语气副词	才3、刚好、偏、恰好			
371	六级	词类	介词	引出时间、处所	于			
372	六级	词类	介词	引出方向、路径	沿(着)			
373	六级	词类	介词	引出对象	同1、与1			
374	六级	词类	介词	引出对象	至于			
375	六级	词类	介词	引出目的、原因	因			

376	Level	Item			
		Ittili	Type	Detail	Content
277	六级	词类	介词	表示排除	除
377	六级	词类	介词	引出凭借、依据	据
378	六级	词类	连词	连接词或词组	而2、同2、与2
379	六级	词类	连词	连接分句或句子	不料、可3、若
380	六级	词类	助词	结构助词	所
381	六级	词类	助词	语气助词	罢了、啦、嘛
382	六级	短语	结构分类型	基本结构类型	数词+形容词+量词
383	六级	短语	固定短语	四字格	或A或B
384	六级	短语	固定短语	四字格	无A无B
385	六级	短语	固定短语	四字格	A这A那
386	六级	短语	固定短语	四字格	左A右B
387	六级	短语	固定短语	其他	不怎么
388	六级	短语	固定短语	其他	不怎么样
389	六级	短语	固定短语	其他	好(不)容易
390	六级	短语	固定短语	其他	那倒(也)是
391	六级	短语	固定短语	其他	就是说/这就是说
392	六级 六级	短语	固定短语	其他	算了 A→+量词,B→+量词
393	ハ级	固定格式	固定格式		
394 395	六级	固定格式 固定格式	固定格式 固定格式		东一A,西一A 为了而
393 396	六级	回足俗式 句子成分	回 定 俗 八 宾 语	宾语	万 · · · · · · · · · · · · · · · · · · ·
390	八级	H)] /////	六归	共归	的语义类型2: (1) 处例共后 #共后的语义类型2: (2) 结果宾语
397	六级	句子成分	补语	趋向补语	趋向补语5表示状态意义(引申用法):
	,	3 4 79433	,,,,,	70171174	动词/形容词+下来了/下去/起来/过来
398	六级	句子的类型	特殊句型	"把"字句	"把"字句4: 表致使(1)主语(非生物
					体)+把+宾语+动词+其他成分#"把"字
					句4: (2) 主语+把+宾语(施事)+动
) . (27)		11 74. (-) TI	<u> </u>	词+其他成分
399	六级	句子的类型	特殊句型	被动句	被动句4:主语+被/叫/让+宾语+给+动词+其他成分
400	六级	句子的类型	复句	并列复句	时而,时而
400	六级	句子的类型	复句	并列复句	一时一时
402	六级	句子的类型	复句	承接复句	便
403	六级	句子的类型	复句	递进复句	
404	六级	句子的类型	复句	递进复句	不是, 还/还是
405	六级	句子的类型	复句	递进复句	连
406	六级	句子的类型	复句	选择复句	要么 要么
407	六级	句子的类型	复句	转折复句	虽,但/可/却/也
408	六级	句子的类型	复句	假设复句	, 要不然/不然
409	六级	句子的类型	复句	条件复句	凡是,都
410	六级	句子的类型	复句	让步复句	就算/就是也
411	六级	句子的类型	复句	紧缩复句	不不
412	六级	句子的类型	复句	多重复句	二重复句2: 复句+复句
413	六级	强调的方法			用"非不可"表示强调

Table 6. The checklist of grammar items with HSK level 6 (Appendix B).

