WHIOCE PUBLISHING PTE. LTD.
PROVIDING FIRST-CLASS SCIENTIFIC INFORMATION FOR TOP SCHOLARS
Universe Scientific Publishing

Forum for Linguistic Studies

2022 VOLUME 4 ISSUE 1
EISSN: 2705-0602

- Joint Publication -
SICHUAN UNIVERSITY

Editorial Board

Editor-in-Chief

Editorial Board Members

2022 | Volume 4 | Issue 1

Forum for Linguistic Studies

http://ojs.whioce.com/index.php/FLS

Contents

2022 | Volume 4 | Issue 1

Forum for Linguistic Studies

Editor-in-Chief

Prof. Feng Duan

Sichuan University
China

WHIOCE PUBLISHING PTE. LTD.
PROVIDING
FIRST-CLASS SCIENTIFIC INFORMATION
FOR TOP SCHOLARS

ORIGINAL ARTICLE

A comparison of L2 and L1 speakers' production of adverb positions in the Cardiff variety of Welsh

Bethan Lines*

School of Linguistics, University of Cambridge, Cambridge, United Kingdom

Abstract: The grammaticality of adverb positions varies by language. Consequently, L2 and L1 speakers may differ from each other in their acquisition of adverb positions. Given that L2 Welsh speakers outnumber L1 Welsh speakers in Wales, differences in acquisition may change which adverb positions occur in contemporary Welsh. This study compares which adverb positions L2 and L1 speakers produce in the spoken data from Cardiff in the *CorCenCC* corpus (Knight et al., 2020) in order to identify any differences in acquisition. Comparisons of L2 and L1 English speakers find that L2 speakers consistently acquire novel adverb positions yet they frequently use ungrammatical adverb positions. They also do not acquire additional constraints on adverb positions. This study largely reinforces these findings. First, L2 Welsh speakers produce every adverb position that L1 speakers produce. Secondly, although the definiteness constraint that Borsley et al. (2009: 50) describe is not productive in the sample of Cardiff Welsh speakers, L1 speakers exhibit a heaviness constraint on V-Adv-O that L2 speakers do not. Therefore, L1 transfer neither inhibits the acquisition of adverb positions nor facilitates the acquisition of additional constraints. However, unlike L2 English speakers, L2 Welsh speakers do not produce ungrammatical adverb positions. This likely derives from the lack of transferable adverb positions between Welsh and English rather than a lack of transfer. Therefore, this sample of Cardiff Welsh reinforces the crosslinguistic consistency of L2 speakers' acquisition of adverb positions. It also suggests that L2 Welsh speakers most likely diverge from L1 speakers in the contexts in which they use adverb positions rather than the adverb positions that they use.

Keywords: second language acquisition; adverbs; syntax; Welsh; corpus study

***Corresponding author:** Bethan Lines, School of Linguistics, University of Cambridge, Cambridge, United Kingdom; bl449@cam.ac.uk

Received: April 6, 2022; **Accepted:** June 3, 2022; **Available online:** August 4, 2022

Citation: Lines B (2022) A comparison of L2 and L1 speakers' production of adverb positions in the Cardiff variety of Welsh. *Forum for Linguistic Studies* 4(1): 1–17. DOI: 10.18063/fls.v4i1.1451

1. Introduction

As part of the revitalisation of the Welsh language, the Welsh government has emphasised how English-speaking monolinguals may acquire Welsh, as Williams (2014) summarises. Consequently,

according to the 2018-19 National Survey for Wales (Welsh Government, 2020: 5), 57% of Welsh speakers acquired Welsh as their L2. Much of the literature on the Welsh language concerns how the L2 and L1 varieties of Welsh diverge from each other and how this may influence the contemporary use of Welsh. Robert (2009) and Binks and Thomas (2019) identify phonetic, semantic, lexical, morphological, and pragmatic differences between L2 and L1 Welsh speakers. However, the extent to which L2 speakers diverge from L1 speakers in syntax remains understudied. The present study aims to address this gap by comparing L2 and L1 Welsh speakers' production of a syntactic construction. Specifically, it investigates which "adverb positions" L2 and L1 speakers consider to be grammatical. Adverb positions are the positions in which adverbs occur relative to obligatory clausal elements. For example, in (1), the adverb follows all obligatory clausal elements and, in (2), the adverb intervenes between the subject and the lexical verb: S-Adv-V.

1. Catrin draws the cat **happily**.

2. Catrin **happily** draws the cat.

The grammaticality of adverb positions may vary between languages. In French, adverbs typically cannot occur in S-Adv-V, as (3) illustrates.

3. *Antoine **probablement** confond le poèm avec un autre. (Schifano, 2018: 63)
   *Antoine **probably** confuse.*$_{3SG.PRES}$ *the poem with an other*
   "Antoine is probably confusing the poem with another."

Authors like White (1991), Osborne (2008), Lardiere (2018), and Larsson et al. (2020) compare L2 and L1 speakers' grammaticality judgements and production of adverb positions. However, these authors investigate L2 and L1 speakers of SVO languages whereas Welsh is a VSO language. It remains to be seen whether their findings recur when a speaker's L2 and L1 use different word orders. Therefore, this study replicates the methodology of Osborne (2008) and Larsson et al. (2020) with Welsh-English multilinguals. A corpus study is conducted with the spoken data from Cardiff in the *CorCenCC* corpus (Knight et al., 2020) in order to compare the production of adverb positions between 30 L2 Welsh speakers and 17 L1 Welsh speakers. This allows the investigation into structural constraints on the formation of adverb positions in each variety. Although Cinque (1999) emphasises the role of semantics in adverbial syntax, for the purposes of studying structural constraints, this study does not address semantic constraints on the selection of adverb positions.

Having presented the premise and aims of this study, the next section outlines the structural constraints on adverb positions in Welsh that this study investigates. In §3, investigations into the L2 acquisition of similar constraints in English are reviewed and the research questions and hypotheses of this study are presented as informed by their findings. §4 presents the design of the corpus study and §5 outlines the results. In §6, these results are used to evaluate structural constraints on adverb positions in the L2 and L1 varieties of Welsh in Cardiff. Finally, in §7, the study concludes with the implications of its findings on the L2 acquisition of adverb positions and on the adverbial syntax of Welsh.

2. Adverb positions in Welsh

In order to investigate the grammaticality of adverb positions between L2 and L1 speakers, it is

necessary to outline the adverb positions of Welsh. There has been no extensive analysis of adverb positions in contemporary Welsh. However, King (2003), Borsley and Jones (2005), Roberts (2005), and Borsley et al. (2009) discuss adverbial constructions and they highlight several constraints on the adverbial syntax of Welsh that are reviewed in this section.

In order to discuss adverb positions in Welsh, it is important to distinguish "synthetic" and "periphrastic" constructions (Borsley et al., 2009: 38). These involve different word orders and (4, 5) illustrate the potential adverb positions in each construction. In synthetic constructions, the lexical verb is finite and occurs clause-initially, which produces VSO word order, like in (4).

4. (Adv) **V** (Adv) **S** (Adv) **O** (Adv)
   Cafodd hi damwain
   Have.$_{3SG.PST}$ $_{3SG.FEM}$ *accident*
   "She had an accident."

In periphrastic constructions, the initial verb is a finite auxiliary verb and the lexical verb is a nonfinite "verbnoun" (Deuchar, 2005: 255) that obligatorily follows an aspectual particle. This produces AuxSAspVO word order, like in (5).

5. (Adv) **Aux** (Adv) **S** (Adv) **Asp** (Adv) **V** (Adv) **O** (Adv)
   Mae Non yn licio cwrw
   $_{AUX.3SG.PRES}$ *Non* $_{PROG}$ *like.*$_{INF}$ *beer*
   "Non likes beer."

In King (2003), Borsley and Jones (2005), Roberts (2005), and Borsley et al.'s (2009) analyses of Welsh grammar, they discuss the grammaticality of several of these potential adverb positions. First, (6, 7) illustrate that adverbs may follow VSO or AuxSAspVO.

6. Mae Dafydd **yn** licio cwrw **yn** **fawr** **iawn.** (Adapted from Borsley and Jones, 2005: 97)
   $_{AUX.3SG.PRES}$ *Dafydd* $_{PROG}$ *like.*$_{INF}$ *beer* $_{ADV.PRT}$ ***big*** ***very***
   "Dafydd likes beer very much."

7. Awn ni **eto** mis nesaf. (King, 2003: 256)
   Go.$_{1PL.FUT}$ $_{1PL}$ ***again*** *month* *next*
   "We'll go again next month."

In this study, these positions are referred to as "clause-final" because the adverb follows all obligatory clausal elements. However, optional phrases may also follow the adverb, like in (7). King (2003: 262–263) also identifies that an "afterthought" category of adverbs may occur clause-finally or clause-initially. However, he elaborates that certain afterthought adverbs like *efallai* "maybe" must precede a subordinate clause if they occur clause-initially, like in (8).

8. **Efallai** fod e 'n sâl, King (2003: 263)
   Perhaps $_{COMP}$ $_{3SG.MASC}$ $_{PRED}$ *ill*
   "Perhaps he is ill."

None of the authors that are discussed in this section examine clause-initial adverb positions beyond the context of afterthought adverbs. However, Cinque (1999) observes that clause-initial ad-

verb positions are likely ubiquitous crosslinguistically, which suggests that they are grammatical in Welsh for most adverb categories.

Regarding clause-medial adverb positions, (10, 11) illustrate that Welsh adverbs may precede or follow the subject in periphrastic constructions.

10. Mae ‘r bws **eisoes** wedi gadael (Borsley et al., 2009: 50)
*AUX.3SG.PRES the bus **already** PERF leave.INF*
“The bus has already left.”

11. Mae **wastad** lefrith yn y ffrij. (Roberts, 2005: 10)
*AUX.3SG.PRES **always** milk in the fridge.*
“There’s always milk in the fridge.”

However, Roberts (2005: 10) and Borsley et al. (2009: 50) contend that Welsh adverbs may only precede an indefinite subject. Borsley et al. (2009: 50) compare (11) with (12) in which *yfory* “tomorrow” cannot precede the definite subject.

12. *Gwelith **yfory** Emrys ddraig. (Borsley et al., 2009: 50)
*see.FUT.3SG **tomorrow** Emrys dragon*
“Emrys will see a dragon tomorrow.”

Note that (11) is a periphrastic construction and (12) is a synthetic construction. By equating (11, 12), Borsley et al. (2009) suggest that adverbs may precede or follow the subject regardless of construction and that additional constraints like the definiteness constraint apply regardless of construction.

Of the six potential clause-medial adverb positions, two do not appear in any example sentence in King (2003), Borsley and Jones (2005), Roberts (2005), or Borsley et al. (2009): S-Adv-O or V-Adv-O. Although this suggests that they are ungrammatical, it does not confirm it. However, in Cinque (1999) and Schifano’s (2018) crosslinguistic observations of adverb positions, they contend that clause-medial adverb positions do not co-exist in a language with equal markedness. For instance, S-Adv-V is grammatical in English where V-Adv-O is not. In French, the converse is true, as (2, 3, 13, 14) illustrate.

2. Catrin **happily** draws the cat.

13. *Catrin draws **happily** the cat.

3. *Antoine **probablement** confond le poèm avec un autre. (Schifano, 2018: 63)
*Antoine **probably** confuse.3SG.PRES the poem with an other*
“Antoine is probably confusing the poem with another.”

14. Antoine confond **probablement** le poèm avec un autre. (Schifano, 2018: 63)
*Antoine confuse.3SG.PRES **probably** the poem with an other*
“Antoine is probably confusing the poem with another.”

Cinque (1999: 33) concedes that further research must confirm the mutual exclusivity of clause-medial adverb positions outside of SVO languages. This weakens the applicability of their

findings to Welsh. However, having considered the literature, the adverbial syntax of Welsh is expected to employ clause-initial and clause-final adverb positions, a definiteness constraint on adverb positions relative to the subject, and an exclusivity constraint on adverb positions that precede the object.

3. The L2 acquisition of adverb positions

The majority of clause-medial adverb positions and the definiteness constraint that were outlined in §2 are not present in English. Therefore, L2 Welsh speakers must acquire these structural constraints. Whether an author believes L2 and L1 acquisition to be fundamentally distinct like Lardiere (2018) or a product of context like Cook (2016), it is largely uncontroversial that L2 and L1 speakers acquire different varieties of a language. This section reviews four investigations into the L2 acquisition of adverb positions in English. These inform the hypotheses of this study for how L2 Welsh speakers acquire adverb positions, which are presented at the end of this section.

White (1991) and Lardiere (2018) compare L2 and L1 English speakers' grammaticality judgements of adverb positions. White investigates French L2 speakers of English and Lardiere investigates a Hokkien and Mandarin L2 speaker of English. S-Adv-O is typically ungrammatical in French whereas it is grammatical in Hokkien and Mandarin. However, in both studies, L2 speakers consistently judge S-Adv-O to be grammatical in English. A speaker's L1 does not obviously inhibit their acquisition of adverb positions. L2 speakers more frequently diverge from L1 speakers in their acceptance of V-Adv-O, which is ungrammatical in English. French L2 speakers of English frequently accept V-Adv-O like they would accept it in French. White argues that L2 speakers require more positive evidence to confirm the ungrammaticality of a familiar adverb position than to confirm the grammaticality of an unfamiliar adverb position. However, Hokkien and Mandarin speakers would not typically use V-Adv-O in their L1. Therefore, Lardiere (2018: 43) attributes the L2 speaker's acceptance of V-Adv-O to misanalysis of the object instead. Lardiere notes that in (16), adverbs may intervene between the verb and an optional prepositional phrase but, in (17), adverbs cannot intervene between the verb and an optional direct object.

