BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

Chinese Undergraduate Students' Acquisition of Verb Classes and Their Features: Using the Rasch Model

Lili Feng , ***Heechul Lee*** *, ***Jaewoo Shim*** *

Department of English Education, Jeonbuk National University, Jeonju 54896, Republic of Korea

ABSTRACT

The development of the second language (L2) learners' competence in verb classes and their features is one of the most challenging aspects of language acquisition. It is strongly influenced by instruction, L1 lexical aspects, and L2 learners' semantic cognition. This study examined the second language acquisition difficulty of the semantic aspect of six verb classes and four pairs of features within the role and reference grammar (RRG) framework based on L2 English learners' perceptions. A total of 265 Chinese undergraduate students majoring in English participated in this study and responded to a 30-item test that required them to determine the acceptability of sentences concerning verb classes and their features. Utilizing the Rasch model, this study established an L2 English difficulty hierarchy based on lexical aspect acquisition, that is, ranging from very easy to very difficult along a continuum. The findings revealed that participants were proficient at verb classes with the features of [+dynamic] and [−punctual] and had difficulty with those with the properties of [−dynamic] and [+punctual]. This study enriches our understanding of cross-linguistic influences and the complex structures of semantic cognition in the L2 acquisition difficulty of verb classes and their semantic features. The implications for English semantic instruction are discussed; in particular, the semantic meanings of six verb classes and their features should be taught through consciousness-raising (CR) tasks as part of grammar instruction.

Keywords: Verb Class; Verb Feature; Difficulty Hierarchy; English Language Teaching; English Language Acquisition

*CORRESPONDING AUTHOR:

Heechul Lee, Department of English Education, Jeonbuk National University, Jeonju 54896, Republic of Korea; Jaewoo Shim, Department of English Education, Jeonbuk National University, Jeonju 54896, Republic of Korea; hclee@jbnu.ac.kr (H.L.); shimjw@jbnu.ac.kr (J.S.)

ARTICLE INFO

Received: 10 February 2025 | Revised: 21 March 2025 | Accepted: 26 March 2025 | Published Online: 12 April 2025
DOI: https://doi.org/10.30564/fls.v7i4.8705

CITATION

Feng, L., Lee, H., Shim, J., 2025. Chinese Undergraduate Students' Acquisition of Verb Classes and Their Features: Using the Rasch Model. Forum for Linguistic Studies. 7(4): 660–674. DOI: https://doi.org/10.30564/fls.v7i4.8705

1. Introduction

The acquisition of the aspectual classification of verbs and their semantic features has long been regarded as a central component in the acquisition of a second language (L2)[1–4]. Various aspects of L2 lexical aspect development have been extensively investigated[3]. According to Shirai and Nishi[1] and Sun and Rodríguez[4], the acquisition of the L2 lexical aspect is influenced by cross-linguistic differences. Empirical studies have shown that L2 semantic properties impact L2 learners' understanding of semantic information[4]. However, while the L2 acquisition difficulty of these semantic categories has been emphasized, little attention has been paid to the empirical research in this field.

Van Valin's[5] study on verb classes and their properties, developed from Vendler[6] and Smith[7], is a comprehensive analysis of semantic aspect in the role and reference grammar (RRG) system that is used to explore six categories (states (i.e., affairs of state which remain constant), activities (i.e., involving action and the duration of time but no endpoint), achievements (i.e., to occur in a very short time and produce a new result state), accomplishments (i.e., involving the duration of time and a time limit, and producing a new result state), semelfactives (i.e., to occur very briefly and without any result), and active accomplishments (i.e., involving action, duration of time and a terminal point) and four pairs of lexical properties (static (i.e., to distinguish static situation from an event)/non-static, dynamic (i.e., related to the degree of force, energy, and intensity; involving action, which occurs in stages with continuous possibility of a new energy input[7])/non-dynamic, telic (i.e., involving a terminal point)/atelic, and punctual (i.e., to occur in a very short time; events which lack temporal duration)/non-punctual (i.e., involving temporal duration) in English.

Nonetheless, most of the empirical studies have focused on Vendler's four aspect categories of states, activities, achievements, and accomplishments and their semantic features of stativity, telicity, and punctuality, leaving Van Valin's classification of six verb classes (i.e., four Vendlerian and two non-Vendlerian verb classes) and their semantic properties understudied[1, 2, 5, 6]. The relevant studies have suggested that six verb classes and four pairs of properties presented in Van Valin[5] are not universal, particularly compared with Vendler's original four verb classes and their inherent features, highlighting the necessity of exploring learners' acquisition of Van Valin's semantic study of the lexical aspect[8–10]. While some studies on Van Valin's verb classes have examined learners' acquisition and use of their native language[9, 10], there has been a dearth of research on the acquisition of L2 within Van Valin's concept of verb classes in the RRG system. Accordingly, the extension of studies on the acquisition of L2 in the framework of Van Valin's lexical aspect will not only expand the research scope of the semantic aspect but will also contribute to a more comprehensive understanding of L2 acquisition.

Given the scarcity of research in English L2 acquisition and the learning difficulty of Van Valin's[5] aspectual verb classes by L2 learners, this study aims to fill this gap by using the Rasch model to identify a difficulty hierarchy of six verb classes and their semantic properties based on Chinese English-major undergraduate students' collective perceptions of Van Valin's[5] lexical representation. Specifically, the present study seeks to address the following three research questions:

1. Do Chinese university undergraduate learners encounter varying levels of difficulty regarding the semantic aspects of the six verb classes and their semantic features?
2. If so, what is the observed distribution pattern of these verb classes and their features in the experimental task?
3. How can the distribution pattern of verb classes and their features be explicitly taught to facilitate their implicit acquisition?

2. Literature Review

2.1. Van Valin's Verb Classes and Semantic Features

Based on previous research on the lexical aspect from Vendler[6], Dowty[11], Smith[7], and Van Valin and LaPolla[12], Van Valin[5] focused on six verb classes: states, activities, achievements, accomplishments, semelfactives, and active accomplishments and four pairs of verb features: static versus non-static, dynamic versus non-dynamic, telic versus atelic, and punctual versus non-punctual. Van Valin[5] emphasized that the fundamental difference between verbs of states and verbs of non-static verbs is the answer to "What happened?" or "What is happening?" Non-static verbs can be used as answers to this question because they involve the change of state. In examples (2), (3), (4), (5), and (6),

these five sentences (e.g., "run", "run to her school", "pop", "freeze", and "cough") can be used as an answer to this question as they all describe specific actions taking place. In contrast, in (1), Lily understands the fact and remains constant because it does not show any action taking place.

1. State
   Lily knows the fact.
   Li-li zhi dao zhe ge shi shi
   莉莉知道这个事实。
2. Activity
   Emily ran in the playground.
   Ai-mi-li zai cao chang shang pao bu
   艾米丽在操场上跑步。
3. Active accomplishment
   Emily ran to her school.
   Ai-mi-li pao dao ta de xue xiao
   艾米丽跑到他的学校。
4. Achievement
   The bubble popped.
   Pao mo po lie le
   泡沫破裂了。
5. Accomplishment
   The water froze.
   Shui jie bing le
   水结冰了。
6. Semelfactive
   Emily coughed.
   Ai-mi-li ke sou le
   艾米丽咳嗽了。

In turn, activities and active accomplishments are dynamic and compatible with adverbs such as "vigorously" and "actively." In examples (2) and (3), "run" and "run to her school" are dynamic. They are related to the degree of energy, and occur in stages with the continuous potential for new energy input. By contrast, in events (4) and (5), "pop" and "freeze" do not imply internal dynamicity.