	Grammar Items (HSK advanced level)							
No.	Level	Item	Type	Detail	Content			
426	高等	词类	动词	能愿动词	需			
427	高等	词类	代词	疑问代词	何			
428	高等	词类	代词	指示代词	该、另、兹			
429	高等	词类	量词	名量词	(1) 栋、粒、枚、则、盖 # (2) 复合量 词: 人次			
430	高等	词类	副词	关联副词	亦			
431	高等	词类	介词	引出方向、路径	顺着			
432	高等	词类	连词	连接词或词组	及			
433	高等	词类	连词	连接分句或句子	继而、要不是			
434	高等	词类	助词	结构助词	之			
519	高等	词类	副词	程度副词	极为			
520	高等	词类	副词	程度副词	尽 蛮			
521	高等	词类	副词	程度副词				
522	高等	词类	副词	程度副词	颇			
523	高等	词类	副词	程度副词	稍稍			
524	高等	词类	副词	程度副词	尤为			
525	高等	词类	副词	程度副词	越发			
528	高等	词类	副词	范围、协同副词	凡			
529	高等	词类	副词	范围、协同副词	皆			
530	高等	词类	副词	范围、协同副词	统统			
531	高等	词类	副词	范围、协同副词	唯独			
532	高等	词类	副词	方式副词	不由得			
533	高等	词类	副词	方式副词	一连			
534	高等	词类	副词	方式副词	顺便			
535	高等	词类	副词	否定副词	未			
536	高等 亨笠	词类 词类	副词	否定副词 频率 重复副词	勿 频频			
537 538	高等 高等	诃笑 词类	副词 副词	频率、重复副词 频率、重复副词	观观 再度			
543	同守 高等	词类	副词	一 一 一 情 态 副 词	必定			
544	高等	词类	副词	情态副词	不妨			
545	高等	词类	副词	情态副词	何必			
546	高等	词类	副词	情态副词	莫非			
547	高等	词类	副词	情态副词	按说			
548	高等	词类	副词	时间副词	即			
549	高等	词类	副词	时间副词	历来			
550	高等	词类	副词	时间副词	尚			
551	高等	词类	副词	时间副词	向来			
552	高等	词类	介词	引出对象	当着			
553	高等	词类	介词	引出对象	就5			
554	高等	词类	介词	引出凭借、依据	趁			
555	高等	词类	介词	引出凭借、依据	基于			
556	高等	词类	介词	引出凭借、依据	依			
557	高等	词类	副词	语气副词	白白			
558	高等	词类	副词	语气副词	反倒			
559	高等	词类	副词	语气副词	分明			
560	高等	词类	副词	语气副词	怪不得			
561	高等 亨笠	词类	副词	语气副词	好在			
562	高等 亨笠	词类	副词	语气副词	力			
563	高等 三笑	词类 词类	副词	语气副词 语气副词	难怪			
564	高等	词类	副词	语气副词	偏偏			