16. The child walked **slowly** (to school).

17. The child ate **slowly** (*her lunch).

After explicit instruction, the L2 speaker in Lardiere (2018) consistently distinguishes (16, 17) and rejects V-Adv-O. However, the L2 speakers in White (1991) continue to accept V-Adv-O in spite of explicit instruction. Together, White and Lardiere's findings suggest that L1 transfer motivates the persistent use of V-Adv-O. Both authors also highlight that L2 speakers violate the exclusivity constraint on S-Adv-O and V-Adv-O that Cinque (1999) observes in English, French, and Mandarin. This suggests that L2 speakers do not acquire additional constraints on adverb positions as consistently as they acquire novel adverb positions and that the presence of a constraint in their L1 does not facilitate the acquisition of that constraint in their L2.

Whereas White (1991) and Lardiere (2018) investigate L2 speakers' knowledge of adverb positions in English, Osborne (2008) and Larsson et al. (2020) investigate L2 speakers' production of adverb positions in English in their corpus studies. They find that L2 speakers consistently produce S-Adv-O yet they use V-Adv-O significantly more frequently than L1 speakers. V-Adv-O is

grammatical in many of the L1 languages of the L2 speakers that Osborne and Larsson et al. study. Therefore, Osborne and Larsson et al.'s findings reinforce that L1 transfer inhibits which adverb positions L2 speakers reject more than which adverb positions L2 speakers accept. Osborne also addresses L2 speakers' acquisition of additional constraints on adverb positions. Although infrequently, L1 English speakers also produce V-Adv-O and instances of V-Adv-O significantly correlate with a "heavy" object: an object that comprises of numerous phrases (Osborne, 2008: 136). Osborne contends that this heaviness constraint licenses V-Adv-O in English. This explains the grammaticality of (18) in contrast to (19).

18. I recited **quite elegantly** [$_{NP}$ the letter [$_{CP}$ that my sister from Swansea sent me]].

19. *I recited **quite elegantly** [$_{NP}$ the letter].

This does not violate the exclusivity constraint on S-Adv-O and V-Adv-O because V-Adv-O is marked whereas S-Adv-O is unmarked. Despite the frequency of V-Adv-O in the L2 dataset, Osborne (2008: 136) finds no significant relationship between V-Adv-O and the heaviness of the object in the L2 dataset. This simultaneously reinforces that L2 speakers transfer the unmarked grammaticality of V-Adv-O in their L1 into English and that L2 English speakers do not acquire additional constraints like the heaviness constraint or the exclusivity constraint.

3.1. Research questions and hypotheses

Having reviewed four comparisons of adverb positions between L2 and L1 English speakers, this section returns to the central question of this study: do L2 and L1 Welsh speakers consider the same adverb positions to be grammatical? In order to address this, this study investigates whether L2 Welsh speakers significantly differ from L1 Welsh speakers in which adverb positions they produce. White (1991), Osborne (2008), Lardiere (2018), and Larsson et al. (2020) observe that L2 speakers consistently acquire novel adverb positions. However, each study also observes that L2 speakers frequently accept adverb positions that L1 speakers do not. Furthermore, despite the consistency with which L2 speakers acquire novel adverb positions, L2 speakers do not exhibit additional constraints on adverb positions regardless of the presence of that constraint in their L1. However, if L2 English speakers' use of V-Adv-O derives from transfer, the incongruity of structure between Welsh and English may prevent L2 Welsh speakers' use of ungrammatical adverb positions. The possibility remains that this incongruity of structure may also inhibit L2 Welsh speakers' acquisition of adverb positions. However, if the findings of the four studies that are discussed are applicable to a VSO language, the hypotheses of this study are the following:

1. L2 speakers will produce every adverb position that L1 speakers produce.
2. L2 speakers will not produce more adverb positions than L1 speakers.
3. L2 speakers will not exhibit additional constraints on adverb positions that L1 speakers exhibit.

4. Materials and methodology

Having outlined the research questions and hypotheses of this study, this section presents the design of the corpus study that investigates the production of adverb positions in L2 and L1 varie-

ties of Welsh. The methodology largely replicates that of Osborne (2008) and Larsson et al. (2020). This facilitates a statistical analysis through which the significance of the L2/L1 distinction and two structural variables on the frequency of each adverb position is determined. This also facilitates a comparison of the L2 acquisition of adverb positions between the VSO language in this study and the SVO language in Osborne and Larsson et al.'s studies.

4.1. Corpus

The frequency of each adverb position is measured in the unscripted spoken data from the Welsh-language corpus *CorCenCC* (Knight et al., 2020). Exclusively analysing unscripted spoken data minimises variation in register and spontaneity. The spoken data of *CorCenCC* comprises of over 2.8 million words, which makes it the largest spoken Welsh-language corpus available. Knight et al. (2020) sourced their data from pre-organised events and from crowdfunding between July 2016 and September 2020. *CorCenCC* contributors optionally supplied metadata through a questionnaire. This includes whether they identify as a "learner", which is taken to mean 'L2 speaker' in this study. The results from the L2 and L1 datasets are presented separately in order to compare them. The metadata sourced from the questionnaire is occasionally incomplete or erroneously labelled. If it cannot be reliably identified whether a contributor identifies as a learner, they are excluded from analysis. Contributors also supplied their county of origin. Exclusively analysing the Welsh spoken data from Cardiff controls for regional variation. Cardiff was selected because of the number of contributions and L2 contributors. In total, the dataset comprises of 47 Welsh speakers: 30 L2 speakers and 17 L1 speakers.

4.2. Methods

Every word in *CorCenCC* is transcribed and tagged with its lemma and part-of-speech. This facilitates the measurement of the frequency in which adverbs occur in each adverb position. §2 outlines every potential adverb position in periphrastic and synthetic constructions. **Table 1** presents the 11 adverb positions that this study measures.

The subject, object, and auxiliary verb may be omitted in Welsh. Because this obfuscates the position of adverbs relative to obligatory elements, this study delimits adverb positions relative to overt obligatory elements. However, in synthetic constructions, every instance of a verb and one argument is considered to be V-Adv-S because distinguishing V-Adv-S and V-Adv-O would be liable to misinterpretation. Clause-initial and -final adverbs are not distinguished by which clausal elements they precede or follow respectively because omitted elements do not obscure their position.

Regarding the structures that this study examines, constructions are not distinguished by mood or polarity if mood or polarity do not affect the word order. Conversely, constructions that violate the

Table 1. Potential adverb positions in periphrastic and synthetic constructions

Periphrastic	Synthetic
Clause-Initial	Clause-Initial
Aux-Adv-S	V-Adv-S
Aux-Adv-Asp	S-Adv-O
S-Adv-Asp	Clause-Final
Asp-Adv-V	
V-Adv-O	
Clause-Final	

canonical periphrastic or synthetic word orders are excluded. Adverbs are also excluded from analysis if they comprise part of a distinct grammatical construction. For example, *ddim* "not", *erioed* "(n)ever", and *byth* "(n)ever" must intervene between the subject and the aspectual particle in standard negation, like in (20). This does not reflect the typical range of adverb positions that this study investigates.

20. Dydy Gwyn **ddim/byth** yn cysgu. (Adapted from Borsley and Jones, 2005: 96).
AUX.NEG.3SG.PRES *Gwyn* *NEG*/***never*** *PROG* *sleep.INF*
"Gwyn is not sleeping."

Finally, the frequencies of each adverb position are tabulated with each dataset. The significance of any discrepancies in the use of an adverb position between L2 and L1 speakers is determined with χ^2-tests. This study also investigates two variables in order to determine the presence of additional constraints on adverb positions. First, for each instance of Aux-Adv-S, S-Adv-Asp, V-Adv-S, and S-Adv-O, the definiteness of the subject is recorded. Secondly, although Borsley et al. (2009) do not identify a heaviness constraint in Welsh, Osborne (2008) observes it in English. In order to investigate the transfer of additional constraints, this study tests for the heaviness constraint in Welsh. Therefore, for each instance of V-Adv-O, S-Adv-O, and O-Adv, the heaviness of the object is recorded. For the purposes of this study, an object is "heavy" if it contains a prepositional phrase or a relative clause. The frequency of each relevant adverb position is tabulated with the definiteness of the subject and the heaviness of the object respectively in order to determine the significance of these relationships with χ^2-tests. Note that for each of the χ^2-tests that are conducted, if any variable has fewer than 5 tokens, Yate's Correction is applied in order to minimise the influence of outliers. In summary, this corpus study allows the investigation into whether the use of adverb positions and the presence of additional constraints significantly varies between the sample of L2 and L1 Welsh speakers in Cardiff.

5. Results

This section presents the results of the corpus study. First, it presents the frequency of adverb positions in periphrastic and synthetic constructions between the L2 and L1 datasets and it highlights any significant discrepancies in use. Secondly, it presents the significance of the definiteness of the subject and the heaviness of the object in each dataset with additional reference to a grammaticality judgement when there are too few tokens to calculate significance. Frequency is presented as a number of tokens and a percentage.

5.1. Adverb positions

In Osborne (2008) and Larsson et al.'s (2020) corpus studies, L2 speakers produce ungrammatical adverb positions that are grammatical in their L1. Because of the structural incongruity between Welsh and English, in §3.1, the second hypothesis predicts that L2 Welsh speakers do not use more adverb positions than L1 speakers. Regarding periphrastic constructions, **Table 2** presents the frequency of each of the seven potential adverb positions across both datasets.

Table 2 shows that adverbs frequently occur clause-initially and -finally in both datasets, like in (21, 22).

Table 2. Frequency of adverb positions in periphrastic constructions in the L2 and L1 datasets

	L2	L1	Total
Clause-Initial	112	117	229
	26.23%	*33.52%*	*29.51%*
Aux-Adv-S	1	0	1
	0.23%	*0.00%*	*0.13%*
Aux-Adv-Asp	12	5	17
	2.81%	*1.43%*	*2.19%*
S-Adv-Asp	36	35	71
	8.43%	*10.03%*	*9.15%*
Asp-Adv-V	19	10	29
	4.45%	*2.87%*	*3.74%*
V-Adv-O	1	11	12
	0.23%	*3.15%*	*1.55%*
Clause-Final	246	171	417
	57.61%	*49.00%*	*53.74%*
Total	427	349	776
	100.00%	*100.00%*	*100.00%*

21. **Yfory** bydd y cymdeithas Cymraeg Kenya yn cwrdd (lla_gw_170228_001)
Tomorrow *AUX.3SG.FUT the society Welsh Kenya PROG meet.INF*
"Tomorrow, the Welsh society of Kenya will meet."

22. Mae 'n glawio **'n aml.** (lla_gw_170228_001)
AUX.3SG.PRES PROG rain.INF ***ADV.PRT*** ***often***
"It rains often."

Adverbs also precede and follow the aspectual particle, like in (23, 24).

23. […] eu bod nhw **jyst** yn gofyn cwestiynau (lla_gw_170517_001)
[…] 3PL.GEN COMP 3PL ***just*** *PROG ask.INF questions*
"[…] that they just ask questions."

24. Dw i 'n **jyst** dod lan (lla_gw_170228_001)
AUX.1SG.PRES 1SG PROG ***just*** *come.INF* down
"I'm just coming down."

However, there are two adverb positions in periphrastic constructions that seldom occur in one of the datasets. First, Aux-Adv-S only occurs once in the L2 dataset, in (25), and it does not occur in the L1 dataset. Secondly, V-Adv-O only occurs once in the L2 dataset, in (26), yet it occurs 11 times in the L1 dataset.

25. Mae **jyst** yr holl peth yn aneglur (lla_gw_170221_001_1)
AUX.3SG.PRES ***jyst*** *the whole thing PRED unclear*
"The whole thing is just unclear."

26. Fi wedi gweld **hefyd** ferswin yr Jungle Book (lla_gw_170221_001_1)
1SG PERF see.INF ***too*** *version the Jungle Book*
"I've seen that version of the Jungle Book too."

A χ^2-test was performed with the frequency of each clause-medial adverb position in periphrastic constructions and each dataset. With Yate's Correction, it indicates that the relationship between these variables is highly significant (n = 130, df = 4, χ^2 = 19.52, p <0.001). However, when V-Adv-O is included in the test and Aux-Adv-Asp is excluded, the relationship remains significant (n = 129, df = 3, χ^2 = 13.84, p = 0.003). Conversely, when Aux-Adv-Asp is included and V-Adv-O is excluded, the relationship is no longer significant (n = 118, df = 3, χ^2 = 5.52, p = 0.137). Therefore, L2 and L1 speakers likely significantly differ in their use of V-Adv-O but not in their use of Aux-Adv-Asp.

Regarding synthetic constructions, **Table 3** presents the frequency of each of the four potential adverb positions between the datasets.

As in periphrastic constructions, Table 3 illustrates that adverbs may occur clause-initially or -finally in synthetic constructions in both datasets, like in (27).

27. **Unwaith** cerddais i **adre.** (lla_gw_170221_001_1)
Once *walk.$_{1SG.PST}$ $_{1SG}$* ***homeward***
"Once I walked home."

However, in the L2 dataset, adverbs do not occur clause-medially. In the L1 dataset, V-Adv-S also does not occur and S-Adv-O only occurs once, in (28).

28. Planna di **yna** ein cawrn (lla_gw_170228_001)
Plant-$_{HAB.2SG}$ $_{2SG}$ ***there*** *$_{1PL.GEN}$ love*
"You'll plant **there** our love."

A χ^2-test was performed with the frequency of each adverb position in synthetic constructions and each dataset. V-Adv-S was omitted from analysis owing to its lack of tokens. With Yate's Correction, it indicates that their relationship is not significant (n = 89, df = 2, χ^2 = 3.48, p = 0.176). Therefore, χ^2-tests only indicate one significant difference in adverb positions between the L2 and L1 datasets in the production of V-Adv-O.