The difference between achievements and accomplishments involves the duration of time. In (4), the bubble shows the action of popping in a very short time, but in (5), the water needs more time to turn into ice (a result state). Similarly, in (2) and (3), "run" and "run to school" have internal duration. In contrast, in (6), "cough" doesn't have internal duration because "cough," as a semelfactive verb, occurs very briefly and without any result in Smith[7]. Despite the difference in temporal duration, achievements and accomplishments share the same property of having an inherent endpoint. Examples (4) and (5) both have a time limit (+telic) and produce a new result state. Similarly, "run to school" has a terminal point in (2), but "run" and "cough" denote events without terminal points in (2) and (6).

According to Van Valin[5], there are six verb classes, and they carry four pairs of properties: *states* (static, non-dynamic, atelic, non-punctual), *activities* (non-static, dynamic, atelic, non-punctual), *achievements* (non-static, non-dynamic, telic, punctual), *accomplishments* (non-static, non-dynamic, telic, non-punctual), *semelfactive* (non-static, dynamic/non-dynamic, atelic, punctual), and *active accomplishment* (non-static, dynamic, telic, non-punctual) (see **Table 1**). Building on Vendler's[6] framework, Van Valin[5] added semelfactives and active accomplishments to the RRG theory and their properties were previously classified as *static versus dynamic* to include *static versus non-static*, and *dynamic versus non-dynamic*, considering complex situations.

Utilizing Vendler's[6] concept of four verb classes, other studies have analyzed the differences in the aspectual structure of verb classes and the semantic features between L2 and native language L1[1, 4]. These studies have also been applied to Mandarin Chinese. However, unlike in previous research, Xiao and McEnery[13] and Chu[14] incorporated the "resultative" property proposed by Dowty[11] and Smith[7, 15] to further analyze the lexical aspect of Mandarin Chinese (see **Table 1**). Current research on semantic knowledge related to verb classes in Van Valin[5] is largely based on the perspective of L1 English speakers. L2 learners may feel confused by the inherent meanings of the six verb classes and their properties. There is a notable lack of research on L2 acquisition in this field, particularly concerning Chinese undergraduate learners.

2.2. Aspectual Structures of Verb Classes Between English and Chinese

The aspectual element in Van Valin's[5] verb classes in both English and Mandarin Chinese is largely adapted to Vendler's[6] basic classification of verbs. In terms of the differences between non-static verbs (*activities, achievements, accomplishments, semelfactives,* and *active accomplishments*), the dynamicity involves action. Specifically, it is influenced by the continuous fluctuation of intensity

Table 1. Verb classes in English and Mandarin Chinese.

	English Verb Classes in Van Valin (2005)		Mandarin Chinese	
	Property	**Example**	**Property**	**Example**
States	+static	be tall	−dynamic(+static)	have
	−dynamic	love	+durative(−punctual)	be happy
	−telic	know	−telic	know
	−punctual	believe	−resultative	like
Activities	−static	march	+dynamic(−static)	run
	+dynamic	walk	+durative(−punctual)	say
	−telic	swim	−telic	
	−punctual	write	−resultative	eat
Accomplishments	−static	melt	+dynamic(−static)	build a house
	−dynamic	freeze	+durative(−punctual)	go home
	+telic	dry	+telic	
	−punctual	learn	+resultative	
Achievements	−static	pop	+dynamic(-static)	break
	−dynamic	explode	−durative(+punctual)	fall asleep
	+telic	shatter	+telic	
	+punctual		+resultative	
Semelfactives	−static	flash	+dynamic(−static)	kick
	+dynamic and −dynamic	cough [+dynamic] glimpse [−dynamic]	−durative(+punctual)	hit
	-telic	tap	−telic	put
	+punctual		−resultative	
Active accomplishment	−static	eat the fish		
	+dynamic	walk to the park		
	+telic	paint Mary's portrait		
	−punctual	run to the park		

with new force input. For example, "dance" (activity) and "walk to her campus" (active accomplishment) are dynamic—they involve action and can co-occur with adverbs like actively, energetically and so on[5]. In contrast, "melt" (accomplishment) and "pop" (achievement) are not dynamic. In terms of punctuality, Van Valin proposed that adverbs, such as "quickly," "rapidly," and "slowly," focus on distinguishing punctuality—for example, "freeze" (accomplishment), "walk" (activity), and "run to the park" (active accomplishment) are non-punctual and they can co-occur with such pace adverbs, which imply internal duration (e.g., quickly, slowly, and so on)[5], but "pop" (achievement) and "flash" (semelfactive) are inherently punctual. In addition, two other verb classes in English—active accomplishments and semelfactives—do not correspond directly with Vendler's[6] classification. Active accomplishments are telic and non-punctual, whereas semelfactives are atelic and punctual.

According to Xiao and McEnery[13], Chu[14], and others, Mandarin Chinese generally corresponds with Vendler's[6] verb classes and Smith's[7] semelfactives. Regarding the distinction between *states* and non-static verbs (*activities, achievements, accomplishments,* and *semelfactives*), Smith[7] highlighted dynamicity as a notable feature. The durative feature further distinguishes two pairs of events (i.e., one pair of activities and accomplishments and the other pair of achievements and semelfactives)[13]. The term *durative* refers to events that involve a certain temporal length. For instance, verbs such as "run" and "build a house" contain durative feature[13]. It should be noted that studies on Chinese lexical aspect typically employ the terms *durative versus punctual*[13, 14], whereas the present study adopts *punctual versus non-punctual,* following Van Valin's framework[5]. The property of involving an endpoint further differentiates achievements from semelfactives. An additional feature, resultativity, further separates semelfactive verbs from achievement verbs: semelfactives (such as "hit") do not cause both a change and an explicit result, but achievements (such as "break") cause a change and yield an explicit result.

In terms of the comparison of verb classes and the semantic features from Van Valin[5] in English with those in

Mandarin Chinese (see **Table 1**), Van Valin[5] highlighted the static feature as a key factor in distinguishing states and events, rather than dynamicity, in Mandarin Chinese. Correspondingly, dynamicity in Van Valin's[5] framework further separates non-static verb classes based on the classification of the continuous possibility of new energy input and the presence of different degrees of intensity. Moreover, active accomplishments in Van Valin's[5] framework are distinguished from accomplishments and activities. For instance, they are dynamic and telic, such as "run to school" (active accomplishment), while "run" (activity) is dynamic and atelic, and "melt" (accomplishment) is telic but not dynamic. From the perspective of semantic structure, verb classes (*states*, *activities, achievements, accomplishments,* and *semelfactives*) coincide in both languages in their non-punctual and telic properties. However, there are differences in the tests for the punctual/non-punctual property of verb classes between Van Valin's[5] test for *Aktionsart* (i.e., originally proposed by Vendler[6]; representing action in German and involving temporal characteristics) and Mandarin Chinese[13]. Van Valin[5] proposed adverbs such as "quickly" in English, while Xiao and McEnery[13] proposed "zhe" (i.e., a sign of the ongoing progression) in Mandarin Chinese.

Given these similarities and differences between English and Mandarin Chinese in the aspectual structures of verb classes and their semantic properties, the question arises as to whether L2 learners encounter difficulties in the acquisition of the lexical aspect in the target language due to positive and negative cross-linguistic influences, which can facilitate and affect their L2 learning process.

2.3. L2 Acquisition Difficulty of Verb Classes and Semantic Features

Many previous studies have demonstrated the difficulty in L2 acquisition and the use of the aspectual structure of verb classes[1, 16, 17]. For example, the work of Shirai and Nishi[1] focused on the acquisition of lexicalization of aspectual structure in English and Japanese. That study revealed cross-linguistic influences by employing case studies to analyze different lexical expressions in English and Japanese in terms of the concept of lexical aspect. The authors indicated that L2 learners encounter more challenges in acquiring the inherent aspect when there is imperfect correspondence between these languages. Other studies, such as those by Jiang[16, 17], have affirmed that L1 transfer from Mandarin Chinese and Korean negatively influences the L2 acquisition of English lexical representation.