No. Level Item Type Detail Content 565 高等 词类 副词 语气副词 索性 566 高等 词类 副词 语气副词 东方 567 高等 词类 副词 语气副词 未免 568 高等 词类 副词 语气副词 未免 569 高等 词类 副词 语气副词 未免 569 高等 词类 副词 语气副词 无好 570 高等等 词类 副词 语气副词 表好 571 高等等 词类 副词 语气机制 要 第 第 第 第 第 第 第 第 第 第 第 第 第 第 第 第 未半 <t< th=""><th colspan="7">Grammar Items (HSK advanced level) (continued)</th></t<>	Grammar Items (HSK advanced level) (continued)						
566 高等 词类 副词 语气副词 东免 567 高等 词类 副词 语气副词 未免 568 高等 词类 副词 语气副词 无非 569 高等 词类 副词 语气副词 来安 570 高等 词类 副词 语气副词 李亏 571 高等 词类 副词 语气副词 李亏 571 高等 词类 副词 语气副词 李亏 571 高等 词类 副词 语气副词 终究 572 高等 词类 勘词 语气制词 终究 573 高等 词类 助词 语气助词 参究 知十量河 抽象 事 43 高等 知十量河 数元 数元 本结构类型 数词+量词+抽象事物 要名 30 未4半A半B 五等 数名不A 半A半B 四字格 非本非路 44 44 高等 短语 四字格 非本非路 独名 独名 基本结节 四字格 四字格 44日							
566 高等 词类 副词 语气副词 东免 567 高等 词类 副词 语气副词 未免 568 高等 词类 副词 语气副词 无非 569 高等 词类 副词 语气副词 来安 570 高等 词类 副词 语气副词 李亏 571 高等 词类 副词 语气副词 李亏 571 高等 词类 副词 语气副词 李亏 571 高等 词类 副词 语气副词 终究 572 高等 词类 勘词 语气别词 终究 573 高等 词类 数词 量 数别 量 数别 量 数别 量 数别 量 数别 #量 对 对 <							
567 高等 词类 副词 语气副词 未免 568 高等 词类 副词 语气副词 未免 569 高等 词类 副词 语气副词 幸好 570 高等 词类 副词 语气副词 幸好 571 高等 词类 助词 语气副词 李究 572 高等 词类 助词 语气副词 李察 573 高等 词类 助词 据令 如己 会察 元 司 大 公司 会察 元 日 公司 会察 元 日 表 知少 十 公司 会 公名公 公司 当 本 本名本各 当 本名本各 国 生 本名本各 国 生 本名本各 国 要 本名本各 国 事 生 本名本各 国 事 本名本各 国 国 本名本各 国 事 生 本名本各 国 主 本名本普 国 主 本名本等 <td< td=""><td></td></td<>							
568 高等 词类 副词 语气副词 无非 李好 569 高等 词类 副词 语气副词 李好 570 高等 词类 副词 语气副词 李安 571 高等 词类 助词 语气制词 终究 572 高等 词类 助词 语气制词 终究 573 高等 词类 助词 接气助词 矣 435 高等 短语 齿皮短语 数周十量词+抽象事物 436 高等 短语 固定短语 型字格 要A不A 437 高等 短语 固定短语 四字格 来A半B 438 高等 短语 国定短语 四字格 东A西B 439 高等 短语 国定短语 四字格 东A西B 448 448 高等 短语 四字格 近路 五本本非路 四字格 440 高等 短语 四字格 近路 世本市 日本市							
569 高等 词类 副词 语气副词 幸好 570 高等 词类 副词 语气副词 幸亏 571 高等 词类 副词 语气副词 幸亏 572 高等 词类 助词 语气助词 疾 573 高等 短语 助词 接 数词+量词+抽象事物 573 高等 短语 齿定短语 型字格 数词+量词+抽象事物 435 高等 短语 固定短短语 四字格 光本半B 表不A 436 高等 短语 固定短短语 四字格 半A半B 本名四字格 半A半B 本名四字格 半A半B 五半 44B 高等 短语 面定短短语 四字格 本名四路 基本4B 44B 基本4事B 2A公图B 444 自、等 短语 四字格 连A中的 基本名图B 441 自、等 短语 五定短语 四字格 连A中的 基本名图B 442 高等 短语 五定短语 五定短语 五定短语 五定短语 五定短语 五定短语 五定短语 五定短语 五定短语							
570 高等 词类 副词 语气副词 幸亏 571 高等 词类 助词 语气副词 续究 572 高等 词类 助词 语气副词 续究 573 高等 词类 助词 要 435 高等 短语 结内类型 基本结构类型 数词+量词+抽象事物 436 高等 短语 四字格 学人平台 要 437 高等 短语 四字格 半A半B 要 A							
571 高等 词类 副词 语气副词 终究 572 高等 词类 助词 语气助词 奏 573 高等 词类 助词 差本结构类型 数词+量词+抽象事物 435 高等 短语 固定短短语 四字格 安A不A 436 高等 短语 固定短短语 四字格 半A半B 437 高等 短语 四字格 半A半B 438 高等 短语 四字格 半A半B 438 高等 短语 四字格 半A非B 440 高等 短语 四字格 2A·忽B 441 高等 短语 西定短短语 四字格 2A·忽B 441 高等 短语 西定短短语 四字格 日本非局 2A·常期 442 高等 短语 西定短语 四字格 日本市 日本带 443 高等 短语 国定短语 其他 巴木得 444 高等 短语 其他 月本 日本市 日本市 日本市 日本市							
572 高等 词类 助词 语气助词 而已 矣 573 高等 词类 助词 语气助词 矣 435 高等 短语 结构类型 数词+量词+抽象事物 436 高等 短语 四字格 数A不A 437 高等 短语 四字格 半A半B 438 高等 短语 四字格 半A半B 438 高等 短语 四字格 半A半B 439 高等 短语 四字格 半A半B 440 高等 短语 四字格 2名A忽B 441 高等 短语 四字格 2名A忽B 441 高等 短语 四字格 2名本带出 442 高等 短语 四字格 2名本带出 443 高等 短语 四字格 时A自B 444 高等 短语 其他 四字格 445 高等 短语 其他 房地 446 高等 短语 其他 投资的 448							