Table 3. Frequency of adverb positions in synthetic constructions in the L2 and L1 datasets

	L2	**L1**	**Total**
Clause-Initial	12	7	19
	26.09%	*15.91%*	*21.35%*
V-Adv-S	0	0	0
	0.00%	*0.00%*	*0.00%*
S-Adv-O	0	1	1
	0.00%	*2.27%*	*1.12%*
Clause-Final	33	36	69
	71.74%	81.82%	77.53%
Total	45	44	89
	100.00%	*100.00%*	*100.00%*

5.2. Additional constraints

Roberts (2005: 10) and Borsley et al. (2007: 50) contend that Welsh adverbs may precede the subject if the subject is indefinite. This study had intended to investigate the significance of the definiteness of the subject on adverb positions relative to the subject. However, Aux-Adv-S, V-Adv-S, and S-Adv-O do not occur more than once each, which limits comparison. The one instance of Aux-Adv-S in the L2 dataset in (25) involves a definite subject, which violates the proposed definiteness constraint.

25. Mae **jyst** yr holl peth yn aneglur (lla_gw_170221_001_1)
AUX.3SG.PRES ***jyst*** *the whole thing* *PRED* *unclear*
"The whole thing is just unclear."

The one instance of S-Adv-O in the L1 dataset in (28) also involves a definite subject, which does not violate the proposed constraint.

28. Planna di **yna** Ein cawrn (lla_gw_170228_001)
Plant-HAB.2SG *2SG* ***there*** *1PL.GEN* *love*
"You'll plant **there** our love."

However, χ^2-tests indicate that the one instance of Aux-Adv-S and S-Adv-O are not significant. Likewise, (26, 29) may be outliers that do not reflect L2 and L1 speakers' use of the definiteness constraint.

In order to supplement the lack of data from the corpus study, this study also tested an L1 Welsh speaker's grammaticality judgements of sentences where the definiteness constraint theoretically applies. The L1 speaker is a 60-year-old from north Wales. He speaks a different dialect to the Welsh speakers in Cardiff. However, the author's access to L1 Welsh speakers is limited and neither Roberts (2005) nor Borsley et al. (2009) claim that the constraint is region-dependent. Therefore, his judgements may still inform the discussion of the definiteness constraint in this study. The L1 speaker was presented with the two example sentences that Borsley et al. (2009: 50) use to justify the definiteness constraint as well as each sentence's alternative adverb position. The L1 speaker judges (11, 12, 29, 30) as follows.

11. ?Mae **wastad** lefrith yn y ffrij. (Roberts, 2005: 10)
AUX.3SG.PRES ***always*** *milk in the fridge.*
"There's always milk in the fridge."

29. Mae lefrith **wastad** yn y ffrij.
AUX.3SG.PRES *milk* ***always*** *in the fridge.*
"There's always milk in the fridge."

12. *Gwelith **yfory** Emrys ddraig. (Borsley et al., 2009: 5)
see.FUT.3SG ***tomorrow*** *Emrys dragon*
"Emrys will see a dragon tomorrow."

30. *Gwelith Emrys **yfory** ddraig.
see.$_{FUT.3SG}$ *Emrys* ***tomorrow*** *dragon*
"Emrys will see a dragon tomorrow."

The L1 speaker considers (29) to be "correct" and he judges (11) to be "incorrect but [he has] heard it being used and it makes clear sense". He does not conform to the definiteness constraint but he may recognise it. However, he rejects both (12, 30). This mirrors the absence of clause-medial adverb positions in Table 3 but it contradicts Borsley et al.'s (2009: 50) claim that the ungrammaticality of (12) derives from the definiteness constraint. The infrequency of adverb tokens in the dataset and the author's limited access to L1 Welsh speakers limits the strength of this evidence. However, overall, the L2 and L1 datasets resemble each other in this dearth of evidence.

The second additional variable that this study investigates is the significance of the heaviness of the object on V-Adv-O, S-Adv-O, and O-Adv. Although S-Adv-O only occurs once in total and V-Adv-O only occurs once in the L2 dataset, V-Adv-O occurs 11 times in the L1 dataset. This allows the calculation of the significance of the heaviness constraint in periphrastic constructions in the L1 dataset. In the L1 dataset, 63.64% of instances of V-Adv-O involve a heavy object, like the relative clause in (31).

31. Da ni 'n gallu gweld **yn** **amlwg** (lla_gw_170302_001)
AUX.3PL.PRES *3PL* *PROG* *able*$_{.INF}$ *see*$_{.INF}$ ***ADV.PRT*** ***obviously***
[sydd yn rhan o gymuned wleidyddol]
COMP *PRED* *part* *of* *community* *political*
"We are able to see obviously [who is part of a political community]."

Furthermore, **Table 4** illustrates that V-Adv-O more frequently occurs with heavy objects than O-Adv. A χ^2-test was performed with the frequency of both adverb positions and the heaviness of the object. With Yate's Correction, it indicates that their relationship is highly significant ($n = 49$, $df = 1$, $\chi^2 = 14.13$, $p < 0.001$). The lack of tokens means that the significance of the heaviness of the object cannot be calculated in the L2 dataset or in synthetic constructions. However, this lack of tokens reveals a discrepancy between the L2 and L1 datasets in periphrastic constructions.

6. Discussion

Having presented the results of the corpus study, first, this section evaluates which adverb positions and constraints vary in grammaticality between L2 and L1 Welsh speakers. Secondly, it discusses whether these findings conform to the hypotheses in §3.1.1, which facilitates the evaluation

Table 4. The frequency of V-Adv-O and O-Adv and the heaviness of the object in the L1 dataset

	Heavy	**Light**	**Total**
V-Adv-O	11	1	12
	55.00	*3.45*	*24.49*
O-Adv	9	28	37
	45.00	*96.55*	*75.51*
Total	20	29	49
	100.00	*100.00*	*100.00*

of whether L2 Welsh speakers significantly differ from L1 Welsh speakers in which adverb positions they acquire.

Regarding the production of adverb positions, in §5, there are three discrepancies between the datasets. S-Adv-O and Aux-Adv-Asp both occur once in the L1 and L2 datasets respectively yet they do not occur in the other dataset. Initially, this may suggest that these adverb positions vary in grammaticality between L2 and L1 speakers. However, first, χ^2-tests indicate that L2 and L1 speakers do not significantly differ from each other in their production of adverb positions in synthetic constructions. Secondly, although L2 and L1 speakers significantly differ from each other in their production of clause-medial adverb positions in periphrastic constructions, this relationship remains significant if Aux-Adv-Asp is excluded from analysis. L2 and L1 speakers likely do not significantly differ in their production of S-Adv-O or Aux-Adv-Asp. Therefore, either (25, 28) are outliers and each adverb position is ungrammatical for both L2 and L1 speakers or both L2 and L1 speakers would produce each adverb position in a larger dataset.

25. Mae **jyst** yr holl peth yn aneglur (lla_gw_170221_001_1)
AUX.3SG.PRES ***jyst*** *the whole thing* *PRED* *unclear*
"The whole thing is just unclear."

28. Planna di **yna** ein cawrn (lla_gw_170228_001)
Plant-HAB.2SG *2SG* ***there*** *1PL.GEN* *love*
"You'll plant **there** our love."

Adverbs seldom occur clause-medially in synthetic constructions and grammaticality judgements, which initially suggests that S-Adv-O and V-Adv-O are both ungrammatical. However, the grammaticality judgements specifically reject *yfory* "tomorrow" in S-Adv-O and V-Adv-O. Schifano (2018) observes that the grammaticality of clause-medial adverb positions depends on the semantics of the adverb in Romance languages. For instance, in Spanish, *siempre* "always" may occur in V-Adv-O but not in S-Adv-O. The converse is true for *probablemente* "probably", as (32, 33) illustrate.

32. Sergio (***siempre**) confunde (**siempre**) este poema. (Schifano, 2018: 67)
Sergio ***always*** *confuse.3SG.PRES* ***always*** *this Poem*
"Sergio is always confusing this poem."

33. Sergio (**probablemente**) confunde (***probablemente**) este poema. (Schifano, 2018: 67)
Sergio ***probably*** *confuse.3SG.PRES* ***probably*** *this poem*
"Sergio is probably confusing this poem."

Therefore, S-Adv-O may be grammatical for other adverbs. However, overall, L2 and L1 speakers resemble each other's production of both S-Adv-O and Aux-Adv-Asp regardless of their grammaticality.

The third discrepancy is that V-Adv-O occurs once in the L2 dataset yet it occurs 11 times in the L1 dataset. χ^2-tests determine that L2 and L1 speakers significantly differ from each other in the production of adverb positions in periphrastic constructions. Furthermore, this relationship becomes insignificant if V-Adv-O is excluded from analysis. L2 and L1 speakers likely significantly differ

from each other in their use of V-Adv-O. Therefore, (26) may be an outlier and V-Adv-O may be ungrammatical in L2 varieties of Welsh.

26. Fi wedi gweld **hefyd** ferswin yr Jungle Book (lla_gw_170221_001_1)
1SG PERF see.INF ***too*** *version the Jungle Book*
"I've seen that version of the Jungle Book too."

Alternatively, V-Adv-O may occur more frequently in a larger sample. Regardless, in this study, V-Adv-O is the only adverb position that significantly differs in production between the L2 and L1 datasets.

These findings largely conform to the first two hypotheses of this study. First, seeing that L2 English speakers consistently acquire S-Adv-O in White (1991), Osborne (2008), Lardiere (2018), and Larsson et al. (2020), the hypotheses in §3.1 predict that L2 speakers would produce every adverb position that L1 speakers produce. Although S-Adv-O occurs in the L1 dataset but not in the L2 dataset, this difference is not significant. Therefore, L2 speakers likely acquire each grammatical adverb position in Welsh. This suggests that the incongruity between SVO and VSO languages does not inhibit the acquisition of adverb positions. However, although V-Adv-O occurs once in the L2 dataset, it is significantly less frequent than in the L1 dataset. If it is ungrammatical in the L2 variety of Welsh, this would deviate from the first hypothesis as well as Osborne and Larsson et al.'s findings that L2 speakers consistently produce adverb positions that are ungrammatical in their L1. Furthermore, L2 and L1 Welsh speakers produce every other adverb position at similar frequencies. In order to account for the discrepancies between these findings, an alternative explanation for the infrequency of V-Adv-O is proposed in the discussion of additional constraints after the discussion of the second hypothesis.

The second hypothesis predicted that L2 speakers would not use more adverb positions than L1 speakers. This hypothesis deviates from L2 English speakers' use of V-Adv-O in White (1991), Osborne (2008), and Larsson et al. (2020). However, in these three studies, V-Adv-O is grammatical in the multilinguals' L1 whereas in Lardiere (2018) and in this study, it is not. L2 speakers use V-Adv-O more persistently if V-Adv-O is present in their L1. Therefore, L2 speakers' use of ungrammatical adverb positions likely derives from L1 transfer. The adverb positions that do not occur in Welsh are not grammatical in English. Therefore, English transfer would not motivate L2 Welsh speakers to produce ungrammatical adverb positions. Likewise, L2 and L1 speakers in this study do not significantly differ from each other in which adverb positions they do not produce. Clause-medial adverb positions in synthetic constructions and Aux-Adv-S seldom occur in both datasets, which conforms to the second hypothesis.

Finally, regarding the third hypothesis, White (1991) and Osborne (2008) argue that L2 speakers do not acquire the clause-medial exclusivity constraint or the heaviness constraint. Therefore, the third hypothesis predicts that L2 Welsh speakers would not acquire additional constraints on adverb positions. First, this study investigates L2 and L1 speakers' production of the definiteness constraint that Borsley et al. (2009: 50) describe. However, in both datasets, Aux-Adv-S, V-Adv-S, and S-Adv-O seldom occur. The Welsh speakers in the present dataset do not productively distinguish adverb positions relative to the subject. The fact that the L1 informant in this study recognises yet rejects (11, 12, 31) supports the idea that the definiteness constraint is present in Welsh, which Borsley et al. describe, yet it is not ubiquitous, which explains its absence in the present dataset.

11. ?Mae **wastad** lefrith yn y ffrij. (Roberts, 2005: 10)
AUX.3SG.PRES ***always*** *milk in the fridge.*
"There's always milk in the fridge."

12. *Gwelith **yfory** Emrys ddraig. (Borsley et al., 2009: 5)
see.FUT.3SG ***tomorrow*** *Emrys dragon*
"Emrys will see a dragon tomorrow."

31. *Gwelith Emrys **yfory** ddraig.
see.FUT.3SG *Emrys* ***tomorrow*** *dragon*
"Emrys will see a dragon tomorrow."

The definiteness constraint may not be present in every dialect of Welsh. Therefore, it would not be present in Cardiff or in the variety that the L1 informant speaks. Alternatively, it may depend on the adverb. Although the L1 speaker rejects *yfory* "tomorrow" and *wastad* "always", the constraint may become apparent in a larger sample with other adverbs. Regarding L2 speakers' acquisition of this constraint, L2 speakers violate the definiteness constraint in the one instance of Aux-Adv-S and L1 speakers conform to the constraint in the one instance of S-Adv-O. However, the lack of evidence for the definiteness constraint indicates that it is unproductive in both datasets.

However, this study also investigates the heaviness constraint in Welsh. S-Adv-O does not occur more than once in either dataset, which suggests that a heaviness constraint is not productive in synthetic constructions in either variety. This conforms to the absence of clause-medial adverb positions in synthetic constructions. However, in periphrastic constructions, V-Adv-O significantly correlates with heavy objects as it does in Osborne (2008). This also conforms to the clause-medial exclusivity constraint that Cinque (1999) and Schifano (2018) observe. In Welsh and in English, V-Adv-O is likely marked while adverbs more freely intervene between the subject and the lexical verb. Furthermore, like in Osborne (2008), the correlation between V-Adv-O and the heaviness of the object is not present in the L2 dataset. This reinforces the idea that L2 speakers do not acquire additional constraints on adverb positions.