In contrast, previous studies in L2 acquisition, including the data analyzed by Sun[18], have shown that dynamicity is favored by Chinese learners over other properties, such as durativity and telicity. Other studies, such as Smith[7] and Xiao and McEnery[13], also observed traces of this phenomenon. Verb classes (except for states) in the Mandarin Chinese lexical aspect are dynamic because they involve change over time, including the degree of force or the changing endpoint. Gennari and Poeppel[19] collected data from 52 participants and found that cognition and semantic meaning were cognitively intertwined. These students took longer to process eventive verbs mentally because eventive verbs have a more complex internal structure than states.

Although Van Valin's verb classes focused on L1 English speakers, this does not mean that L2 learners cannot directly and intuitively acquire Van Valin's lexical representation. Indeed, Van Valin's test for *Aktionsart* is at the heart of his RRG system (see **Table 2**). Sayyad et al.[10] employed Van Valin and LaPolla's[12] verb class tests to determine static structure in Arabic. They found that four out of five of the tests were applicable in Arabic, including the application of adverbs related to force, speed, and time. Other studies, such as that by Phillips and Thiengburanathum[9], have supported the use of Van Valin's tests[5] within the RRG framework to distinguish verb classes in Thai.

These studies indicate that cross-linguistic differences and the complex internal structures of the verb classes influence L2 learners' acquisition. However, little research has focused on L2 acquisition within the framework of Van Valin's verb classes and semantic properties, particularly using Rasch model analysis, which is a model of statistical analysis, rather than linguistic analysis. Therefore, the present study adopted Van Valin's[5] tests for *Aktionsart* and analyzed the use and acquisition of English verbs of different aspectual classes by Chinese English-major learners using the Rasch model. These Rasch model analyses allowed us to capture the distribution of verb classes and their features as perceived by Chinese learners.

2.4. Pedagogy on Semantic Meaning

Semantics is widely acknowledged as an important component of grammar, along with syntax, morphology, pragmatics, phonology, and vocabulary[20]. Indeed, semantic instruc-

tion is commonly applied in grammar teaching. Moreover, Larsen-Freeman[21] challenged the traditional grammar pedagogy and emphasized that three dimensions of form, meaning, and use should be integrated, following a meaning-based or communicative teaching approach. Larsen-Freeman[21] illustrated this with the example of the verb phrase. In terms of form, verb phrases can consist of two parts (verb + particle), such as "to look up", or three parts (verb + particle + preposition), such as "to keep up with." They can also be transitive phrases ("to look at") or intransitive phrases ("to come back"). Regarding meaning, verb phrases can be literal ("to hang up") or figurative ("to run into", interpreted as "to meet by chance"). In terms of use, single verbs are preferred in formal situations ("extinguish the fire"), whereas verb phrases such as "put out the fire" are more common in daily communication. According to her study[21], focusing on form within a semantic context in communication is a better way to teach grammar than teaching form in isolation.

In L2 acquisition, task-based instruction (TBI) is an effective approach to providing learners with opportunities to perceive semantic differences in pre-task or post-task phases, rather than simply presenting rules or forms to L2 learners and then guiding them to practice[22]. Dalpanagioti[23] examined how semantic frameworks are integrated into TBI in further exploring the relationship between cognition and semantics. Shim and Lee[22] collected data from 54 L2 English instructors and proposed consciousness-raising (CR) tasks on semantic meaning to help learners perceive cross-linguistic influence in semantic instruction. Woll and Paquet[24] gathered data from 47 Spanish and French classes and investigated the effectiveness of CR tasks in cross-linguistic research. Their findings suggested that multilingual CR tasks can help learners establish connections between their existing linguistic knowledge and the target language, thereby enhancing language acquisition. However, the literature on teaching L2 English verb classes through semantic meaning—particularly within Van Valin's RRG framework—remains largely unexplored, both theoretically and empirically. Moreover, English semantics are insufficiently addressed at the undergraduate stage and are typically integrated into English linguistics courses in China[25].

3. Methods

3.1. Participants

The present study employed the Rasch model method using English-major undergraduates from a university located in southern China as research participants. The undergraduate students ranged from Year 1 to Year 3 and were being trained in English listening, speaking, reading, and writing skills. Data on verb classes and their characteristics were collected through questionnaires during a 30-minute on-site examination held in regular campus classrooms. A total of 298 participants, aged 18–20 years, were recruited, of whom 33 were excluded due to invalid responses (i.e., questionnaires with the same answers for all items or answers with a repetition rate of at least 90%). Finally, 265 valid responses remained, with an effective rate of 88.93%.

3.2. Instrument

To investigate Chinese EFL (English as a foreign language) students' acquisition of the verb classes (states, activities, achievements, accomplishments, active accomplishments, and semelfactives) and their features (±static, ±dynamic, ±punctual, ±telic), a specialized test for verb classes revised by Van Valin (2005) was adopted. A questionnaire consisting of states (five items), activities (four items), achievements (four items), accomplishments (five items), active accomplishments (seven items), and semelfactives (five items) was designed to measure the students' understanding and recognition of verb classes and their corresponding features (see **Appendix A**). To ensure the validity and authenticity of the test, the participants were required to complete a 30-item test within 30 minutes held in their classrooms during regular class time. Below are some examples, with √ indicating acceptable answers and X indicating unacceptable answers. After all of the tests were finished, we applied Van Valin's tests for *Aktionsart* to check the L2 learners' answers, as illustrated in **Table 2**.

1. We are having only two realistic choices. (X)
2. A small blue square flashed near the cross quickly. (X)
3. They are running quickly. (√)
4. Rome collapsed actively. (X)
5. Jim coughed once violently. (√)
6. James painted Mary's portrait in one hour. (√)

Table 2. Van Valin's[5] tests for Aktionsart.

Criterion	State	Achiev	Seml	Accomp	Activity	Active Accomp
1. Occurs with Progressive	No*	No*	No*	Yes	Yes	Yes
2. Occurs with adverbs like vigorously,	No	No	Some*	No	Yes	Yes
3. Occurs with adverbs like quickly, slowly, etc.	No	No*	No*	Yes	Yes	Yes
4. Occurs with X for an hour, spend an hour Xing	Yes*	No*	Yes*	Irrelevant*	Yes	Irrelevant*
5. Occurs with X in an hour	No	No*	No*	Yes	No	Yes

3.3. Data Collection and Analysis

The questionnaire instructions were written in Chinese to ensure full comprehension among the respondents. After the data collection, the checked answers of 30 items were scored as 1 for correctness and 0 for incorrectness using SPSS Version 27.0 (Armonk, NY: IBM Corp), with female participants represented by 0 and male participants represented by 1. Next, the data from SPSS were transferred to WINSTEPS Version 3.68.2 for a series of analyses, including person and item reliability, item statistics, a unidimensionality test, and an item map. WINSTEPS Version 3.68.2 efficiently facilitated the analysis of the summary statistics of measured items, the difficulty hierarchy of items, the measurement of items, and the unidimensionality test[26–28].

The Rasch model was selected as the primary means of analysis for several reasons. First, in contrast with traditional test analyses, the estimated item difficulty in the Rasch model analysis is stable, independent of the respondents' ability, because the measures of item difficulty are independent of the sample[29]. Second, using the Rasch model analysis, reliability coefficients enable researchers to check the degree of errors of item measures, and model fit statistics help researchers to evaluate the validity. That is, the collected data conform to the model expectations. The estimation of the degree of model fit is expressed quantitatively by the infit mean square (MNSQ), outfit MNSQ, infit Z-standard (ZSTD), and outfit ZSTD. Third, the Rasch model is particularly useful because items in the present study involve different features based on L2 learners' perceptions[30].

To answer the first research question, L2 learners' acquisition difficulty logit values of verb classes were estimated by the Rasch model analysis, including item reliability coefficients, item statistics, a unidimensionality test, and the item map. Next, to answer the second research question, the distribution of English verb classes and the semantic features was estimated, discussed, and analyzed. To answer the third question, based on the finding of the difficulty hierarchy of verb classes and their properties, compared with L1 Mandarin Chinese and combined with complex cognitive English semantics, the practical implications of semantic instruction were provided.