573 高等 词类 助词 每气助词 矣 435 高等 短语 结构类型 基本结构类型 数词+量词+抽象事物 436 高等 短语 固定短语 四字格 爱A不A 437 高等 短语 四字格 光本半B 438 高等 短语 四字格 并A半B 438 高等 短语 四字格 非A非B 439 高等 短语 四字格 2名忽B 440 高等 短语 四字格 2名忽B 441 高等 短语 四字格 2名忽B 441 高等 短语 四字格 2名忽B 442 高等 短语 四字格 连A带B 443 高等 短语 四字格 连A带B 444 高等 短语 四字格 EA带B 444 高等 短语 其他 四不得 445 高等 短语 其他 月根型 分 446 高等 短语 其他 没说的							
436 高等 短语 固定短语 四字格 爱A不A 437 高等 短语 固定短语 四字格 半A半B 438 高等 短语 固定短语 四字格 东A西B 439 高等 短语 固定短语 四字格 京本非B 440 高等 短语 固定短语 四字格 空A忽B 441 高等 短语 固定短语 四字格 连A带B 441 高等 短语 固定短语 四字格 自A自B 442 高等 短语 固定短语 四字格 自A自B 443 高等 短语 固定短语 其他 四字格 自A自B 444 高等 短语 固定短语 其他 以此一条 445 高等 短语 固定短语 其他 以此一条 447 高等 短语 固定短语 其他 投说的 448 高等 短语 固定短语 其他 大论如何 450 高等 短语 其他 与此同时 540 高等 短语 其他 这样一来 541 高等 短语							
436 高等 短语 固定短语 四字格 爱A不A 437 高等 短语 固定短语 四字格 半A半B 438 高等 短语 固定短语 四字格 来A西B 439 高等 短语 固定短语 四字格 求A和B 440 高等 短语 四字格 忽A忽B 441 高等 短语 四字格 连A带B 441 高等 短语 四字格 连A带B 442 高等 短语 四字格 由A时B 443 高等 短语 固定短语 四字格 自A自B 444 高等 短语 固定短语 其他 四字格 自A自B 445 高等 短语 固定短语 其他 別提了 中代 日本							
437 高等 短语 固定短语 四字格 半A半B 438 高等 短语 固定短语 四字格 东A西B 439 高等 短语 固定短语 四字格 非A非B 440 高等 短语 四字格 空A忽B 441 高等 短语 四字格 空A忽A忽B 441 高等 短语 四字格 连A带B 442 高等 短语 四字格 时A时B 443 高等 短语 四字格 自A自B 444 高等 短语 四字格 自A自B 444 高等 短语 其他 四子得 445 高等 短语 四定短语 其他 四子得 446 高等 短语 其他 別提了 所述之外 447 高等 短语 其他 知识 知识 448 高等 短语 其他 投资的 无论如何 450 高等 短语 其他 与此同时 与此同时 539 高等 短语 其他 与此同时 与此同 与此同 与此同 上述一来							
438 高等 短语 固定短语 四字格 东A西B 439 高等 短语 固定短语 四字格 非A非B 440 高等 短语 固定短语 四字格 经A帮B 441 高等 短语 固定短语 四字格 时A时B 442 高等 短语 固定短语 四字格 日A自B 443 高等 短语 固定短语 其他 巴不得 444 高等 短语 其他 別提了 445 高等 短语 其他 別提了 446 高等 短语 其他 別提了 447 高等 短语 其他 別提了 448 高等 短语 其他 以间根到底 449 高等 短语 其他 没说的 450 高等 短语 其他 无论如何 451 高等 短语 其他 与此同时 540 高等 短语 其他 完成 与此同时 541 高等 短语 其他 总的来说/总而言之 542 高等 固定短语 其他 总的来说							
439 高等 短语 固定短语 四字格 非A非B 440 高等 短语 固定短语 四字格 忽A忽B 441 高等 短语 固定短语 四字格 连A带B 442 高等 短语 固定短语 四字格 时A时B 443 高等 短语 固定短语 四字格 日A自B 444 高等 短语 固定短语 其他 巴不得 445 高等 短语 固定短语 其他 所述 以此列底 446 高等 短语 固定短语 其他 只见相到底 可不是 投说的 447 高等 短语 固定短语 其他 无论如何 无论如何 无论如何 无论如何 450 高等 短语 固定短语 其他 与此同时 540 高等 短语 互应短语 其他 这样一来 541 高等 短语 互应短语 其他 总的来说/总而言之 不知好 452 高等 固定短语 其他 总的来说/总而言之 不知好 452 高等 固定短语 其他 总的来说/总而言之 不知好 452 高等 固定短语 其他 总的来说/总面记 26							
440 高等 短语 固定短语 四字格 忽A忽B 441 高等 短语 固定短语 四字格 连A带B 442 高等 短语 固定短语 四字格 时A时B 443 高等 短语 固定短语 四字格 日A自B 444 高等 短语 固定短语 其他 巴不得 445 高等 短语 固定短语 其他 別提了 446 高等 短语 固定短语 其他 四子人 以此之外 日本 447 高等 短语 固定短语 其他 只有 现据 四字格 时未时 日本							
441高等短语固定短语四字格连A带B442高等短语四字格时A时B443高等短语固定短语四字格自A自B444高等短语其他巴不得445高等短语其他别提了446高等短语其他归根到底447高等短语其他可不是448高等短语其他没说的450高等短语其他无论如何451高等短语其他与此同时539高等短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等超语其他总的来说/总而言之452高等固定短语其他总的来说/总而言之452高等固定格式工公							