Now that a potential heaviness constraint on V-Adv-O in Welsh and its absence in the L2 dataset has been identified, another explanation for the significant infrequency of V-Adv-O in the L1 dataset becomes possible. In Osborne (2008), L2 speakers produce V-Adv-O yet they do not exhibit the heaviness constraint. Therefore, Osborne argues that L1 transfer motivates L2 speakers' production of V-Adv-O. V-Adv-O is not grammatical in English, however. The heaviness constraint licenses V-Adv-O in the L1 variety of Welsh yet, without the heaviness constraint and without L1 transfer, L2 speakers have no motivation to produce V-Adv-O. Therefore, the significant infrequency of V-Adv-O in the L2 dataset likely reflects differences in L2 speakers' acquisition of additional constraints rather than differences in their acquisition of adverb positions. This conforms to the first and third hypotheses. It is of further importance that both Welsh and English L1 speakers exhibit the heaviness constraint in the present dataset and in Osborne (2008) respectively. However, L2 Welsh speakers do not obviously transfer the heaviness constraint from English into Welsh. This mirrors White (1991) and Osborne's findings that L2 speakers do not exhibit the mutual exclusivity of clause-medial adverb positions despite the presence of that constraint in their L1. Therefore, although L1 transfer may motivate the use of ungrammatical adverb positions if the L2 has a structur-

ally-equivalent adverb position, it does not facilitate the acquisition of additional constraints. Consequently, L2 Welsh speakers likely do not differ from L1 speakers in which adverb positions they acquire yet they diverge in their use of those adverb positions.

7. Conclusion

This study has compared the frequency of adverb positions between L2 and L1 Welsh speakers from Cardiff in the *CorCenCC* corpus in order to identify structural constraints on the adverb positions of each variety. This section concludes with the main implications of the findings of this study on the L2 acquisition of adverb positions in a VSO language and on the adverbial syntax of Welsh.

The findings of this study largely reinforce those from previous investigations into the L2 acquisition of adverb positions. L2 speakers of Welsh and English consistently produce the same adverb positions as L1 speakers. Therefore, the structural congruity between L2 and L1 languages neither facilitates nor inhibits the acquisition of novel adverb positions and these patterns likely recur across L2 speakers regardless of language. However, L2 English speakers produce ungrammatical adverb positions whereas L2 Welsh speakers do not. An examination of structurally incongruous languages suggests that L1 transfer may motivate the use of ungrammatical adverb positions but, without L1 transfer, L2 speakers only acquire grammatical adverb positions. Finally, following White (1991) and Osborne (2008), this study concludes that L2 speakers likely do not acquire additional constraints on adverb positions like the heaviness constraint and the exclusivity constraint. The presence of the heaviness constraint in both Welsh and English provides more evidence that L1 transfer neither motivates the use of additional constraints nor facilitates their acquisition. However, the applicability of these findings for both Welsh and VSO languages at large is limited by the sample of Cardiff Welsh in this study. First, the infrequency of adverb tokens limits the ability to discern outliers. Secondly, the findings do not reflect regional variation or further sociolinguistic variation that may determine some of the patterns that are attributed to L2 and L1 varieties of Welsh. Thirdly, although Osborne and Larsson et al.'s corpus studies largely reflect the findings of White and Lardiere's grammaticality judgements, the corpus study cannot precisely determine which adverb positions L2 and L1 speakers consider to be grammatical.

Finally, by comparing L2 and L1 speakers' production of adverb positions, this study also tests for the presence of constraints on the adverbial syntax of Cardiff Welsh. First, the definiteness constraint that Borsley et al. (2009: 50) describe is not productive in the present dataset. Therefore, the constraint is likely restricted by region or semantics. Further grammaticality judgements are necessary to confirm this. Secondly, a potential heaviness constraint on V-Adv-O is identified and clause-medial adverb positions are likely to be restricted in synthetic constructions, if not ungrammatical. The constraint on clause-medial adverb positions in synthetic constructions recurs more strongly in the present dataset than either the definiteness constraint or the heaviness constraint. This study must be replicated with other dialects of Welsh in order to investigate how widespread these constraints are in larger datasets and across Wales. Regarding Cardiff Welsh, however, the L2-majority speaker population likely does not affect which adverb positions are produced yet L2 speakers diverge from L1 speakers in the contexts in which these adverb positions are used.

Conflict of interest

The author declared no conflict of interest.

References

Binks HL and Thomas EM (2019) Long-term outcomes for bilinguals in minority language contexts: Welsh-English teenagers' performance on measures of grammatical gender and plural morphology in Welsh. *Applied Psycholinguistics* (40): 1019–1049.

Borsley RD and Jones BM (2005) *Welsh Negation and Grammatical Theory*. Cardiff: University of Wales Press.

Borsley RD, Tallerman M, and Willis D (2009) *The Syntax of Welsh*. Cambridge: Cambridge University Press.

Cinque G (1999) *Adverbs and Functional Heads: A Crosslinguistic Perspective*. Oxford: Oxford University Press.

Cook V (2016) Premises of multi-competence. In: Cook V and Wei L (eds.) *The Cambridge Handbook of Linguistic Multi-Competence*. Cambridge: Cambridge University Press, pp.1–25.

Deuchar M (2005) Congruence and Welsh-English code-switching. *Bilingualism: Language and Cognition* 8(3): 255–269.

King G (2003) *Modern Welsh: A Comprehensive Grammar* (2nd edn.). London, New York: Routledge.

Knight D, Morris S, Fitzpatrick T, et al. (2020) *CorCenCC: Corpws Cenedlaethol Cymraeg Cyfoes – The National Corpus of Contemporary Welsh (Version 1.0.0)*. Cardiff University. DOI: 10.17035/d.2020.0119878310

Lardiere D (2018) Establishing ultimate attainment in a particular second language grammar. In: Han Z and Odlin T (eds.) *Studies of Fossilization in Second Language Acquisition*. Bristol, Blueridge Summit: Multilingual Matters, pp.35–55.

Larsson T, Callies M, Hasselgård H, et al. (2020) Adverb placement in EFL academic writing: Going beyond syntactic transfer. *International Journal of Corpus Linguistics* 25(2): 155–184.

Osborne J (2008) Adverb placement in post-intermediate learner English: A contrastive study of learner corpora. In: Gilquin G, Papp S, and Díez-Bedmar MB (eds.) *Linking up Contrastive and Learner Corpus Research*. Amsterdam: Brill, pp.127–146.

Robert E (2009) Accommodating "new" speakers? An attitudinal investigation of L2 speakers of Welsh in southeast Wales. *International Journal of the Sociology of Language* 195(195): 93–115.

Roberts IG (2005) *Principles and Parameters in a VSO Language: A Case Study in Welsh*. Oxford: Oxford University Press.

Schifano N (2018) *Verb Movement in Romance: A Comparative Study*. Oxford: Oxford University Press.

Welsh Government (2020) Where and when people learn to speak Welsh. *National Survey for Wales, 2018-19* [online]. Statistics for Wales. Available at: https://gov.wales/sites/default/files/statistics-and-research/2020-04/where-and-when-people-learn-speak-welsh-national-survey-wales-april-2018-march-2019-182.pdf

White L (1991) Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research* 7(2): 133–161.

Williams CH (2014) The lightening veil: Language revitalization in Wales. *Review of Research in Education* 38(1): 252–272.

REVIEW ARTICLE

A bibliometric study of the research field of experimental philosophy of language

Jincai Li[1], Xiaozhen Zhu[2*]

[1] School of Foreign Languages, East China Normal University, Shanghai, China
[2] School of English Education, Guangdong University of Foreign Studies, Guangzhou, China

Abstract: The past eighteen years witnessed the rapid development of experimental philosophy of language. Adopting a bibliometric approach, this study examines the research trends and status quo of this burgeoning field based on a corpus of 237 publications retrieved from PhilPapers. It is observed that experimental philosophy of language has undergone three stages, the initiation stage, the development stage, and the extension stage, across which there is a clear upward trend in the annual number of publications. Michael Devitt, Edouard Machery, John Turri, Nat Hansen, et al., are found to be the most productive philosophers, testifying their leading positions in this field. Journals, instead of books, are the major homes of works in this area. The analysis also yields a list of influential works, including the seminal work "Semantics, Cross-cultural Style" and other significant publications on the semantics of various types of expressions. Relatedly, the major research themes are found to include not only intuitions about the reference of proper names, but also a wide array of philosophically and linguistically interesting issues like the meaning of color adjectives, epistemic modals, and predicates of personal taste, the norms of assertions and the essence of lies, etc. These findings showcase that experimental philosophy of language has broadened the research territory and offered deep insights into central issues of philosophy of language that are beyond the reach of the conventional armchair methodology.

Keywords: experimental philosophy of language; bibliometric analysis; research trend; research themes

***Corresponding author:** Xiaozhen Zhu, lecturer at School of English Education, Guangdong University of Foreign Studies, Guangzhou 510006, Guangdong, China; zhuxiaozhen@oamail.gdufs.edu.cn

Received: May 23, 2022; **Accepted:** June 21, 2022; **Available online:** August 27, 2022

Citation: Li J and Zhu X (2022) A bibliometric study of the research field of experimental philosophy of language. *Forum for Linguistic Studies* 4(1): 18–35. DOI: 10.18063/fls.v4i1.1462

1. Introduction

Experimental philosophy of language, also interchangeably referred to as experimental semantics, is a newly emerged area in philosophy. Experimental philosophers of language typically adopt research methods commonly used in psychology and cognitive science to address philosophically

interesting issues related to language, such as meaning and reference of various linguistic expressions. If we take the publication of the short paper "Semantics, cross-cultural style" by Machery et al. (2004) as the herald of experimental philosophy of language movement, then this year 2022 marks the 18th year of its development. Young though it is, it has attracted increasing attention among philosophers, linguists and psycholinguists interested in language, and has matured into an influential and indispensable way of doing semantics and philosophy.

While the field of experimental philosophy of language is ever expanding, producing a growing body of scholarly output, there have been few studies that systematically review this burgeoning area. Hansen (2015) and Machery (2021) have each provided a survey of the studies in experimental philosophy of language, but they have nonetheless failed to offer a panoramic view of the whole field, primarily due to their narrow focus on the empirical investigations into the issue of reference of names, which in reality represents just one strand of the experimental research. As a result, it is still unclear, at least to those who are not familiar with this field, what the most commonly explored research topics are in experimental philosophy of language, who have been the most productive and influential researchers, which have been the most important works, and what the major venues for works in experimental philosophy of language are. It is also unknown whether there are changes in these aspects.

To better answer these questions, we turn our attention to bibliometrics, an approach that is originally developed and commonly used in library and information science. A bibliometric analysis is in essence a quantitative analysis that applies "mathematics and statistical methods to the analysis of academic publications" (Pritchard, 1969: 348). It can be employed to uncover research trends and publication patterns, and thus give researchers a broad overview of a certain academic field, such as the number of publications within a certain time period, the productivity of individual authors and institutions, the publishing capacity of academic publishers and journals, and the influence of researchers and published works (de Bellis, 2009). When used in combination with machine learning and natural language processing techniques, bibliometrics can further offer deeper insights into the major research themes and their evolution over a specified time window, and hence delineate the boundary and the developmental trajectory of the whole field in an objective manner. Over the years, bibliometrics has been widely used in different disciplines, such as computer science (Xie and Willett, 2013), business management (Liu et al., 2015), education (Chang et al., 2020), and linguistics (Lei and Liu, 2019; Zhang, 2020; Hyland and Jiang, 2021, etc.). But, to the best of our knowledge, there have been no such studies even in the broad field of philosophy, let alone its sub-field experimental philosophy of language.

There are several studies in philosophy that have a savor of bibliometric analysis, which might be considered as primitive form of bibliometric studies in this field. For instance, Andow (2015) investigated the explosion of the word "intuition" and its cognates based on articles retrieved from JSTOR database. It was observed that the intuition talk has been prevalent since the 1900s and this phenomenon is not confined to the field of philosophy. Other academic areas of finance, marketing, linguistics, economics, etc., have all witnessed the surge of intuition talks. But philosophers in particular seemed to have enjoyed using intuitions in their works the most. In this way, Andow (2015) provided evidence of the widespread use of intuitions in academic research and hence laid ground for further studies on the conventional armchair theorizing methodology. However, with a focus on the use of intuitions, the study cannot inform us of the other hot research topics in philosophy in the

examined time span.

Another case in point is Knobe (2015), in which around 400 influential publications in philosophy of mind from 20 reputable philosophy journals in the periods of 1960–1999 and 2009–2013 were examined to see what contemporary philosophers are actually doing. Knobe found that works in these time frames are quite different in terms of both research topics and methods. While the studies in the 20th century mostly adopted the pure a priori armchair-based method, the latest contemporary works have widely appealed to empirical approaches. With this shift in methodology, there is also a shift of attention from broad metaphysical questions of mind to more specific and cognition-related aspects of mind. Knobe thus concluded that philosophers nowadays are doing things substantially different from their predecessors and philosophy has evolved to be an interdisciplinary research area. Still another empirical bibliometric-like study is Knobe (2016). In this study, Knobe offered a quantitative analysis of 379 papers documenting 453 experimental studies indexed under the category "experimental philosophy" in the PhilPapers Database. Through manual examination and categorization, Knobe found that a vast majority of these studies (88%) do not belong to either the positive or negative research program in experimental philosophy. Instead, they are best construed as falling into the broad area of cognitive science. This finding has led Knobe to claim that experimental philosophy is part of cognitive science. These two quantitative studies conducted by Knobe (2015, 2016) provided valuable data on two subfields of contemporary research in philosophy. But owing to constraints in scope and methods, these studies have not provided a comprehensive review of the relevant areas.

Thus, this present study aims to fill the niche by carrying out a bibliometric analysis of the research outputs in the specific area of experimental philosophy of language. The questions guiding this project include the following:

1) What is the overall publication trend in experimental philosophy of language?
2) What are the popular venues for works in experimental philosophy of language?
3) Which philosophers and documents in experimental philosophy of language have been most influential?
4) What have been the major research topics in experimental philosophy of language?

In the rest of this paper, we will first introduce the research methods in detail, including the process of data collection and data cleansing in Section 2. In Section 3 we will show the major findings, with statistical analysis and brief discussions of the results, structured according to the research questions listed above. Section 4 concludes the study with predictions and implications for future bibliometric research in experimental philosophy of language.