4. Results

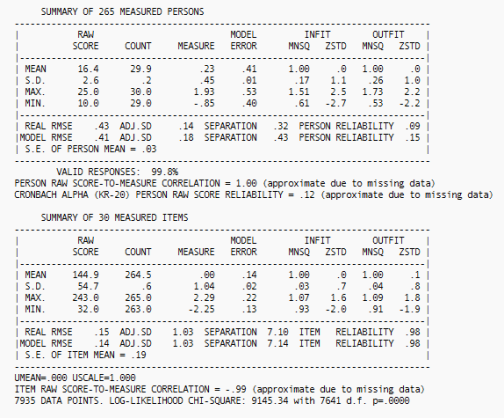
We conducted item reliability, item statistics, unidimensionality test, and item map to answer the first question: Do L2 participants perceive the item difficulty hierarchy of six verb classes and their semantic properties?

4.1. Person Reliability and Item Reliability

Figure 1 presents the summary statistics of 265 individuals and 30 measured items. The mean person measure was 0.23 logits, indicating the average ability of all students to answer the 30 items. In other words, a person measure of 0.23 logits greater than .00 logits shows that the average students' abilities were slightly higher than the difficulty level of the questions. According to Nuryanti et al.[31], a standard deviation of 1.04 in the summary of 30 measured items demonstrates that students tend to have the ability to answer the questions correctly. As Sumintono and Widhiarso[32] observed, the standard values for person and item reliability are as follows: A value > 0.94 is considered special; 0.91–0.94 is excellent; 0.81–0.90 is good; 0.67–0.80 is sufficient; and <0.67 indicates low reliability.

The person reliability of 0.09 and the item reliability of 0.98 are contrasted in **Figure 1**. However, as the primary focus of the study was on the hierarchical distribution of item difficulty across verb classes and their features, greater emphasis is placed on item reliability. Notably, **Figure 1** demonstrates that the item reliability measure, at 0.98, was significantly higher than the person reliability measure. This indicates varying levels of difficulty among the test items,

with a separation index of approximately 7.10. Consequently, the quality of these 30 items is considered to be quite high.



SUMMARY OF 265 MEASURED PERSONS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	16.4	29.9	.23	.41	1.00	.0	1.00	.0
S.D.	2.6	.2	.45	.01	.17	1.1	.26	1.0
MAX.	25.0	30.0	1.93	.53	1.51	2.5	1.73	2.2
MIN.	10.0	29.0	-.85	.40	.61	-2.7	.53	-2.2

REAL RMSE	.43	ADJ.SD	.14	SEPARATION	.32	PERSON RELIABILITY	.09	
MODEL RMSE	.41	ADJ.SD	.18	SEPARATION	.43	PERSON RELIABILITY	.15	
S.E. OF PERSON MEAN = .03								

VALID RESPONSES: 99.8%

PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00 (approximate due to missing data)

CRONBACH ALPHA (KR-20) PERSON RAW SCORE RELIABILITY = .12 (approximate due to missing data)

SUMMARY OF 30 MEASURED ITEMS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	144.9	264.5	.00	.14	1.00	.0	1.00	.1
S.D.	54.7	.6	1.04	.02	.03	.7	.04	.8
MAX.	243.0	265.0	2.29	.22	1.07	1.6	1.09	1.8
MIN.	32.0	263.0	-2.25	.13	.93	-2.0	.91	-1.9

REAL RMSE	.15	ADJ.SD	1.03	SEPARATION	7.10	ITEM RELIABILITY	.98	
MODEL RMSE	.14	ADJ.SD	1.03	SEPARATION	7.14	ITEM RELIABILITY	.98	
S.E. OF ITEM MEAN = .19								

UMEAN=.000 USCALE=1.000

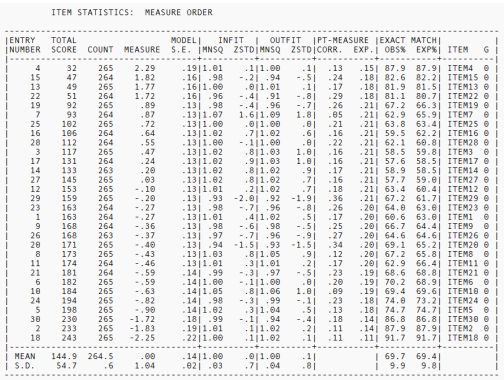
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.99 (approximate due to missing data)

7935 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 9145.34 with 7641 d.f. p=.0000

Figure 1. Person and item reliability.

4.2. Rasch Analysis of Item Statistics

Figure 2 presents the Rasch measure of item statistics with infit and outfit. There were 30 items, and the column order of the measure aligns with the logit of the person–item map, which had a standard deviation of 1.04 and a mean item measure of 0.00. The range varied from the most difficult item, at 2.29, to the easiest item, at −2.25, allowing for clear detection of the difficulty hierarchy. The important columns that follow include the standard error, infit MNSQ and infit ZSTD values, outfit MNSQ and ZSTD values, and the real and anticipated correlations for each item's outcome. According to Boone et al.[33], it is important to consider the following criteria for outfit values: MNSQ outfit values between 0.5 and +1.5 and ZSTD outfit values falling between −2 and +2. Additionally, if the value of the point–measure correlation (PTME) is positive (+), the item is considered to effectively assess the intended construct[30].



ITEM STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
4	32	265	2.29	.19	1.01	.1	1.00	.1	.13	.15	87.9	87.9	ITEM4	0
15	47	264	1.82	.16	.98	-.2	.94	-.5	.24	.18	82.6	82.2	ITEM15	0
13	49	265	1.77	.16	1.00	.0	1.01	.1	.17	.18	81.9	81.5	ITEM13	0
22	51	264	1.72	.16	.96	-.4	.91	-.8	.29	.18	81.1	80.7	ITEM22	0
19	92	265	.89	.13	.98	-.4	.96	-.7	.26	.21	67.2	66.3	ITEM19	0
7	93	264	.87	.13	1.07	1.6	1.09	1.8	.05	.21	62.9	65.9	ITEM7	0
25	102	265	.72	.13	1.00	.0	1.00	.0	.21	.21	63.8	63.4	ITEM25	0
16	106	264	.64	.13	1.02	.7	1.02	.6	.16	.21	59.5	62.2	ITEM16	0
28	112	264	.55	.13	1.00	-.1	1.00	.0	.22	.21	62.1	60.8	ITEM28	0
3	117	265	.47	.13	1.02	.8	1.03	1.0	.16	.21	58.5	59.8	ITEM3	0
17	131	264	.24	.13	1.02	.9	1.03	1.0	.16	.21	57.6	58.5	ITEM17	0
14	133	263	.20	.13	1.02	.9	1.02	.9	.17	.21	58.9	58.5	ITEM14	0
27	145	265	.03	.13	1.02	.8	1.02	.7	.16	.21	57.7	59.0	ITEM27	0
12	153	265	-.10	.13	1.01	.2	1.02	.7	.18	.21	63.4	60.4	ITEM12	0
29	159	265	-.20	.13	.93	-2.0	.92	-1.9	.36	.21	69.2	61.7	ITEM29	0
23	163	264	-.27	.13	.98	-.7	.96	-.8	.26	.20	64.0	63.0	ITEM23	0
1	163	264	-.27	.13	1.01	.4	1.02	.5	.17	.20	60.6	63.0	ITEM1	0
9	168	264	-.36	.13	.98	-.6	.98	-.5	.25	.20	66.7	64.4	ITEM9	0
26	168	263	-.37	.13	.97	-.7	.96	-.9	.27	.20	64.6	64.6	ITEM26	0
20	171	265	-.40	.13	.94	-1.5	.93	-1.5	.34	.20	69.1	65.2	ITEM20	0
8	173	265	-.43	.13	1.03	.8	1.05	.9	.12	.20	67.2	65.8	ITEM8	0
11	174	264	-.46	.13	1.01	.3	1.01	.2	.17	.20	62.9	66.4	ITEM11	0
21	181	264	-.59	.14	.99	-.3	.97	-.5	.23	.19	68.6	68.0	ITEM21	0
6	182	265	-.59	.14	1.00	.1	1.00	.2	.20	.19	70.2	68.0	ITEM6	0
10	184	265	-.63	.14	1.05	.8	1.06	1.0	.09	.19	69.4	69.6	ITEM10	0
24	194	265	-.82	.14	.98	-.3	.99	-.1	.23	.18	74.0	73.2	ITEM24	0
5	198	265	-.90	.14	1.02	.3	1.04	.5	.13	.18	74.7	74.7	ITEM5	0
30	230	265	-1.72	.18	.99	-.1	.94	-.4	.18	.14	86.8	86.8	ITEM30	0
2	233	265	-1.83	.19	1.01	.1	1.02	.2	.11	.14	87.9	87.9	ITEM2	0
18	243	265	-2.25	.22	1.00	.1	1.02	.1	.11	.11	91.7	91.7	ITEM18	0
MEAN	144.9	264.5	.00	.14	1.00	.0	1.00	.1			69.7	69.4		
S.D.	54.7	.6	1.04	.02	.03	.7	.04	.8			9.0	9.0		