442高等短语固定短语四字格时A时B443高等短语四字格自A自B444高等短语其他巴不得445高等短语其他别提了446高等短语其他归根到底447高等短语其他可不是448高等短语其他没说的450高等短语其他无论如何451高等短语其他由此可见539高等短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等短语其他总的来说/总而言之452高等固定短语其他总的来说/总而言之452高等固定格式其他总的来说/总而言之							
443高等短语固定短语四字格自A自B444高等短语其他巴不得445高等短语其他别提了446高等短语其他除此之外447高等短语其他归根到底448高等短语其他投说的450高等短语其他无论如何451高等短语其他五次如何451高等短语其他与此同时539高等短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等超语其他总的来说/总而言之452高等固定格式其他总的来说/总而言之							
444 高等 短语 固定短语 其他 巴不得 445 高等 短语 固定短语 其他 别提了 446 高等 短语 固定短语 其他 除此之外 447 高等 短语 固定短语 其他 归根到底 448 高等 短语 固定短语 其他 没说的 449 高等 短语 固定短语 其他 无论如何 450 高等 短语 固定短语 其他 与此同时 539 高等 短语 固定短语 其他 与此同时 540 高等 短语 固定短语 其他 综上所述 541 高等 短语 固定短语 其他 总的来说/总而言之 452 高等 固定格式 其他 总的来说/总而言之 452 高等 固定格式 其他 总的来说/总而言之							
445高等短语固定短语其他别提了446高等短语其他异根到底447高等短语其他归根到底448高等短语固定短语其他没说的450高等短语其他无论如何451高等短语其他由此可见539高等短语其他与此同时540高等短语其他这样一来541高等短语其他这样一来542高等短语其他总的来说/总而言之452高等固定短语其他总的来说/总而言之452高等固定格式工工							
446 高等 短语 其他 除此之外 447 高等 短语 其他 归根到底 448 高等 短语 其他 可不是 449 高等 短语 其他 没说的 450 高等 短语 其他 无论如何 451 高等 短语 其他 由此可见 539 高等 短语 其他 与此同时 540 高等 短语 其他 这样一来 541 高等 短语 其他 综上所述 542 高等 短语 其他 总的来说/总而言之 452 高等 固定格式 其他 总的来说/总而言之							
447高等短语其他归根到底448高等短语其他可不是449高等短语其他没说的450高等短语其他无论如何451高等短语其他由此可见539高等短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等短语其他总的来说/总而言之452高等固定格式工工							
448 高等 短语 其他 可不是 449 高等 短语 其他 没说的 450 高等 短语 其他 无论如何 451 高等 短语 其他 由此可见 539 高等 短语 其他 与此同时 540 高等 短语 其他 这样一来 541 高等 短语 其他 综上所述 542 高等 短语 其他 总的来说/总而言之 452 高等 固定格式 工知好							
449高等短语其他没说的450高等短语其他无论如何451高等短语其他由此可见539高等短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等短语其他总的来说/总而言之452高等固定格式不知好							
450 高等 短语 固定短语 其他 无论如何 451 高等 短语 其他 由此可见 539 高等 短语 固定短语 其他 与此同时 540 高等 短语 固定短语 其他 这样一来 541 高等 短语 固定短语 其他 综上所述 542 高等 短语 固定短语 其他 总的来说/总而言之 452 高等 固定格式 不知好							
451高等短语固定短语其他由此可见539高等短语固定短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等短语其他总的来说/总而言之452高等固定格式不知好							
539高等短语固定短语其他与此同时540高等短语其他这样一来541高等短语其他综上所述542高等短语其他总的来说/总而言之452高等固定格式不知好							
540高等短语固定短语其他这样一来541高等短语其他综上所述542高等短语其他总的来说/总而言之452高等固定格式不知好							
541 高等短语固定短语其他综上所述542 高等短语固定短语其他总的来说/总而言之452 高等固定格式不知好							
542 高等 短语 固定短语 其他 总的来说/总而言之 452 高等 固定格式 不知好							
452 高等 固定格式 不知好							
454 高等 固定格式 无非/不过/只不过/只是而已	罢了						
455 高等 固定格式 以为							
456 高等 固定格式 因而							
457 高等 句子成分 宾语 宾语的语义类型3: (1) 方式宾语 #(2) 工具宾语 #	(3) 材						
料宾语 # (4) 目的宾语 458 高等 句子成分 补语 程度补语3: (1) 形容词/动词+得+不行 # (2) 词/动词+得+要命/要死	形容						
459 高等 句子成分 补语 状态补语: "个"引导的补语							
460 高等 句子的类型 特殊句型 "把"字句5: 表致使(主语+)把+宾语(施事)	+动						
词+了							
461 高等 句子的类型 特殊句型 被动句 (1)被所 #(2) 为所 为所							
462 高等 句子的类型 特殊句型 比较句6: (1) 比起(来) # (2) A+ 词+于+B # (3) A+比+名词+还+	形容 呂词						
463 高等 句子的类型 复句 并列复句 一面,一面							
464 高等 句子的类型 复句 承接复句,此后							
465 高等 句子的类型 复句 承接复句 起初,才							