2. Methodology

2.1. Corpus: Experimental philosophy of language

The first step of the bibliometric analysis is to create a corpus of the research outputs in experimental philosophy of language. But how do we delimit experimental philosophy of language? For the current purpose, we roughly follow the definition offered by Hansen (2015: 1):

> Experimental philosophy of language applies experimental methods used in the cognitive sciences (experimental psychology, psycholinguistics) to topics of interest to philosophers of language, such as the meaning of particular kinds of expressions (names, determiners, natural kind terms, adjectives, and so on), pragmatic phenomena (implicature, presupposition, metaphor, the semantics-pragmatics boundary, for example), and methodological issues (the reliability of informal versus formal experimental methods, the reliability of expert judgments versus the judgments of ordinary speakers, for example).

However, we exclude from the current study the empirical research on pragmatic phenomena (implicature, presupposition, metaphor, the semantics-pragmatics boundary, etc.), as these phenomena are typically described as being the core part of "experimental pragmatics" (Noveck, 2018) and are mostly investigated by researchers in the field of linguistics. Here, it is also worthwhile to point out that in experimental studies on the semantics-pragmatics boundary/interface issue, researchers also use the term "experimental semantics" as a coordinate of "experimental pragmatics" (see the book titles *Experimental Pragmatics/Semantics* by Meibauer and Steinbach, and *The Oxford Handbook of Experimental Semantics and Pragmatics* by Cummins and Katsos. However, there is an anthropological difference to be noted here. The studies to this type of experimental semantics differs from what we focus on in the current research, in that while these studies are mostly carried out by linguists or psycholinguists, what we are concerned about are principally done by philosophers.

Having set the boundary, we next sampled the scholarly works indexed on the PhilPapers website (https://philpapers.org/). We targeted this resource pool instead of the commonly used databases like Google Scholar and Web of Science, because thus far it has been the largest and the most comprehensive database of philosophy maintained by professional philosophers. Further, the works indexed are all categorized under different categories and subcategories based on the central topics. For instance, experimental philosophy is placed under philosophical methods section, which is under metaphilosophy. To obtain the relevant literature, we located the subcategory of "experimental philosophy of language" that is among the 13 subcategories of experimental philosophy. As of May 16, 2022, there are 209 entries in this category. We retrieved the bibliometric data of these works, including title, author, publication year, document type, source, abstract, keywords (if there are) and citation numbers.

2.2. Data pre-processing

We next pre-processed the data for further bibliometric analysis. As a first step, duplicate entries, book reviews, book introductions, conference presentations, short commentaries and those that were published before 2004 were manually removed. In the meanwhile, we also added some scholarly works that in our opinion clearly belong to experimental philosophy of language but are nonetheless missing in the index. This resulted in a total number of 237 publications in our corpus spanning the period of 2004–2021. Among these works, there are 9 books, 35 book chapters, 193 journal articles, suggesting journals are the most popular medium of disseminating knowledge in experimental philosophy of language.

In addition, while conducting manual check of these works, we supplemented the core bibliometric information such as publication year, abstract and keywords for works that do contain such data but were somehow incomplete as indexed on the PhilPapers website. We also corrected the citation counts based on Google Scholar (as of May 15, 2022), which is up-to-date and comparatively

speaking more accurate. In addition, the author names were re-coded in case the same researcher's name is listed differently (e.g., Devitt, M, Michael Devitt, J. Knobe and Joshua Knobe, Nathaniel Hansen, Nat Hansen, Hansen, N.) in the literature. The journal titles that were abbreviated in some works were also changed into the full title for easy referencing and analysis.

3. Results and analysis

In this section, we report the results and analysis of our bibliometric study, which includes: 1) the number of publications in experimental philosophy of philosophy across the past 18 years, divided into three 6-year periods, 2) the most productive researchers in experimental philosophy of language, 3) the major publication venues of works in experimental philosophy of language, 4) the main research themes in experimental philosophy of language over the last two decades.

3.1. Number of publications over the past 18 years

What is the developmental trajectory of experimental philosophy of language like? **Figure 1** below presents the annual number of publications over the years. It is evident that despite several falls, the number of publications in general has kept rising over the last two decades, rendering experimental philosophy of language a burgeoning field. In particular, since 2015, the average number of yearly research outputs has stayed around 22, which is almost triple of the average annual number of publications (i.e.,7.5) in the previous 11 years.

When we divide the past 18 years into three 6-year periods, the increasing trend across the three periods is easily observable. As shown in **Figure 2**, in the first period there are only 15 publications, while in the last period the number has risen to 133. This sharp increase in the number of publications suggests that experimental philosophy of language as a young and developing sub-discipline has attracted greater and greater attention over the examined time span.

3.2. Most productive philosophers of language

In the field of experimental philosophy of language, who have been the leading and driving forces? To answer this question, we looked at the productivity of researchers in this area. As shown in **Table 1**, Michael Devitt, Edouard Machery, John Turri and Nat Hansen are the top five most prolific philosophers of language who have authored or co-authored more than 10 works in the last two

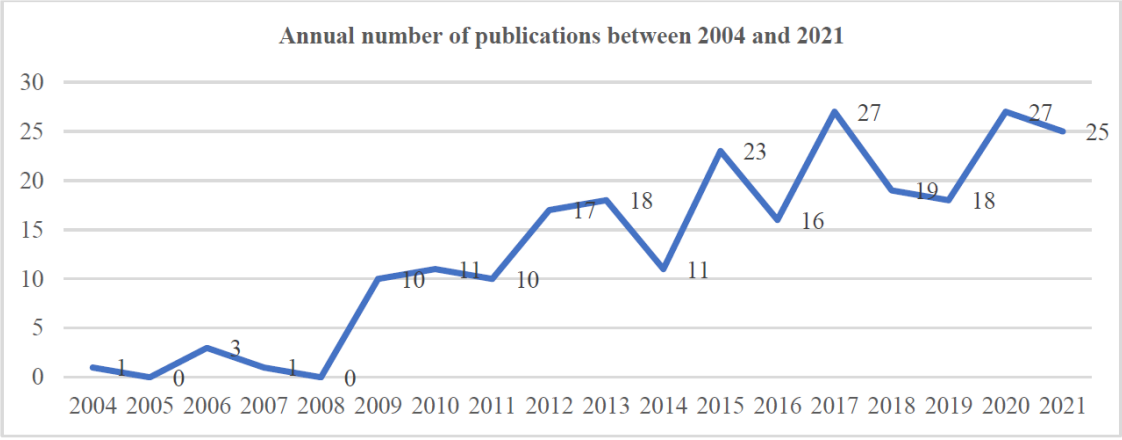


Figure 1. Annual number of publications in experimental philosophy of language between 2004 and 2021.

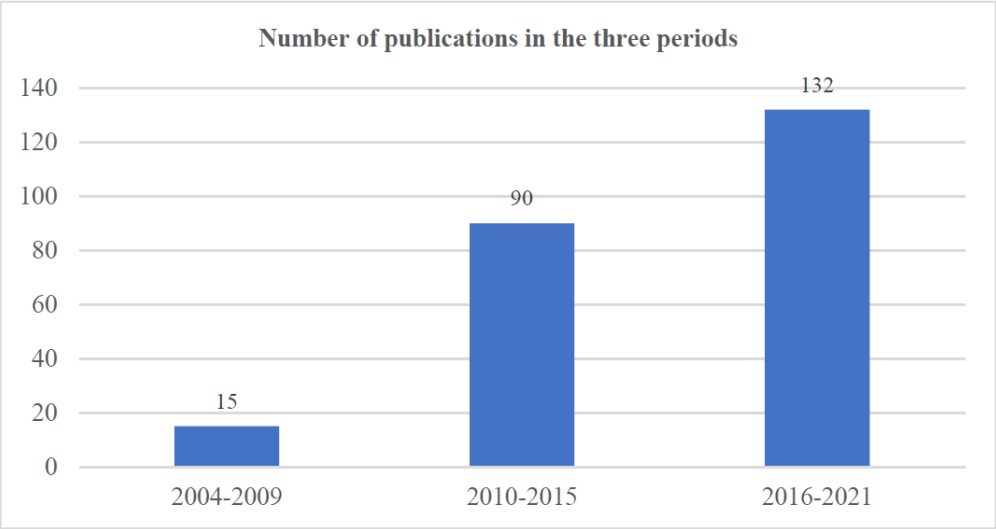


Figure 2. Number of publications in experimental philosophy of language in the three periods.

Table 1. The top 10 productive philosophers of language

	As sole author or co-author	Total number of publications
Michael Devitt	13	13
Edouard Machery	11	12
John Turri	9	14
Nat Hansen	9	12
Eugen Fischer	7	7
Genoveva Marti	6	6
Emmanuel Chemla	5	9
Justin Khoo	5	5
Markus Kneer	5	5
Daniel Cohnitz	4	4
Jeffrey Maynes	4	4

decades. In particular, Michael Devitt has published 13 works as the sole author or the first author on topics of linguistic intuitions and their evidential value, issues in experimental semantics, testing theories of reference of names (including proper names and natural kind terms), etc. Edouard Machery, one of the leading figures in experimental philosophy movement, produced 12 works in philosophy of language, with a special focus on the reference of proper names. His work "Semantics, Cross-Cultural Style" (2004), has remained the most influential and provocative study, replications and criticism of which have not ceased to appear. It is also this research that initiated the experimental turn of philosophy of language (Li and Liu, 2015). Intriguingly, over the years, we have seen heated debates on testing reference of proper names between Edouard Machery, Michael Devitt and Genoveva Marti, which has no doubt facilitated the development of experimental philosophy of language.

In the last decade, Nat Hansen has reviewed the experimental philosophy of language movement,

conducted experiments on contextualism and also examined issues and claims in ordinary language philosophy. John Turri's research mostly touches the norms of assertion, moral judgements, lying, etc. More recently, Eugen Fischer has published works concerning topics like intuitions and inferences, conceptual engineering, and more generally, ways of doing experimental philosophy (of language). Also in recent years, Markus Kneer has conducted several studies on belief ascriptions, predicate of personal tastes, epistemic modals, assertions, debates between contextualism and relativism. Similarly, Justin Khoo explored the issues concerning modal disagreements, epistemic modals, indexicals, moral semantics, etc. Emmanuel Chemla differs from the above philosophers in that her works on presuppositions and scalar implicatures, the interface between semantics and pragmatics, etc., are most often discussed and cited in the linguistic field. Nonetheless, these topics also have a philosophical ring in themselves, which is why they were indexed in the experimental philosophy of language category in PhilPapers.

3.3. Major publication venues

People who are interested in the experimental philosophy of language movement may be eager to know what the popular venues that host the emerging works in the field are. In our database, the 237 works includes 9 books, 35 book chapters and 193 journal articles, indicating journals are the major homes of scholarly works in experimental philosophy of language. Among these publications, the books are mostly quite young, which mostly appeared after the year 2015. These new books include *Advances in Experimental Philosophy of Language* (Haukioja, 2015), *Puzzles of Reference* (Cappelen and Dever, 2018), *The Oxford Handbook of Experimental Semantics and Pragmatics* (Cummins and Katsos, 2019), *Experimental Philosophy: A Critical Study* (Mukerji, 2019), *Linguistic Intuitions: Evidence and Method* (Schindler et al., 2020), and *The Routledge Handbook of Linguistic Reference* (Biggs and Geirsson, 2021). Similarly, most of the book chapters also came out in the last ten years. Their sources, besides the above-mentioned edited book volumes, include *Current Controversies in Experimental Philosophy* (O'Neill and Machery, 2014), *On Reference* (Bianchi, 2015), *Advances in Experimental Philosophy and Philosophical Methodology* (Nado, 2016), *A Companion to Experimental Philosophy* (Sytsma and Buckwalter, 2016), *Methodological Advances in Experimental Philosophy* (Fischer and Curtis, 2019), and *Language and Reality from a Naturalistic Perspective* (Bianchi, 2020). These recent books and book chapters, though small in number, present the status quo and the advances of experimental philosophy of language in a rather systematic and coherent manner, making important contributions to the filed.

Different from the recent books and book chapters, journal articles started to appear early in around the beginning of this century, presenting the most recent research of the time. What are the major journals that host these articles? **Table 2** presents the top 15 venues that have published research in experimental philosophy of language. Most notably, *Synthese*, *Cognition*, *Review of Philosophy and Psychology*, *Mind and Language*, *Philosophical Psychology* have been the most popular venues, probably because these journals enjoy a wider readership among philosophers, linguists, and psychologists as well. Indeed, the rest of the journals in the list such as *Cognitive Science*, *Philosophy and Phenomenological Research*, and *Philosophy Compass* are also interdisciplinary to varying degree, rather than being confined to a narrow and specific discipline. It seems natural that these journals form the logical homes of works in experimental philosophy of language, primarily due to their interdisciplinary nature, i.e., employing methods commonly used in psychology and cognitive science to address issues in philosophy of language that are of interest to philosophers,

Table 2. The most popular venues of works of experimental philosophy of language

No.	Journal title	No. of publications
1	*Synthese*	15
2	*Cognition*	14
3	*Review of Philosophy and Psychology*	13
4	*Mind and Language*	12
5	*Philosophical Psychology*	9
6	*Philosophy and Phenomenological Research*	7
7	*Philosophy Compass*	7
8	*Journal of Semantics*	6
9	*Australasian Journal of Philosophy*	5
10	*Erkenntnis*	5
11	*Inquiry*	5
12	*Semantics and Pragmatics*	5
13	*Cognitive Science*	4
14	*Philosophical Studies*	4
15	*Analysis*	4

linguists and psychologists.