Figure 2. Item statistics.

Figure 2 indicates that the MNSQ outfit values ranged from 0.91 to 1.09 and that the ZSTD outfit values ranged from −1.9 to +1.8. Therefore, all items met the established criteria of outfit items as per Boone et al.[33]. Furthermore, the PTME column did not contain any negative values, and a positive index indicates a consistent positive correlation and support for the internal coherence of the 30 items. Consequently, the hierarchical structure of these 30 items was statistically coherent and robust, confirming the fitness of all items.

4.3. Unidimensionality Test

Figure 3 presents an analysis of item unidimensionality, which assesses the construct validity of the items by examining the evolution of multiple dimensions through both explained and unexplained variances. According to Widyaningsih and Yusuf[34], the standard for interpreting raw variance values explained by measures is as follows: Values between 20% and 40% are good, and the unexplained variance should be less than 15%. The analysis in **Figure 3** indicates that the raw variance explained by the measure was 20.5%, which is considered good. The unexplained variances for the first through fifth measures were 7.5%, 5.2%, 4.7%, 4.3%, and 3.9%, respectively, all of which were below 15%. Therefore, based on these results and using the Rasch model through WINSTEPS Version 3.68.2, the validity criteria were well met in terms of explained and unexplained variances.

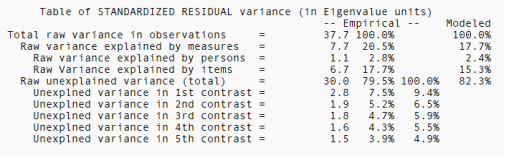


Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --		Modeled
Total raw variance in observations	=	37.7	100.0%	100.0%
Raw variance explained by measures	=	7.7	20.5%	17.7%
Raw variance explained by persons	=	1.1	2.8%	2.4%
Raw Variance explained by items	=	6.7	17.7%	15.3%
Raw unexplained variance (total)	=	30.0	79.5% 100.0%	82.3%
Unexplned variance in 1st contrast	=	2.8	7.5% 9.4%	
Unexplned variance in 2nd contrast	=	1.9	5.2% 6.5%	
Unexplned variance in 3rd contrast	=	1.8	4.7% 5.9%	
Unexplned variance in 4th contrast	=	1.6	4.3% 5.5%	
Unexplned variance in 5th contrast	=	1.5	3.9% 4.9%	

Figure 3. Item unidimensionality.

4.4. Item Map

Figure 4 illustrates the hierarchy of difficulty in the person–item map (SD = 1.04; M = 0). The symbols # and . represent individuals in the left column, whereas the lettered items with corresponding numbers on the right column indicate their respective difficulty levels among the 30 items. If participants aligned with or were above an item, they had a 50% or greater chance of answering it correctly. For instance, there were 17 participants aligned with Item 25, indicating

that these individuals had a fifty–fifty likelihood of answering the question correctly. Those positioned above Item 25 had a higher probability of answering it correctly, whereas those below this item had a lower possibility. According to Sumintono and Widhiarso[32], the item map is divided into the following four levels based on specific criteria: i) very difficult level > 1 SD; ii) difficult level, between 1 SD and mean logit value; iii) easy level, between mean logit value and −1SD; and iv) very easy level < −1SD (see **Tables 3** and **4**).

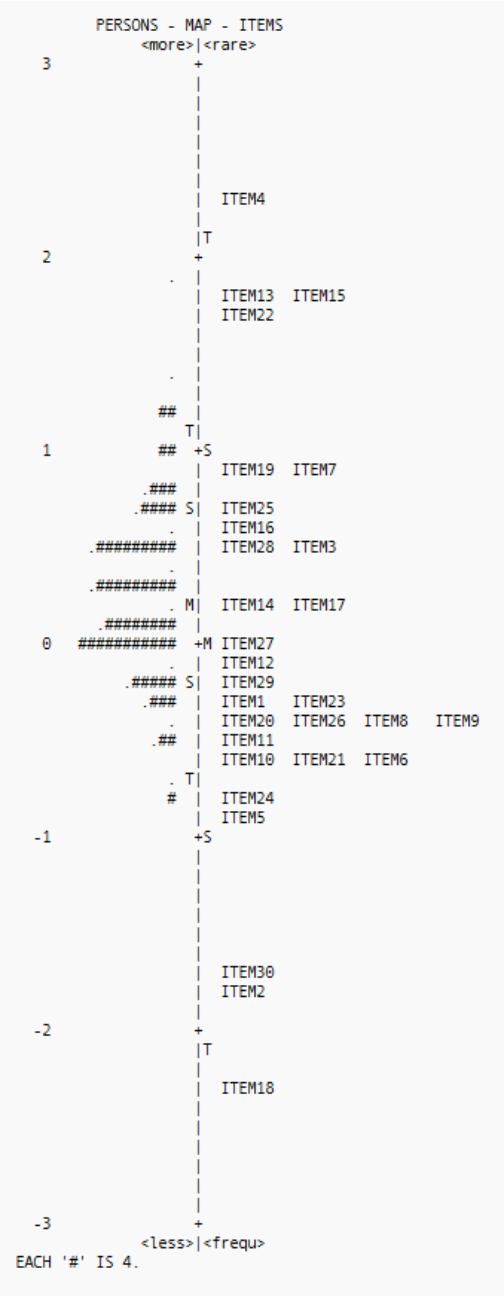


Figure 4. Item map.

The first research question can be addressed using the results presented above. Non-native English speakers were capable of discerning the difficulty hierarchy among six verb classes and four features. However, although the 30 sentences in the questionnaires were easily understandable for the 265 Chinese participants majoring in English, answering all 30 items correctly remained a challenge. This is attributed to the need for careful analysis of each verb class and its corresponding features, which requires relevant semantic knowledge. The remaining two research questions are addressed in the subsequent section.

4.5. Distribution of Verb Classes and Semantic Features

Research question 2 examines the L2 acquisition difficulty distribution of verb classes and their semantic features. **Figures 2** and **4** indicate the students' perceptions of verb classes and their features based on the difficulty hierarchy of items, with a mean of 0 and a standard deviation of 1.04. To identify the specific difficulty levels of verb classes and the generalization of corresponding features, **Tables 3** and **4** present the detailed distributions according to **Figures 2** and **4**.