Grammar Items (HSK advanced level) (continued)						
No.	Level	Item	Type	Detail	Content	
466	高等	句子的类型	复句	递进复句	别说,连也/都# 连也/都,别说#别 说,即使也#即 使也,别说	
467	高等	句子的类型	复句	递进复句	,况且	
468	高等	句子的类型	复句	递进复句	连 更不用说	
469	高等	句子的类型	复句	递进复句	,乃至	
470	高等	句子的类型	复句	递进复句	,且	
471	高等	句子的类型	复句	递进复句	,甚至于	
472	高等	句子的类型	复句	选择复句	或 或	
473	高等	句子的类型	复句	选择复句	宁可/宁愿 也	
474	高等	句子的类型	复句	选择复句	与其,不如	
475	高等	句子的类型	复句	选择复句	与其,宁可/宁愿	
476	高等	句子的类型	复句	转折复句	,而(则)	
477	高等	句子的类型	复句	转折复句	,倒/反倒	
478	高等	句子的类型	复句	假设复句	倘若/若,	
479	高等	句子的类型	复句	假设复句	倘若/假设/假使/若,就/那么	
480	高等	句子的类型	复句	假设复句	幸亏,要不然/不然/要不/否则	
481	高等	句子的类型	复句	条件复句	别管,都	
482	高等	句子的类型	复句	条件复句	任,也	
483	高等	句子的类型	复句	因果复句	(因),故	
484	高等	句子的类型	复句	因果复句	鉴于	
485	高等	句子的类型	复句	因果复句	(由于), 以致	
486	高等	句子的类型	复句	因果复句	, 以至于	
487	高等	句子的类型	复句	因果复句	之所以,是因为/是由于	
488	高等	句子的类型	复句	让步复句	固然,也	
489	高等 高等	句子的类型	复句 复句	让步复句	固然,但是/可是/不过	
490	同守 高等	句子的类型 句子的类型	复印复句	让步复句 让步复句	即便,也 虽说,但是/可是/不过	
491 492	同守 高等	句子的类型	复印复句	让少复印 让步复句	虽远,但定/可定/不过 纵然,也	
492	高等	句子的类型	复句	目的复句	······, ······· ······, 以······	
494	高等	句子的类型	复句	目的复句	,以免/免得	
495	高等	句子的类型	复句	解说复句	, 也就是说	
496	高等	句子的类型	复句	紧缩复句	(要+) 动词+就+动词+个+补语	
497	高等	句子的类型	复句	紧缩复句	动词 (+宾语1) +就+动词 (+宾语1),	
T 21	1111 24,	可工的大王	≪ □	水川久 刊	别	
498	高等	句子的类型	复句	多重复句	三重或三重以上的复句	
526	高等	句子的类型	复句	递进复句	何况	
527	高等	句子的类型	复句	递进复句	,进而	
499	高等	强调的方法			用反问句表示强调 反问句3:何必/何苦呢?	