3.4. Influential works

How to gauge the impact of individual academic research output? Citation counts are always employed as a crucial means to measure the significance of a publication. However, while raw citation numbers are intuitive, they are less informative and meaningful when we compare works that appear in different time periods. For instance, a paper published in the last two years is less likely to have as high a citation count as those appeared fifteen years ago. But the relatively low count of the latest work does not mean it is not important. In bibliometric studies, a common strategy is to retrieve all the references cited in the works being included in the corpora and normalize the frequency of each of them based on their raw frequencies and the total number of publications during the time span under investigation. On the basis of these data, researchers can also compare the highly cited work in any period of interest in the development of the field, which may help reveal the diachronic changes over the years. In the current study, however, we failed to retrieve the references in each of the works included in our database from the PhilPapers website. As a remedy, in order to control for the time effect, we compared the citations (retrieved from Google Scholar) of works of similar age, namely those that came out during the same time period. We divided the past eighteen years roughly into five windows, namely 2004–2007, 2009–2011, 2012–2014, 2015–2017, 2018–2021, because there are only 6 publications in the first period (and there is none in the year 2008), and then normalized the frequencies within each time window. Below we present the list of the top 20 most influential works in experimental philosophy of language in terms of their relative citation rate (shortened as RCR), which is computed according to the following formula proposed by Li and Lei (2019):

$$\text{Relative citation rate (RCR)} = \frac{\text{Observed citation counts (OCC)}}{\text{Expected citation counts (ECC)}}$$

Here, OCC is the raw citation counts of a given research output, whereas ECC is the expected citations of any research output in the time period it was published. For example, the years between 2009 and 2011 have seen the publication of 32 pieces of works which were cited 2,345 times in total, then for any paper published in this period the expected citation count is approximately 73. Thus, if a paper published in 2010 has actually been cited 182 times thus far, then the relative citation rate of this paper will be 2.48. We think this normalized citation rate can be used as a rough measure to compare the impact of works across the time span we investigated, while controlling for the effect of the number of years an article has been published. Nonetheless, the list below in **Table 3** should not be read as an absolute rank of the influence of the publications.

Several interesting observations can be made from the table. First, among the 20 impactful publications, only one of them is a book chapter, while the rest are all journal articles, suggesting journal articles are the primary sources of references in experimental studies in philosophy of language. Second, judging from the publication years, half of the influential works appeared during the latest period of 2018–2021. This makes sense as with the advances of experimental philosophy of language the recent works may offer a better review of the prior literature and deeper insights into the issues under consideration. Third, the influential studies center on several hot topics in experimental philosophy of language. For instance, on the reference of proper names, the influential publications include the seminal work "Semantics, cross-cultural style" by Machery et al. (2004) and those in the response literature like "Reference in the land of the rising sun: A cross-cultural study on the reference of proper names" (Sytsma et al., 2015), and "Speaker's reference and cross-cultural semantics" (Machery et al., 2015), and more recently, Devitt and Porot's work "The reference of proper names: Testing usage and intuitions" (2018).

Some studies focus on concepts, normality and assertions, such as "Dual character concepts" (Reuter, 2019), "Moral disagreement and moral semantics" (Khoo and Knobe, 2016), "Normality: Part descriptive, part prescriptive" (Bear and Knobe, 2017), "Dual character concepts and the normative dimension of conceptual representation" (Knobe et al., 2013), and "The norm of assertion: Empirical data" (Kneer, 2018). Others deal with general problems in experimental philosophy (of language), like "Remarks on the experimental turn in the study of scalar implicature, Part 1" (Chemla and Singh, 2014), "Experimental ordinary language philosophy: A cross-linguistic study of defeasible default inferences" (Fischer et al., 2019), and "Carnapian explications, experimental philosophy, and fruitful concepts" (Koch, 2019). There are also significant studies on the semantics of particular types of expressions like epistemic modals, quantifiers, adjectives, and appositives, such as "Modal disagreements" (Khoo, 2015), "Must, knowledge, and (in)directness" (Lassiter, 2016), "Might do better: Flexible relativism and the QUD" (Beddor and Egan, 2018), "The semantics of many, much, few, and little" (Rett, 2018), "Evaluational adjectives" (Silk, 2021), and "Experimental evidence for the truth conditional contribution and shifting information status of appositives" (Syrett and Koev, 2015). Finally, it is notable that some philosophers like Edouard Machery, Joshua Knobe have appeared several times in the influential works, showcasing their leading position in the experimental philosophy movement.

3.5. Major research themes

What are the major research themes in experimental philosophy of language? In this section,

Table 3. Top 20 influential works in experimental philosophy of language

Title	Author	Year	Source	OCC	RCR
Normality: Part descriptive, part prescriptive	Bear and Knobe	2017	*Cognition*	127	6.38
The semantics of many, much, few, and little	Rett	2018	*Language and Linguistics Compass*	38	6.18
David Lewis in the lab: Experimental results on the emergence of meaning	Bruner et al.	2018	*Synthese*	35	5.70
The reference of proper names: Testing usage and intuitions	Devitt and Porot	2018	*Cognitive Science*	30	4.88
Experimental evidence for the truth conditional contribution and shifting information status of appositives	Syrett and Koev	2015	*Journal of Semantics*	95	4.77
Remarks on the experimental turn in the study of scalar implicature, Part I	Chemla and Singh	2014	*Language and Linguistics Compass*	95	4.77
Are false implicatures lies? An empirical investigation	Weissman and Terkourafi	2019	*Mind and Language*	28	4.56
Moral disagreement and moral semantics	Khoo and Knobe	2016	*Noûs*	87	4.37
Semantics, cross-cultural style	Machery et al.	2004	*Cognition*	723	4.33
Dual character concepts	Reuter	2019	*Philosophy Compass*	26	4.23
Must, knowledge, and (in)directness	Lassiter	2016	*Natural Language Semantics*	83	4.17
The norm of assertion: Empirical data	Kneer	2018	*Cognition*	25	4.07
Carnapian explications, experimental philosophy, and fruitful concepts	Koch	2019	*Inquiry*	22	3.58
Speaker's reference and cross-cultural semantics	Machery et al.	2015	In: Bianchi A (ed.) *On Reference*	63	3.16
Might do better: Flexible relativism and the QUD	Beddor and Egan	2018	*Semantics and Pragmatics*	18	2.93
Modal disagreements	Khoo	2015	*Inquiry*	57	2.86
Experimental ordinary language philosophy: A cross-linguistic study of defeasible default inferences	Fischer et al.	2019	*Synthese*	17	2.77
Dual character concepts and the normative dimension of conceptual representation	Knobe et al.	2013	*Cognition*	133	2.69
Reference in the land of the rising sun: A cross-cultural study on the reference of proper names	Sytsma et al.	2015	*Review of Philosophy and Psychology*	53	2.66
Evaluational adjectives	Silk	2021	*Philosophy and Phenomenological Research*	16	2.60

we first analyzed the keywords provided by authors in some of the works. Since some articles published in journals like *Review of Philosophy and Psychology*, *Analysis*, *Philosophy Compass*, and *Mind and Language*, do not contain keywords, we also analyzed the texts of all the abstracts of the journal articles and book chapters as well as book introductions with the help of natural language

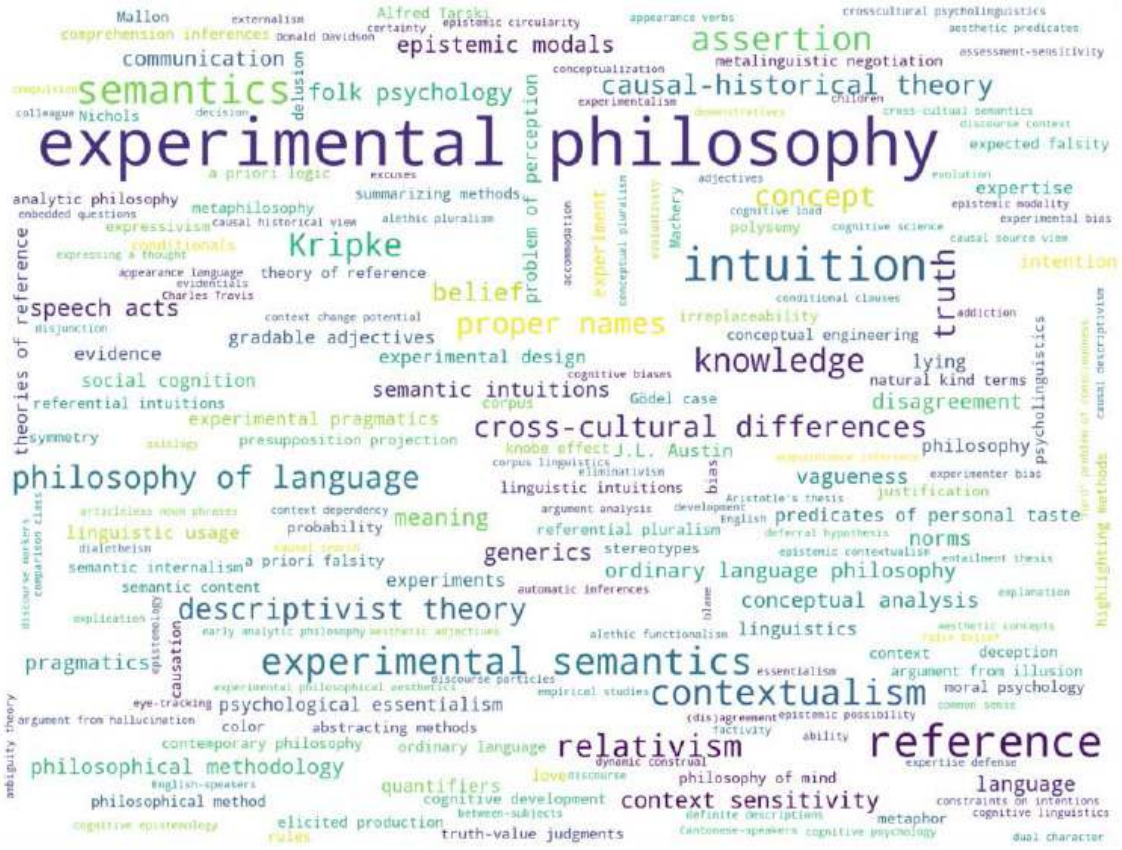


Figure 3. Word cloud based on the keywords provided by authors.

processing techniques. Specifically, we first cleaned the keywords list and the texts of abstract using the programming language Python script. They were then lemmatized[1] via spaCy, which is an open-source library designed to help build natural language processing applications. In the next step, we calculated the frequency of the lemmatized keywords and ranked them in order. Manual checks were then conducted to consolidate the different key terms that in essence denote the same concept. For example, we manually changed all instances of "causal-historical", "causal historical", "causal view of reference", "causal theory of reference" into the oft-used label "causal-historical theory of reference". Similarly, all tokens of "descriptivism", "description theory of names" and "descriptivist theory of names" were modified into "descriptivist theory of names"; and "cultural difference", "(cross) cultural variation" were changed into "cross-cultural differences". After applying these procedures, a list of 384 keywords were produced, a word cloud representation of which is displayed above in **Figure 3**.

With regard to the abstracts, we extracted the n-grams in the lemmatized texts, which includes monograms, 2-grams, 3-grams, and 4-grams. For mono grams, only nouns and adjectives were studied while all the other words were removed from the 1-gram list because they are unlikely to be used as research themes. Also included in the procedures was a step to remove all the functional words like propositions, articles, and auxiliary verbs, using a stop word list as these words rarely

[1] Lemmatization is a commonly used pre-processing technique in natural language processing that can change any form of a word into its base form. For example, forms like "goes", "going", "went" can all be reduced to the base "go" for further analysis and computation through a simple lemmatization process.

occur in keywords. Subsequently, to get an insight from the data about the hot research themes in experimental philosophy of language, we contemplated the frequency of the n-grams and nouns chunks for them to be counted as meaningful topics. Initially, the words and word strings that occur more than 10 times were selected, which resulted in a list of 505 words and phrases. However, quite a number of these strings, though very frequent, were either not meaningful topics (like *proper*, *paper presents*, *reference proper*, *intuition reference proper*) or just too broad (like *study*, *philosophy*, *language*, *theory*, and *case*) and hence were manually removed. This procedure yielded a list of 60 meaningful concepts (see Appendix for the full list).

In **Table 4**, we list the top 20 research themes from both the author-supplied keyword list and the n-grams extracted from the abstracts, which also feature prominently in the word cloud in Figure 3. It is evident in Table 4 that the umbrella terms "experimental philosophy" and "experimental semantics" are most often used in the literature, as these are the central concepts that researchers cannot possibly eschew when they are introducing the research background, discussing their own empirical findings, and drawing implications for the development of the field. What is also easily noticeable from both columns of themes in the table is that "intuition" and "reference" are among the most popularly discussed topics in experimental philosophy of language. Also present in the list are terms like "cross-cultural difference", "descriptivist", "causal-historical theory", "theory of reference" as

Table 4. The top 20 research themes based on the n-grams and the keywords supplied by authors

Keywords		**N-grams**	
Theme	**Frequency**	**Theme**	**Frequency**
experimental philosophy	51	intuition	251
intuition	20	reference	193
reference	18	experiment	172
experimental semantics	13	participant	111
contextualism	12	truth	92
assertion	11	judgment	81
knowledge	10	Machery	74
Kripke	10	descriptivist	74
relativism	9	context	71
truth	9	assertion	65
concept	8	theory (of) reference	60
cross-cultural differences	8	knowledge	59
descriptivist theory	8	adjective	49
proper names	8	epistemic	48
causal-historical theory	7	Kripke	48
belief	6	experimental study	39
context sensitivity	5	cross-cultural difference	35
epistemic modals	5	epistemic modal	34
folk psychology	5	predicate	34
generics	5	presupposition	34
meaning	5	causal-historical theory	34

well as person proper names like Kripke and Machery.