According to the mean of logit value, **Table 4** shows that the difficulty hierarchy is as follows: *achievements* (1.05) > *semelfactives* (0.50) > *accomplishments* (−0.03) > *states* (−0.06) > *active accomplishments* (−0.36) > *activities* (−0.92). Moreover, based on the analysis of the generalization of features in the very difficult, difficult, easy, and very easy levels, **Table 4** presents the following general property order: *non-dynamic and punctual* (very difficult) → *non-static* (difficult) → non-punctual (easy) → *dynamic and non-punctual* (very easy). The students performed well in verb classes with the features of dynamicity and non-punctuality, but they found it challenging to comprehend verb classes with the features of non-dynamicity and punctuality. Notably, the students still felt confused about the non-static verb classes.

5. Discussion

Research Question 1 was addressed by presenting the results of item reliability, item statistics, the unidimensionality test, and the item map. The item reliability coefficient

Table 3. Item distribution of verb classes and their features.

Difficulty Level	Items (Verb Classes and Features)					Logit Value
Very difficult	Item 4 (2.29) Accomp −dynamic	Item 15 (1.82) Achiev +Telic	Item 13 (1.77) Achiev +punctual	Item 22 (1.72) Seml +punctual		2.29∼1.72
Difficult	Item 19 (0.89) Achiev −dynamic	Item 7 (0.87) State −telic	Item 25 (0.72) Seml −dynamic	Item 16 (0.64) Seml (once) −telic	Item 28 (0.55) Active Accomp +telic	0.89∼0.03
	Item 3 (0.47) Active Accomp +telic	Item 17 (0.24) Accomp −static	Item 14 (0.20) Accomp +telic	Item 27 (0.03) Seml −static		
Easy	Item 12 (−0.10) State −telic	Item 29 (−0.20) Active accomp +telic	Item 1 (−0.27) State −dynamic	Item 23 (−0.27) Achiev +telic	Item 9 (−0.36) State −dynamic	−0.10∼−0.90
	Item 263 (−0.37) Activity −static	Item 203 (−0.40) Active accomp +telic	Item 83 (−0.43) Active accomp +dynamic	Item 113 (−0.46) State −dynamic	Item 213 (−0.59) Seml +dynamic	
	Item 6 (−0.59) Activity −static	Item 10 (−0.63) Accomp −static	Item 24 (−0.82) Active accomp +dynamic	Item 5 (−0.90) Activity −telic		
Very easy	Item 30 (−1.72) Active accomp −punctual	Item 2 (−1.83) Activity +dynamic	Item 18 (−2.25) Accomp −punctual			−1.72∼−2.25

was 0.98, indicating a high degree of measurement accuracy for the 30 items. The results of item statistics, the unidimensionality test, and the item map suggest a reasonable fit of the data. However, it is also noted that the person reliability coefficient (0.09) was inconsistent with item reliability (0.98). Two possible reasons may explain this difference: similar participant abilities and items that were either too easy or too hard items. They seemed to have acquired similar abilities in terms of verb classes either implicitly or explicitly. It is also noted that participants may lack semantic knowledge of verb classes due to lack of the explicit focus on this topic idea in the context of semantic teaching in China. Semantic is only a minor component of English linguistics courses. What's more, according to the ability and difficulty map, about 23% (7 out of 30 items) of all the items turned out to be either too difficult or too easy. Because of this mismatch between ability and item difficulty, the person separation index yielded only 0.43.Therefore, the present research suggests that semantic knowledge—such as understanding the inherent meaning of verb classes and their features—should be incorporated into English teaching programs to help English learners better understand the semantic differences between English and their native language in L2 acquisition environments.

Research Question 2 examined the distribution pattern of verb classes and the semantic features perceived by L2 learners. Based on the perceptions of Chinese English-major undergraduate students, this study established an item hierarchy of L2 English verb class difficulty. A Rasch analysis of the students' perceptions revealed that the verbs involving *semelfactives, achievements,* and *accomplishments* and the verb classes with the properties of non-dynamicity, punc-

Table 4. Frequency, mean of logit value, and general properties, separated according to difficulty hierarchy.

	Verb Classes							General Properties
	Achiev	**Seml**	**Accomp**	**State**	**Active Accomp**	**Activity**	**Total**	
Very difficult	2	1	1				4	−dynamic +punctual
Difficult	1	3	2	1	2		9	−static
Easy	1	1	1	4	4	3	14	−punctual
Very easy			1		1	1	3	+dynamic −punctual
M (logit value)	1.05	0.50	−0.03	−0.06	−0.36	−0.92		

tuality, and non-stativity represented the higher end of the item difficulty hierarchy. The aspectual classification of the verbs involving *activities* and *active accomplishments* and verb classes with the properties of dynamicity and non-punctuality on the lower end of the item difficulty hierarchy were considered processable by average L2 participants.

Although not all non-static verb classes were ranked as the most difficult ones in this study, eventive verbs generally represented the "difficult" level, as opposed to states, and these findings are substantiated by empirical results in semantic research [19]. That is, English and Chinese stative verbs share similarities to some degree. Specifically, in Mandarin Chinese just as in English, a static property is expressed as being steady and persistent over time. An example is as follows: **She was knowing the answer*; *ta zhengzai zhidao daan*. The progressive test appears reliable for L2 learners of Chinese. It cannot take the progressive "zai" in Chinese because *know* ("zhidao") is a stative verb. However, according to Gennari and Poeppel's empirical study, learners take longer to communicate eventive verbs. An example is as follows: **The bubble is popping* and *The bubbles are popping.* In these examples, "pop" as an achievement verb carries the property of punctuality, whereas if the subject "bubbles" is a plural noun, it can take the progressive. Thus, eventive verbs are processible by L2 learners through complex mental activities involving semantic meaning.

The results of this study demonstrate that properties of lexical semantics are the primary factors that impact L2 learners' perceptions of verb classes. The verbs involving *achievements, semelfactives*, and *accomplishments* with the property of non-dynamicity are at the top of the item difficulty hierarchy. This result indicates that there are differences between Van Valin's [5] verb classes in English and those in Mandarin Chinese. To be specific, in Mandarin Chinese, a dynamic event is expressed by the initial distinction of states and events, and a dynamic situation involves change over time—either through a heterogeneous internal structure or a changing endpoint. However, Van Valin's dynamic event in English is expressed by the second distinction based on action —continuous new energy input with diverse intensities. The verbs of achievement (e.g., *The house collapsed* in English or *fangzi daota le* in Mandarin Chinese), semelfactive (e.g., *The police glimpsed the thief* in English or *jingcha piejian le zhege xiaotou* in Mandarin Chinese), and accomplishment (e.g., *The water is freezing* in English or *shui zai jiebing* in Mandarin Chinese) are not considered dynamic based on the criteria (with the application of adverbs, such as "vigorously") of Van Valin's [5] RRG system. These three types of verb classes are not compatible with force adverbs in Van Valin's classification. It is worth noting that some semelfactives are also dynamic (e.g., *Emily coughed once violently* in English or *aimili julie de kesou le yisheng* in Mandarin Chinese). Considering this restriction in the lexical structure of dynamicity, which is different from the Mandarin Chinese classification of dynamicity/non-stativity, it is not unusual that achievements, semelfactives, and accomplishments were the least easily recognized verbs with non-dynamic features when compared with other verb classes in the study.

In contrast to *achievements, semelfactives,* and *accomplishments, activities* and *active accomplishments* with the property of dynamicity were at the bottom end of the item difficulty hierarchy. In this study, the activities and active accomplishments observed in the results were perceived by average L2 participants; for example*: He ran to her actively*, *I ran energetically*. In these examples, activities and active accomplishments are dynamic. Moreover, the Chinese general dynamic event shares similarities with Van Valin's narrow dynamic concept in activities and active accomplish-

ments. Having, thus, due to the positive cross-linguistic influence in their lexical semantics, these two types of verbs with dynamic features are better processed by L2 learners.