Table 7. The checklist of grammar items with HSK advanced level (Appendix B).

The prompt for generating instruction data based on grammar items.

[ROLE]

你是一位经验丰富的中文写作教学专家、擅长设计写作型语法教学任务。

[Requirements]

我将会提供给你一个语法项目的结构化描述,请你根据这个语法项目,完成以下任务:

- ①请理解"语法项目"、"类别"、"细目"的可能的教学意图,将语法内容作为一个单独的练习对象,设计成一个用于学生写作练习的教学指令数据。
- ②请你根据不同的HSK等级,设计不同难度的指令任务,旨在指引学生写一篇"XXX",也可以是写一句/一段"XXX"。
- ③每个语法内容都可能包含不同的用法,对于此情况你可以用不同的用法来构造多样的教学指令任务。
- ④由于每个语法项目都需要生成若干条不同的教学指令数据,请你每一次输出新的教学指令数据时,不同与指令池里已有的教学指令重复。指令池已有内容: [Instruction Pool]
- ⑤你生成的指令数据必须以json格式输出。

[Examples]

```
例子1:
```

例子2:

```
下面是一个语法项目的结构化描述:

{
    "HSK等级": "3级",
    "语法项目": "词类",
    "类别": "动词",
    "细目": "离合词(动宾式)",
    "语法内容": "见面"

}

你的输出 (json):

{
    "instruction": "请描述你上一次和朋友见面的经历,要求使用'见面'这个词。",
    "input": "语法点: 离合词(动宾式); 类别: 动词; 词条: 见面",
    "output": "上个周末我和老同学在市中心见面,我们聊了很多高中时的趣事,还一起吃了火锅。"
}
```

```
下面是一个语法项目的结构化描述:
{
  "HSK等级": "3级",
  "语法项目": "句子成分".
  "类别": "补语",
  "细目": "可能补语",
  "语法内容": "可能补语: 动词+得/不+动词/形容词"
}
你的输出 (json):
{
  "instruction": "请写一段关于你尝试学习一项新技能的经历,并在描述中使用"动词+得/不+形容词"这一可
能补语结构,表达你当时是否学得会、做得好。",
  "input": "语法点:可能补语;类别:补语;结构:动词+得/不+动词/形容词",
  "output": "去年我开始学习打羽毛球。一开始我动作做得不标准,总是接不到球。但我每天坚持练习,后
来发球发得越来越好,跑动也变得很灵活。虽然比赛还赢不了,但教练说我学得挺快。"
## [Task Descriptions]
现在,我将会提供给你一个新的语法项目,请你遵循json格式输出1条指令数据,最后的keys只包括 "in-
struction"、"input"、"output"。其中,指令任务可以指导学生写"一篇/一段/一句"某个主题的内容,但是要确保"output"内容的词数要符合这个指令的要求。请注意:你生成新的教学指令数据时,不能与指令池里的已有
指令重复,尽可能变化多样(不同场景,不同指令风格,不同要求)去设计这样的教学指令,所有内容都必
须是中文。
新的语法项目:
[New Grammar Point]
你的输出 (json):
```

Figure 1: The prompt of generating instruction data based on the level-based grammar items (Appendix C).

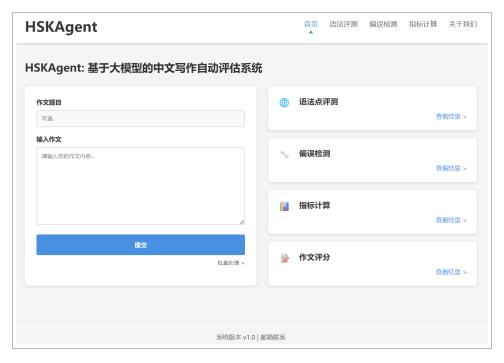


Figure 2: The homepage in the HSKAgent platform (Appendix G).



Figure 3: The function of the assessment of grammar items in the HSKAgent platform (Appendix G).



Figure 4: The function of error detection in the HSKAgent platform (Appendix G).



Figure 5: The functions of metric calculation and holistic scoring in the HSKAgent platform (Appendix G).

References

Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F. L.; Almeida, D.; Altenschmidt, J.; Altman, S.; Anadkat, S.; et al. 2023. Gpt-4 Technical Report. *arXiv* preprint arXiv:2303.08774.

Cui, Y.; Yang, Z.; and Yao, X. 2024. Efficient and Effective Text Encoding for Chinese LLaMA and Alpaca. arXiv:2304.08177.

Jiang, A. Q.; Sablayrolles, A.; Mensch, A.; Bamford, C.; Chaplot, D. S.; de las Casas, D.; Bressand, F.; Lengyel, G.; Lample, G.; Saulnier, L.; Lavaud, L. R.; Lachaux, M.-A.; Stock, P.; Scao, T. L.; Lavril, T.; Wang, T.; Lacroix, T.; and Sayed, W. E. 2023. Mistral 7B. arXiv:2310.06825.

Liu, A.; Feng, B.; Xue, B.; Wang, B.; Wu, B.; Lu, C.; Zhao, C.; Deng, C.; Zhang, C.; Ruan, C.; et al. 2024. Deepseek-v3 Technical Report. *arXiv preprint arXiv:2412.19437*.

Team, G.; Anil, R.; Borgeaud, S.; Alayrac, J.-B.; Yu, J.; Soricut, R.; Schalkwyk, J.; Dai, A. M.; Hauth, A.; Millican, K.; et al. 2023. Gemini: A Family of Highly Capable Multimodal Models. *arXiv preprint arXiv:2312.11805*.