The presence of these words as major research themes come as no surprise, given the fact that the whole movement of experimental philosophy of language was first initiated by the seminal work "Semantics, cross-cultural style" of Machery et al. (2004) which targeted the armchair theorizing methodology that Saul Kripke appealed to in advancing his causal-historical view of reference of names, and has subsequently been propelled by the growing body of response literature that center around the issues in the original study of Machery et al. (2004). In this expanding pool of literature on the reference of proper names, methodological issues are also widely discussed, hence besides words like "experiment", "participants" that are unavoidable in experimental studies, key terms like "truth", "judgment" are also very frequent as they may have occurred in expressions like "truth-value judgment" and "intuitive judgment"[2].

Besides the groups of works on the reference of proper names, studies in experimental philosophy of language also paid much attention to debates on contextualism and other opposing camps like relativism. For instance, in the last decade, a growing number of experiments have been conducted on epistemic modals, color adjectives, and predicates of personal taste, etc. Notably, Hansen and Chemla (2013) critically discussed the design of *the context shifting experiment* which has been employed as the central method of amassing evidence for contextualism. Based on their revised and improved experimental design, Hansen and Chemla (2013) tested a series of scenarios concerning knowledge ascriptions and color adjectives, providing insights into the impact of changing contexts on the evaluation of sentences containing the philosophically interesting terms like "know" and "green" and the critical features of the design and implementation of philosophical thought experiments and quantitative questionnaires. In several other works, based on experimental results, Hansen and Chemla (2017) as well as Adams and Hansen (2020) furthered our understanding of the nature and essence of color adjectives as well as the debates on context sensitivity. In addition to works on color adjectives, recent years have also seen the publication of research on epistemic modals. Just as mentioned in the list of influential works in Table 3, Beddor and Egan (2018) reported a set of experimental findings on epistemic modals that better support the flexible forms of relativism than contextualism. Knobe and Yalcin (2014) and Khoo and Phillips (2019) have also documented empirical results on epistemic modals that help adjudicate on the disputes surrounding contextualism. In more recent years, Markus Kneer also joined the heated debates between contextualism and relativism. Specifically, Kneer (2021a, 2022) provided large sets of data on perspectival claims including epistemic modals and predicate of personal taste, adding evidence in favor of the contextualist claims rather than relativism.

What is also attracting increasing attention lately is the issue of the norms of assertions. In particular, John Turri, one of the most productive experimental philosophers of language, has produced a series of empirical studies on the proper conditions under which assertions can be legitimately made. For example, Turri (2013) presented experimental evidence from six studies supporting the factive norms of assertion, and subsequently Turri (2016a, 2016b, 2016c, 2017, 2021) offered ample empirical data in favor of the knowledge account of assertions and compared the respective impact of knowledge and certainty on assertability. Reuter and Brössel (2019), however, challenged the knowledge account by presenting a set of evidence showing that the norm of assertion is justified

[2] The word "judgement" could also be used alone the in experimental studies as a synonym to "response" or "answer".

belief, with truth or even knowledge not being required in making assertions. Marsili and Wiegmann (2021) posed further challenges on the factive norms of assertion by questioning the common assumptions that participants' judgements about what an agent "should say" are evidence of their intuitions about assertability. Marsili and Wiegmann (2021) developed versions of stimuli that are more likely to be interpreted as intended by researchers, then experimentally tested the validity of these new measures, and offered evidence for the non-factive accounts of assertion. Differing from these above findings, Kneer (2018) reported data that indicate that knowledge cannot predict assertability reliably and the factive constraints seem unwarranted either. More recently, in a large-scale cross-cultural study in the United States, Germany, and Japan, Kneer (2021b) found that the speakers are expected to carry the epistemic responsibility of the speakers in asserting certain claims.

Closely related to assertions, lies which are regarded as dishonest assertions have also gained much attention in experimental philosophy of language in recent years. Turri and Turri (2016) assessed the common view that to tell a lie is to make false assertion in order to deceive others in a series of behavioral experiments. What they found is that attributions of lies are subjected to the influence of audience uptake and are based on attributions of assertion rather than attribution of deceptive intentions. Similarly, Marsili (2016) and Reins and Wiegmann (2021) have both critically discussed the traditional definitions of lies and tested their proposals against ordinary speakers' intuitions, thus offered alternative accounts of lies. Undoubtedly, these above-mentioned empirical studies in experimental philosophy of language have greatly advanced our knowledge and understanding of the various linguistic phenomena concerned.

4. General discussions

This bibliometric study of the movement of experimental philosophy of language has produced several crucial findings that carry important implications for future research. First, it is observed that there is an overall growing trend of publications in this burgeoning field over the past 18 years. A close look at the yearly productivity suggests that the developmental pathway of experimental philosophy of language can be characterized as consisting of three stages: the *initiation* period, the *development* period and the *extension* period. Secondly, this bibliometric research produces lists of the most productive researchers, the most popular publication venues and the most impactful works in experimental philosophy of language, which, we believe, are instrumental in promoting the movement and facilitating the advancement of the field. Thirdly, we find that the territory of experimental philosophy is fast expanding, with more and more philosophically and linguistically interesting phenomena being investigated empirically.

At its early stage, experimental philosophy of language is mostly concerned with the reference of proper names, or more specifically, the cross-cultural differences in folks' intuitions about the reference of proper names. During the development stage, while studies on reference of proper names and theories of reference kept coming out, research outputs on color adjectives, epistemic modals, predicates of personal taste, etc., begin to sprout up. And lately in the extension period, while these issues incur even more heated debates, the normative issue of assertions and lies also grabs a lot of attention. It thus becomes evident that there is a gradual shift away from the narrow focus on reference of proper names and intuitions as well as the meta-philosophical discussions of them. For this reason, the research projects in experimental philosophy of language can no longer be simply

judged as negative programs in experimental philosophy. Instead, they form a very active part of the positive programs and are even being broadened into the grand enterprise of cognitive science. Given the productivity and the impact of the research in experimental philosophy of language, we are optimistic that there is absolutely more to come in this booming area.

These having said, informative and illuminating as it is, the current study is also constrained in terms of the corpus and the analytic techniques. In particular, due to the small size of the database and the availability of certain types of bibliometric data (e.g., the references in each of the works in the database), we did not perform statistical analysis on the datasets, nor were we able to compare the data diachronically across the three time periods. Hence, the results and the conclusions might be compromised to some extent. Future studies should try to amass a larger body of bibliometric data of studies in experimental philosophy of language and carry out rigorous statical analysis to better delineate the development and the status quo of this exciting and rapidly developing research field.

Conflict of interest

No conflict of interest was declared by the authors.

Acknowledgements

This work was supported by funding from Shanghai Planning Office of Philosophy and Social Science (2019EZX006), and China Postdoctoral Science Foundation (2019M651444).

References

Adams Z and Hansen N (2020) The myth of the commonsense conception of color. In: Wikforss A and Marques T (eds.) *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability*. Oxford: Oxford University Press, pp.106–127.

Andow J (2015) How "intuition" exploded. *Metaphilosophy* 46(2): 189–212.

Bear A and Knobe J (2017) Normality: Part descriptive, part prescriptive. *Cognition* 167: 25–37.

Beddor B and Egan A (2018) Might do better: Flexible relativism and the QUD. *Semantics and Pragmatics* 11: 7.

Bianchi A (ed.) (2015) *On Reference.* Oxford: Oxford University Press.

Bianchi A (ed.) (2020) *Language and Reality from a Naturalistic Perspective: Themes from Michael Devitt*. Cham: Springer International Publishing.

Biggs S and Geirsson H (eds.) (2021) *The Routledge Handbook of Linguistic Reference*. New York: Routledge, Taylor and Francis Group.

Bruner J, O'Connor C, Rubin H, et al. (2018) David Lewis in the lab: Experimental results on the emergence of meaning. *Synthese* 195(2): 603–621.

Cappelen H and Dever J (2018) *Puzzles of Reference* (1st edn.). Oxford: Oxford University Press.

Chang H-Y, Lin T-J, Lee M-H, et al. (2020) A systematic review of trends and findings in research employing drawing assessment in science education. *Studies in Science Education* 56(1): 77–110.

Chemla E and Singh R (2014) Remarks on the experimental turn in the study of scalar implicature, Part I. *Language and Linguistics Compass* 8(9): 373–386.

Cummins C and Katsos N (eds.) (2019) *The Oxford Handbook of Experimental Semantics and Pragmatics*. Oxford: Oxford University Press.

De Bellis N (2009) *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetics*. Mary-

land: Scarecrow Press.
Devitt M and Porot N (2018) The reference of proper names: Testing usage and intuitions. *Cognitive Science* 42(5): 1552–1585.
Fischer E, Engelhardt PE, Horvath J, et al. (2019) Experimental ordinary language philosophy: A cross-linguistic study of defeasible default inferences. *Synthese* 198: 1029–1070.
Fischer E and Curtis M (eds.) (2019) *Methodological Advances in Experimental Philosophy*. London: Bloomsbury Academic.
Hansen N (2015) Experimental philosophy of language. *Oxford Handbooks Online*. DOI: 10.1093/oxfordhb/9780199935314.013.53
Hansen N and Chemla E (2013) Experimenting on contextualism. *Mind & Language* 28(3): 286–321.
Hansen N and Chemla E (2017) Color adjectives, standards, and thresholds: An experimental investigation. *Linguistics and Philosophy* 40(3): 239–278.
Haukioja J (ed.) (2015) *Advances in Experimental Philosophy of Language*. London: Bloomsbury Academic.
Hyland K and Jiang F (2021) A bibliometric study of EAP research: Who is doing what, where and when? *Journal of English for Academic Purposes* 49: 100929.
Khoo J and Phillips J (2019) New horizons for a theory of epistemic modals. *Australasian Journal of Philosophy* 97(2): 309–324.
Khoo J (2015) Modal disagreements. *Inquiry* 58(5): 511–534.
Khoo J and Knobe J (2016) Moral disagreement and moral semantics. *Noûs* 52(1): 109–143.
Kneer M (2018) The norm of assertion: Empirical data. *Cognition* 177: 165–171.
Kneer M (2021a) Predicates of personal taste: Empirical data. *Synthese* 199(3–4): 6455–6471.
Kneer M (2021b) Norms of assertion in the United States, Germany, and Japan. *Proceedings of the National Academy of Sciences* 118(37): e2105365118.
Kneer M (2022) Contextualism versus relativism: More empirical data. In: Wyatt J, Zakkou J, and Zeman D (eds.) *Perspectives on Taste*. New York: Routledge, pp.109–140.
Knobe J (2015) Philosophers are doing something different now: Quantitative data. *Cognition* 135: 36–38.
Knobe J (2016) Experimental philosophy is cognitive science. In: Sytsma J and Buckwalter W (eds.) *A Companion to Experimental Philosophy*. West Sussex: John Wiley & Sons, Ltd., pp.37–52.
Knobe J, Prasada S, and Newman GE (2013) Dual character concepts and the normative dimension of conceptual representation. *Cognition* 127(2): 242–257.
Knobe J and Yalcin S (2014) Epistemic modals and context: Experimental data. *Semantics and Pragmatics* 7.
Koch S (2019) Carnapian explications, experimental philosophy, and fruitful concepts. *Inquiry* 62(6): 700–717.
Lassiter D (2016) Must, knowledge, and (in)directness. *Natural Language Semantics* 24(2): 117–163.
Li J and Liu L (2015) Shiyan yuyan zhexue: Gegu dingxin yihuo chenchen xiangyin [Experimental philosophy of language: Radical reform or old wine in new bottles?]. *Contemporary Linguistics* 3: 273–283.
Li X and Lei L (2019) A bibliometric analysis of topic modelling studies (2000–2017). *Journal of Information Science* 47: 161–175.
Lei L and Liu D (2019) Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics* 40: 540–561.
Liu X, Xu Q, and Li M (2015) A comparative analysis of scientific publications in management journals by authors from Mainland China, Hong Kong, Taiwan, and Macau: 2003–2012. *Scientometrics* 105: 135–143.
Machery E (2021) Cross cultural semantics at 15. In: Biggs S and Geirsson H (eds.) *The Routledge Handbook of Linguistic Reference*. New York: Routledge, pp.535–550.
Machery E, Mallon R, Nichols S, et al. (2004) Semantics, cross-cultural style. *Cognition* 92(3): B1–B12.
Machery E, Sytsma J, and Deutsch M (2015) Speaker's reference and cross-cultural semantics. In: Bianchi A (ed.) *On Reference*. Oxford: Oxford University Press, pp.62–76.
Marsili N (2016) Lying by promising: A study on insincere illocutionary acts. *International Review of Pragmatics* 8(2): 271–313.

Marsili N and Wiegmann A (2021) Should I say that? An experimental investigation of the norm of assertion. *Cognition* 212: 104657.

Meibauer J and Steinbach M (eds.) (2011) *Experimental Pragmatics/Semantics*. Amsterdam: John Benjamins Publishing Company.

Nado J (ed.) (2016) *Advances in Experimental Philosophy and Philosophical Methodology*. London: Bloomsbury Publishing.

Mukerji N (2019) *Experimental Philosophy: A Critical Study*. London and New York: Rowman & Littlefield International.

Noveck I (2018) *Experimental Pragmatics: The Making of a Cognitive Science*. Cambridge: Cambridge University Press.

O'Neill E and Machery E (eds.) (2014) *Current Controversies in Experimental Philosophy*. New York: Routledge.

Pritchard A (1969) Statistical bibliography or bibliometrics? *Journal of Documentation* 25: 348–9.

Reins LM and Wiegmann A (2021) Is lying bound to commitment? Empirically investigating deceptive presuppositions, implicatures, and actions. *Cognitive Science* 45(2): e12936.

Rett J (2018) The semantics of many, much, few, and little. *Language and Linguistics Compass* 12(1).

Reuter K (2019) Dual character concepts. *Philosophy Compass* 14(1): e12557.

Reuter K and Brössel P (2019) No knowledge required. *Episteme* 16(3): 303–321.

Schindler S, Drożdżowicz A, and Brøcker K (eds.) (2020) *Linguistic Intuitions. Evidence and Method*. Oxford: Oxford University Press.

Silk A (2021) Evaluational adjectives. *Philosophy and Phenomenological Research* 102(1): 127–161.