In addition to verb classes with a non-dynamic nature at the top end of the difficulty hierarchy, *achievements* and *semelfactives* with a punctual property were not well processed by the L2 participants either. Although these two verb classes demonstrate a similar pattern to Mandarin Chinese, the pattern in Van Valin's conceptualization reflects complex semantic cognition， which poses great challenges for Chinese learners of English. In general, *achievements* and *semelfactives* are punctual, whereas if *achievements* and *semelfactives* are based on repetitive reading, *achievements* and *semelfactives* are similar to activities. For example, *The bubbles are popping* rather than *The bubble is popping*; *The bird is flapping a wing* rather than *The bird is flapping a wing once*. This restructuring of semantic knowledge regarding an iterative reading of these two verb classes is harder to process cognitively for L2 learners. Another possible explanation is the form of the test. In Van Valin's classification, pace adverbs that can occur with events with temporal duration are used to test the punctuality of English verb classes (e.g., "quickly," "rapidly," and "slowly"). *For example, Emily slowly/quickly realized the mistake, but semelfactives and achievements cannot co-occur with such pace adverbs as* "*slowly,*" "*quickly*" *and so on. For example, The light flashed. However, in the marginal cases, they may occur together with these expressions representing very short time, such as* "in an instant". In contrast, "zhe" is considered a reliable form to test durativity in Mandarin Chinese.

In contrast, *accomplishments, active accomplishments,* and *activities* with a non-punctual property were easily perceived by the participants in the present study. It is possible that the non-punctual property carried by these verbs is objective, intuitively and obviously, and learners process these three types of verb classes by a translation of the counterpart words in their L1; thus, correspondence between these two languages contributes to their acquisition. Examples include the following: *He walked to the box slowly and opened it*; *They are running quickly*; *Lily is learning how to read a clock quickly*. Therefore, it does not impose a cognitive burden on learners, significantly decreasing the processing difficulty.

In sum, the findings regarding L2 verb classes' and the semantic properties' difficulty lend support to the claim that L2 acquisition difficulty is affected by cross-linguistic influences and complex cognitive structures on semantic meaning. Of particular interest to L2 researchers are the opportunities of discerning semantic instruction. Based on the findings, the present study provides the following suggestions regarding the future instruction of verb classes and their semantic features.

6. Practical Implications

Research Question 3 addressed how the distribution patterns of verb classes and their features can be explicitly taught to facilitate their implicit acquisition.. The findings in the present study may provide direct practical applications for L2 English semantic instruction, curriculum design, and cognitive semantic assessment. To begin, it is beneficial for L2 instructors to be aware of Van Valin's verb classes and their properties related to L2 semantic acquisition difficulty. Multiple factors of lexical aspect difficulty, cross-linguistic influence, and learner-related variables come into play in determining the difficulty of L2 acquisition [1, 2, 4, 9, 16, 17]. Such awareness may help teachers take into consideration these factors and construct well-grounded instructional teaching strategies.

In the daily pedagogical practice of focus-on-form on semantic meaning, it is the difficulty hierarchy perceived by L2 learners that guides the selection and evaluation of the aspectual classification of verbs and the semantic features. That is, it is possible that the perceived English verb classes' difficulty by L2 learners determines the instructional order of verb classes in the pre-task and main task instruction. Therefore, L2 semantic instruction should be systematically organized and conscientiously executed to facilitate learning and acquisition in the development of communicative competence and promote instructors' teaching knowledge.

Although the complex cognitive structure of verb classes and cross-linguistic influences are regarded as essential factors for L2 learners to process, it is significant that difficulty items are examined as a way to understand semantic knowledge in the curriculum and improve the accuracy of L2 English learners' semantic knowledge. It is also important to understand that not all verb classes and semantic properties are automatically acquired through positive L1 transfer, such as verb classes with properties of dynamicity

and non-punctuality. For cross-linguistic influence and the complex cognition of semantic knowledge, including verb classes with properties of non-dynamicity and punctuality, it is possible that learners experience regression or fluctuation in their language fluency. How this semantic knowledge can be effectively reinforced should be a focus of all L2 language instructors.

Based on the present findings, CR tasks provide L2 instructors with an exceptional opportunity to conduct semantic instruction. Woll and Paquet's [24] study, involving 47 multilingual learners, revealed three stages of CR tasks to enhance positive transfer and mitigate negative influence due to interlanguage influence: pre-task (discovery), main task (cross-linguistic CR reflection), and post-task (validation). In the pre-task, L2 learners are guided to discover the basic pattern of the target language. In the main task, learners perceive the similarities and differences of intrinsic rules through the reflection of cross-linguistic CR. In the post-task, L2 learners validate the hypothesis and achieve common ground. In the present study, L2 learners' perceived difficulty hierarchy provides a direction for semantic instruction. For example, in the pre-task communicative stage, L2 learners are guided to be aware of the patterns among achievements, semelfactives, and accomplishments in the non-dynamic property, and active accomplishments and activities in the dynamic property—or the patterns between achievements and semelfactives in the punctual feature and accomplishments, active accomplishments, and activities in the non-punctual feature. As Woll and Paquet [24] suggested, following the discovery stage, L2 learners experience reflection on intrinsic rules to find the correspondences and differences of these verb classes, and they ultimately reach a consensus on semantic meaning in the validation process. The present study suggests that deductive and inductive instruction can serve as the main teaching strategies integrated into the three stages of CR tasks. For example, deductive instruction aims to provide examples of verb classes with dynamicity and non-dynamicity, while inductive instruction aims to distinguish verb classes and semantic features based on the data they attempt to explore. The acquisition of six verb classes and their semantic features is a gradual process. L2 instructors may systematically guide learners to perceive the similarities and discrepancies in the communicative stage and encourage them to develop critical thinking.

7. Conclusions

Based on the collective perceptions of Chinese learners of English, this study has established a difficulty hierarchy of L2 English verb classes and semantic properties within the RRG framework presented in Van Valin [5]. Through examining the general nature of the difficulty of six verb classes and four pairs of inherent features, the research results demonstrated that students with strong abilities in verb classes featuring non-punctuality and dynamicity encounter difficulties in verb classes featuring punctuality and non-dynamicity. Moreover, this study provides evidence for the notion that the acquisition difficulty is influenced by cross-linguistic differences and complex structures of semantic cognition.

These findings have direct implications for L2 English semantic instruction, curriculum design, and cognitive semantic assessment. The present study further proposed the instruction of a CR task on semantic meaning for L2 instructors. However, the findings and implications should be interpreted with caution, given that this study draws on data from Chinese-speaking and English-major learners of English. The difficulty hierarchy of L2 English verb classes established in this study may be inapplicable to non-Chinese-speaking learners of English. Future research endeavors will be directed toward broadening the scope of participants while conducting a more comprehensive exploration of English education regarding verb classes and their features within the context of second language acquisition backgrounds. Moreover, based on the current analysis, we may investigate the acquisition of the Chinese aspect by L1 English speakers and shed light on the effect of L1 transfer in the L2 acquisition of aspectual classification of verbs.

Author Contributions

Conceptualization, L.F., J.S. and H.L.; methodology, J.S., L.F.; software, J.S.; formal analysis, H.L., L.F.; writing—original draft preparation, L.F.; writing—review and editing, L.F., H.L. and J.S.; visualization, L.F. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments

We would like to thank all the individuals who participated in the study project and anonymous reviewers who provided valuable feedback.

Conflicts of Interest

There is no conflict of interest involved in this study.