Touvron, H.; Martin, L.; Stone, K.; Albert, P.; Almahairi, A.; Babaei, Y.; Bashlykov, N.; Batra, S.; Bhargava, P.; Bhosale, S.; Bikel, D.; Blecher, L.; Ferrer, C. C.; Chen, M.; Cucurull, G.; Esiobu, D.; Fernandes, J.; Fu, J.; Fu, W.; Fuller, B.; Gao, C.; Goswami, V.; Goyal, N.; Hartshorn, A.; Hosseini, S.; Hou, R.; Inan, H.; Kardas, M.; Kerkez, V.; Khabsa, M.; Kloumann, I.; Korenev, A.; Koura, P. S.; Lachaux, M.-A.; Lavril, T.; Lee, J.; Liskovich, D.; Lu, Y.; Mao, Y.; Martinet, X.; Mihaylov, T.; Mishra, P.; Molybog, I.; Nie, Y.; Poulton, A.; Reizenstein, J.; Rungta, R.; Saladi, K.; Schelten, A.; Silva, R.; Smith, E. M.; Subramanian, R.; Tan, X. E.; Tang, B.; Taylor, R.; Williams, A.; Kuan, J. X.; Xu, P.; Yan, Z.; Zarov, I.; Zhang, Y.; Fan, A.; Kambadur, M.; Narang, S.; Rodriguez, A.; Stojnic, R.; Edunov, S.; and Scialom, T. 2023. Llama 2: Open Foundation and Fine-Tuned Chat Models. arXiv:2307.09288.

Reproducibility Checklist

Unless specified otherwise, please answer "yes" to each question if the relevant information is described either in the paper itself or in a technical appendix with an explicit reference from the main paper. If you wish to explain an answer further, please do so in a section titled "Reproducibility Checklist" at the end of the technical appendix.

This paper:

Includes a conceptual outline and/or pseudocode description of AI methods introduced (**partial**)

Clearly delineates statements that are opinions, hypothesis, and speculation from objective facts and results (yes)

Provides well marked pedagogical references for less-familiare readers to gain background necessary to replicate the paper (yes)

Does this paper make theoretical contributions? (yes) If yes, please complete the list below.

All assumptions and restrictions are stated clearly and formally. (yes)

All novel claims are stated formally (e.g., in theorem statements). (partial)

Proofs of all novel claims are included. (no)

Proof sketches or intuitions are given for complex and/or novel results. (yes)

Appropriate citations to theoretical tools used are given. (ves)

All theoretical claims are demonstrated empirically to hold. (ves)

All experimental code used to eliminate or disprove claims is included. (yes)

Does this paper rely on one or more datasets? (yes)

If yes, please complete the list below.

A motivation is given for why the experiments are conducted on the selected datasets (**yes**)

All novel datasets introduced in this paper are included in a data appendix. (yes)

All novel datasets introduced in this paper will be made publicly available upon publication of the paper with a license that allows free usage for research purposes. (yes)

All datasets drawn from the existing literature (potentially including authors' own previously published work) are accompanied by appropriate citations. (yes)

All datasets drawn from the existing literature (potentially including authors' own previously published work) are publicly available. (**partial**)

All datasets that are not publicly available are described in detail, with explanation why publicly available alternatives are not scientifically satisficing. (**yes**)

Does this paper include computational experiments? (**yes**) If yes, please complete the list below.

This paper states the number and range of values tried per (hyper-) parameter during development of the paper, along with the criterion used for selecting the final parameter setting. (partial)

Any code required for pre-processing data is included in the appendix. (no).

All source code required for conducting and analyzing the experiments is included in a code appendix. (**no**)

All source code required for conducting and analyzing the experiments will be made publicly available upon publication of the paper with a license that allows free usage for research purposes. (yes)

All source code implementing new methods have comments detailing the implementation, with references to the paper where each step comes from (partial)

If an algorithm depends on randomness, then the method used for setting seeds is described in a way sufficient to allow replication of results. (**no**)

This paper specifies the computing infrastructure used for running experiments (hardware and software), including GPU/CPU models; amount of memory; operating system; names and versions of relevant software libraries and frameworks. (yes)

This paper formally describes evaluation metrics used and explains the motivation for choosing these metrics. (yes)

This paper states the number of algorithm runs used to compute each reported result. (no)

Analysis of experiments goes beyond single-dimensional summaries of performance (e.g., average; median) to include measures of variation, confidence, or other distributional information. (no)

The significance of any improvement or decrease in performance is judged using appropriate statistical tests (e.g., Wilcoxon signed-rank). (no)

This paper lists all final (hyper-)parameters used for each

model/algorithm in the paper's experiments. (partial).