Appendix A

(1) We are having only two realistic choices. ( )

(2) They are running quickly. ( )

(3) They walked into her office in one hour. ( )

(4) Lucy learned technical skills actively. ( )

(5) We walked up the mountain for two hours. ( )

(6) Everything is running along smoothly. ( )

(7) I loved those big cats for an hour. ( )

(8) He ran to her actively. ( )

(9) If you're fluent in French, you quickly have a better chance of getting into a good university. ( )

(10) Researchers are learning that the benefits of the sweet may not stop there. ( )

(11) I actively have a few things to finish up. ( )

(12) I loved her action movies in two hours. ( )

(13) He recognized you quickly. ( )

(14) She learned well for one hour. ( )

(15) He fell several times, and then collapsed in 10 minutes. ( )

(16) The light flashed once for an hour. ( )

(17) I was realizing the new computerization system would not increase our workload. ( )

(18) Lily is learning how to read a clock quickly. ( )

(19) Rome collapsed actively. ( )

(20) She painted a picture on the side of his fire-house in one hour. ( )

(21) Jim coughed once violently. ( )

(22) A small blue square flashed near the cross quickly. ( )

(23) I recognized him from somewhere for one hour. ( )

(24) I ran energetically to the cliff falling down to the beach and ocean below. ( )

(25) Tom glimpsed Lily strongly. ( )

(26) I was marching to the room with feelings of doom and foreboding heavy in my heart. ( )

(27) The light is flashing once. ( )

(28) New England pioneers boiled and ate the nuts for one hour. ( )

(29) James painted Mary's portrait in one hour. ( )

(30) He walked to the box slowly and opened it. ( )

References

[1] Shirai, Y., Nishi, Y., 2002. Lexicalisation of Aspectual Structures in English and Japanese. In: Ramat, A.G. (ed.). Typology and Second Language Acquisition. De Gruyter Mouton: Berlin, NY, United States. pp. 267–290. DOI: https://doi.org/10.1515/9783110891249.267

[2] Alruwaili, R., 2014. Second Language Acquisition of Aspect and Tense by Saudi-Arabic Learners of English [Doctoral Dissertation]. University of York: York, UK.

[3] Rastelli, S., 2020. The Acquisition of Aspect in a Second Language. Cambridge University Press: Cambridge, UK. DOI: https://doi.org/10.1017/9781108903455

[4] Sun, Y., Rodríguez, L.D., 2023. A Study of the Distribution of Classes of Lexical Aspect in Spanish by Chinese Learners of Spanish in Oral Narration. Poznan Studies in Contemporary Linguistics. 59(2), 455–474. DOI: https://doi.org/10.1515/psicl-2022-1006

[5] Van Valin, Jr., R.D., 2005. Exploring the Syntax–Semantics Interface. Cambridge University Press: Cambridge, UK. DOI: https://doi.org/10.1017/CBO9780511610578

[6] Vendler, Z., 1967. Verbs and Times. In: Vendler, Z. Linguistics in Philosophy. Cornell Univer-

sity Press: Ithaca, NY, USA. pp. 97–121. DOI: https://doi.org/10.7591/9781501743726-005

[7] Smith, C.S., 1997. The Parameter of Aspect (2nd edition). Kluwer: Dordrecht, The Netherlands. DOI: https://doi.org/10.1007/978-94-011-5606-6

[8] Matasović, R., 2009. A New Typology of Control Constructions Within the Framework of Role and Reference Grammar. Studies in Role and Reference Grammar. pp. 305–318.

[9] Phillips, A., Thiengburanathum, P., 2007. Verb Classes in Thai. Language and Linguistics. 8(1), 167–191.

[10] Sayyad, I., Mirdehghan Farashah, M., Rezaei, A., 2021. Analyzing Arabic Static Structures Based on Role and Reference Grammar and Using Evidence from the Holy Qur'an. Literary Quranic Researches. 9(2), 131–164.

[11] Dowty, D., 1979. Word Meaning and Montague Grammar. Reidel: Dordrecht, The Netherlands.

[12] Van Valin, R.D., LaPolla, R.J., 1997. Syntax: Structure, Meaning, and Function. Cambridge University Press: Cambridge , UK. DOI: https://doi.org/10.1017/CBO9781139166799

[13] Xiao, R., McEnery, T., 2004. Aspect in Mandarin Chinese: A Corpus-Based Study. John Benjamins Publishing: Amsterdam, the Netherlands. DOI: https://doi.org/10.1075/slcs.73?locatt=mode:legacy

[14] Chu, C., 2006.《汉语篇章语法》A Discourse Grammar of Mandarin Chinese. Beijing Language and Culture University Press: Beijing China.

[15] Smith, C.S., 1991. The Parameter of Aspect, 1st ed. Kluwer: Dordrecht, the Netherlands. DOI: https://doi.org/10.1007/978-94-015-7911-7

[16] Jiang, N., 2002. Form–Meaning Mapping in Vocabulary Acquisition in a Second Language. Studies in Second Language Acquisition. 24(4), 617–637. DOI: https://doi.org/10.1017/S0272263102004047

[17] Jiang, N., 2004. Semantic Transfer and Its Implications for Vocabulary Teaching in a Second Language. The Modern Language Journal. 88(3), 416–432.

[18] Sun, Y., 2019. Estudio Sobre la Adquisición de los Tiempos Aspectuales Simples del Pasado en Español por Aprendices Chinos a Partir de Distintas Perspectivas Lingüísticas [Unpublished Ph.D. Dissertation]. Universitat de Barcelona：Barcelona, Spain.

[19] Gennari, S., Poeppel, D., 2019. Events Versus States: Empirical Correlates of Lexical Classes. Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society. Psychology Press: London, UK. pp. 351–356.

[20] Purpura, J.E., 2004. Assessing Grammar. Cambridge University Press: Cambridge, UK.

[21] Larsen-Freeman, D., 2001. Teaching Grammar. In: Teaching English as a Second or Foreign Language. 3, 251–266.

[22] Shim, J.W., Heechul, L., 2012. An Analysis of Semantic Knowledge of L2 Adult Learners of English Concerning Implicative/Non-Implicative Verbs. The Linguistic Association of Korea. 20(2), 23–40. DOI: https://doi.org/10.24303/lakdoi.2012.20.2.23

[23] Dalpanagioti, T., 2023. What Frame Semantics Can Offer to Task - Based Language Teaching. TESOL Journal. 14(2), e692. DOI: https://doi.org/https://doi.org/10.1002/tesj.692

[24] Woll, N., Paquet, P.L., 2025. Developing Crosslinguistic Awareness Through Plurilingual Consciousness-Raising Tasks. Language Teaching Research. 29(1), 63–87. DOI: https://doi.org/10.1177/13621688211056544

[25] Hu, Z., 2021. A Course in Linguistics. Beijing Book Co. Inc.: Beijing, China.

[26] Nadhirah, N.A., Suryana, D., binti Ahmad, A., 2022. Development and Validation Peaceful Classroom Scale: Rasch Model Analysis. International Journal of Instruction. 15(4), 497–514.

[27] Ramdani, R., Hanurawan, F., Ramli, M., et al., 2021. Development and Validation of Indonesian Academic Resilience Scale Using Rasch Models. International Journal of Instruction. 14(1), 105–120.

[28] Runnels, J., 2012. Using the Rasch Model to Validate a Multiple Choice English Achievement Test. International Journal of Language Studies. 6(4), 141–155.

[29] Clifford, R., Cox, T.L., 2013. Empirical Validation of Reading Proficiency Guidelines. Foreign Language Annals. 46(1), 45–61.

[30] Bond, T.G., Fox, C.M., 2015. Applying the Rasch Model: Fundamental Measurement in the Human Sciences, 3rd ed. Routledge: London, UK.

[31] Nuryanti, S., Masykuri, M., Susilowati, E., 2018. Analisis Iteman dan Model Rasch pada Pengembangan Instrumen Kemampuan Berpikir Kritis Peserta Didik Sekolah Menengah Kejuruan. Jurnal Inovasi Pendidikan IPA. 4(2), 224–233.

[32] Sumintono, B., Widhiarso, W., 2015. Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Trim Komunikata: Cimahi, Indonesia.

[33] Boone, W.J., Staver, R.J., Yale, S.M., 2014. Rasch Analysis in the Human Sciences. Springer: London, UK.

[34] Widyaningsih, S. W., Yusuf, I., 2018. Analisis soal modul laboratorium fisika sekolah I menggunakan racsh model. Gravity: Jurnal Ilmiah Penelitian Dan Pembelajaran Fisika. 4(1), 33–46.