Syrett K and Koev T (2015) Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics* 32(3): 525–577.

Sytsma J and Buckwalter W (eds.) (2016) *A Companion to Experimental Philosophy*. West Sussex: John Wiley & Sons, Ltd.

Sytsma J, Livengood J, Sato R, et al. (2015) Reference in the land of the rising sun: A cross-cultural study on the reference of proper names. *Review of Philosophy and Psychology* 6(2): 213–230.

Turri A and Turri J (2016) Lying, uptake, assertion, and intent. *International Review of Pragmatics* 8(2): 314–333.

Turri J (2013) The test of truth: An experimental investigation of the norm of assertion. *Cognition* 129(2): 279–291.

Turri J (2016a) Knowledge, certainty, and assertion. *Philosophical Psychology* 29(2): 293–299.

Turri J (2016b) The point of assertion is to transmit knowledge. *Analysis* 76(2): 130–136.

Turri J (2016c) Knowledge and assertion in "Gettier" cases. *Philosophical Psychology* 29(5): 759–775.

Turri J (2017) Experimental work on the norms of assertion. *Philosophy Compass* 12(7): e12425.

Turri J (2021) Truth, fallibility, and justification: New studies in the norms of assertion. *Synthese* 198(9): 8073–8084.

Weissman B and Terkourafi M (2019) Are false implicatures lies? An empirical investigation. *Mind & Language* 34(2): 221–246.

Xie Z and Willett P (2013) The development of computer science research in the People's Republic of China 2000–2009: A bibliometric study. *Information Development* 29(3): 251–264.

Zhang X (2020) A bibliometric analysis of second language acquisition between 1997 and 2018. *Studies in Second Language Acquisition* 42(1): 199–222.

Appendix

Table 5. The 60 n-grams with the frequency above 10

N-gram	Count	N-gram	Count
intuition	251	natural kind	18
reference	193	truth value	18
experiment	172	contextualist	17
participant	111	slur	16
truth	92	experimental philosopher	16
judgment	81	semantic intuition	16
Machery	74	semantic pragmatic	16
descriptivist	74	armchair	15
context	71	cultural difference	15
assertion	65	aesthetic	14
theory reference	60	pluralism	14
knowledge	59	westerner	14
adjective	49	empirical evidence	14
epistemic	48	semantic reference	14
Kipke	48	Devitt	13
experimental study	39	dual character concept	12
cross-cultural difference)	35	ordinary language philosophy	12
epistemic modal	34	experimental semantics	12
predicate	34	logical form	12
presupposition	34	norm assertion	12
causal historical theory	34	referential intuition	12
ambiguity	31	Chinese	11
experimental philosophy	30	Japanese	11
competence	30	metalinguistic	11
intuition reference	28	colour adjective	11
contextualism	25	quasi-indexical	11
relativism	25	stereotypical inference	11
Gödel	23	acceptability	10
linguistic intuition	22	ordinary speaker	10
implicature	21	definite description	10

ORIGINAL ARTICLE

A new approach to the study on counterexamples of generic sentences: From the perspective of interactive reference point-target relationship and re-categorization model

Yunfei Liu, Yiting Yu, Siyu Chen*

School of Foreign Languages, China Three Gorges University, Yichang, China

Abstract: Based on deficiencies of existing researches, this paper, aiming at taking the tolerance of counterexamples reflecting seeming syntax-semantic mismatch in generic sentences, and the online cognitive process of these sentences into the same analyzing framework, proposes the Interactive Reference Point-target Relationship and Re-categorization Model (IRPR-RC Model) to give a unified explanation to the main types of counterexample-tolerating generic sentences (GS), thus further fulfilling the generalization commitment of cognitive linguistics. According to this model: 1) there is an interaction relationship between reference points and targets connecting generic words and attribute words in counterexample-tolerating generic sentences (GS); 2) this interactive relationship provides the premise for re-categorization, which selects a particular sub-category and makes it salient. This process can also be viewed as a phenomenon of attribute words coercing the generic words; 3) the model can be divided into three types: Focusing Type, Imbedding Type and Repulsing Type, according to different operation mechanism of IRPR-RC Model in counterexample-tolerating generic sentences (GS).

Keywords: interactive reference point-target relationship and re-categorization model; generic sentence; counterexample-tolerating

***Corresponding author:** Siyu Chen, School of Foreign Languages, China Three Gorges University, Yichang 443002, Hubei Province, China; siyuchen1117@163.com

Received: April 15, 2022; **Accepted:** June 6, 2022; **Available online:** August 12, 2022

Citation: Liu Y, Yu Y, and Chen S (2022) A new approach to the study on counterexamples of generic sentences: From the perspective of interactive reference point-target relationship and re-categorization model. *Forum for Linguistic Studies* 4(1): 36–49. DOI: 10.18063/fls.v4i1.1454

1. Introduction

A Generic Sentence (GS) is a kind of special sentences that generalizes a kind of things rather than an individual with certain attributes or regular characteristics. Its subject-predicate structure is described as "S + P", in which "S" is the subject, which refers to the generic word, with conceptual connotation of categorization, usually used by bare nouns to express certain class of things; "P" refers to the generic attribute. GS is often reflected as "predicative", "verb" and "auxiliary verb". Such

as:

(1) a. Birds fly.

b. Chinese speak Mandarin.

c. Whales are mammals.

For "predicative", when we hear (1a), listeners ignore the fact that ostriches, penguins, and newborn birds cannot fly, while believe that the judgment of "birds fly" is correct (Carlson and Pelletier, 1995: 43). Compared with the generic words, the number of counterexamples cannot be determined in proportion.

The example of "verb" (1b) is similar to the former, which also implies the feature above. In China, due to regional, historical, educational and other reasons, some Chinese can only speak dialects, not Mandarin. However, listeners normally tolerate the counterexample and still regard GS (1b) as reality. In short, the tolerating of counterexamples happens when people use logical methods to carry out syllogism reasoning, the subject cannot always cover the whole category; while the users of GS default to the existence of counterexamples (Zhou, 2004; Xu, 2010a; Lei, 2019).

On the contrary, the "auxiliary verb" (1c) does not reflect the feature above. Since biologically, any type of whale is a mammal. That is to say, only some of generic sentences can tolerate counterexamples.

For the classification of GS, previous scholars mostly focused on the generic word. According to the description of the syntactic morphology of English generic words, Krifka (1987: 4–12) divided the generic sentences into "I-generic (indefinite)" and "D-generic (definite)" or "kind-referring sentence" and "characterizing sentence". Inspired by Lesile (2007), Prasada and Dillingham (2009) and Sandeep et al. (2013: 405) based on the relation between generic words and referential attributes, further divided GS into "minority characteristic generics", "majority characteristic generics" and "majority statistical generics".

However, different from the scholars' classifications above, Xu (2010a) observed the morph of generic sentences and the pragmatic effect of the feature "counterexamples-tolerating". Therefore, the generic sentence is divided into "law-like generic sentence", "implicit generic sentence" and "temporary generic sentence". Such as:

(2) a. xiàng chǐ fén shēn.

Elephants are killed for ivories of good quality. (literal translation)

b. *Everyone went home,* only Kim still working overtime.

The law-like generic sentence is a concise representation which involves law-like expressions of things. As mentioned above, both (1a) and (1b) tolerate counterexamples. For example, language users in (1a) assume that "birds fly" is a conventional judgment of birds' characteristics, but there are still counterexamples to this conventional fact.

"Implicit generic sentence" refers to the implicit use of generic sentences as presuppositions in metaphor, metonymy or idiom poetry. (2a) conveys and presupposes that "elephants are killed for

their valuable teeth"; and also "fén shēn" metaphorically expresses "being hunted", which indicates the tolerant of counterexamples. Not all elephants have valuable teeth, and those with short, broken teeth or even toothless will not be killed for ivories. Only those with high-quality ivories will be killed.

Using the feature "tolerating counterexamples", "temporary generic sentence" is an application of rhetoric to generalize a kind of temporary event/thing into a kind of expression of regularity or integrity (Xu, 2010a). For (2b), "*Everyone went home*" is identified as a class of event, whose generic attribute is "off duty"; while "Kim" who is still working overtime in the company is a salient counterexample with immediacy, thus reflecting a strong contrast.

In a word, the previous classifications of generic sentences mostly focused on the characteristics of generic words, without the attributes. Therefore, they did not pay much attention to the phenomenon of "tolerance of counterexamples". Instead, Xu (2010a) took account of this phenomenon. Since this paper aims to take the cognitive process of generic sentences and the tolerance of counterexamples into a unified framework, it focuses on these three types of generic sentences classified by Xu (2010a). Drawing on the results and reflecting on deficiencies from previous studies, this paper constructs an Interactive Reference Point-target Relationship and Re-categorization Model (IRPR-RC Model) to explain the cognitive motivation of counterexamples in the formation of three types' counterexample-tolerating generic sentences (GS) above.

2. Previous studies

As discussed above, one of the most important features of GS is the tolerating of counterexamples. And it is also the research focus in the study of GS, which has led many scholars to analyze the motivation of this phenomenon from the cognitive perspective.

1) The Categorization View. It holds that the tolerance of counterexamples in GS is that people focus on the whole of some particular category, while neglects other members. Eckardt (1999) proposed that "*GS is the statement about the prototype of category*". Based on this view, some scholars (Xu, 2010a; Li, 2013; Alexander and Frank, 2019), starting from the perspective of pragmatic features and philosophical concepts, studied the prototypical category. They further pointed out that conceptualizer excludes the marginal members in the category referred by generic words and focuses on the overall properties to express the general attributes or features, which leads to the tolerance of counterexamples in GS. Since human focuses on the whole category of the "class" and ignores the counterexample in the category, this kind of general expression is generated (Xu, 2010b).

In addition, that the category is focused as a whole and some counterexamples in the category are neglected indicates that the boundaries of categories are fuzzy and the classification of GS is based on the overall properties of the category rather than the sufficient conditions of each feature (Li, 2012). In other words, the counterexample can be tolerated for this bottom-up categorization process and particular pragmatic needs. This indirectly proves Xu's (2010a) view that the pattern, by which human recognizes things in terms of classes, provides possibilities of tolerating counterexamples in GS. Besides, Li (2013) pointed out that via a series of strategic choices of categorization in the linguistic thinking stage, the conceptualizer transforms the "event" into the "usage event", thereby the counterexample-tolerating GS is formed.

2) Holistic View. This mainly focuses on the overall generation of GS and the pragmatic features to explore the tolerance of counterexamples. Liu (2010), Bu (2012) and Li (2013), borrowing from the Generative Holism in philosophy (part is the embodiment of the whole), pointed out that the semantics of the subject in GS is generated by the whole. In this generating process, human selects restricted model(s) from the potential models provided by the whole within the context of SP. And the selecting process from the whole does not necessarily require the selected model bearing the most important, typical, and obvious features of the category, but this model must serve specific pragmatic purpose (Liao, 2010). Wei (2012) further pointed out that pragmatic purpose can be reflected as pragmatic presupposition, which brings about the generic range of GS that changes with the context, for that people can tolerate the "overgeneralization" of GS.

The overall concept that GS refers to is dominating in information processing, which can be easily solidified in people's cognition. Therefore, the referred concepts are reinforced, and exceptions are ignored (Xu, 2010a; Sun and Cheng, 2013). Focusing on the cognitive operation in the comprehension of GS, the holistic view revels that the GS can tolerate counterexamples while be judged to be true, for the comprehension of subject is constrained by GS as a whole in generation process.

3) The Metonymic View. Centering around the mirroring relationship between "class" and "example", Metonymic View explores the causes of counterexample tolerating in GS. Zou and Zhang (2011) pointed out that the existence of counterexamples is the result of metonymy by replacing essence with highly inductive attribute. In the framework of Idealized Cognitive Model (ICM), Fu (2010) and Gao (2013) proposed that counterexample-tolerating GS is generated by metonymic mechanism and prototype effect, whose truth-value is judged under certain ICM rather than reality. Counterexamples can be accepted within a reasonable range when the activated matrix domain and the category hierarchy used for subject-predicate items in ICM are reasonable (Fu, 2017). Relying on metonym, human experiences, cognizes and generalizes some features in the category referred by the subject, leading to tolerance of counterexamples in GS (Lei, 2019).

4) The Error-tolerance View. It mainly focuses on quantitative expressions to analyze the tolerance of counterexamples in GS. The truth condition for the existence of counterexamples in GS should be the existential quantification, i.e., the existence of an individual X, which belongs to the general attribute P and has the attribute P1 (Krifka, 1987; Wu, 2010). According to some scholars, the quantification process of tolerating counterexamples is either the result of the application of default rule in logic (Krifka, 1987) or the result of pragmatic reasoning in the use of language based on error-tolerance mechanisms and "overgeneralization" (Wu, 2010; Wu and Wei 2012).

Related studies have focused on GS from macro to micro, deepening our understanding of the tolerance of counterexamples, but the following shortcomings still remain: 1) explaining the formation of GS and the counterexample-tolerating in isolation leads to the incompatibility of these two kinds of explanations to each other. Fu (2017: 137) expressed her concern about this analysis method: "*Although the pragmatic study of GS can explain the tolerance of counterexamples in a certain extent, it is difficult to explain the syntactic-semantic features of GS by only referring to the dynamic context for pragmatic analysis.*" This paper argues that the tolerance of counterexamples in GS is formed in the cognitive process of its expression. So the study of the construction and the counterexample of GS should be carried out in a same framework. 2) The relevant studies mainly focus on one type of counterexample-tolerating GS, and the relevant explanations cannot be applied to other

types of GS. This violates the generalization commitment of cognitive linguistics that different dimensions of language share specific and universal foundational organizing principles. If there is no unified explanation for the series of syntactic phenomena (for example, counterexample-tolerating GS), how can we explain the cognitive mechanisms shared by different dimensions of human language? The GS is an organic whole, which should be described and explained in a unified framework. These two shortcomings constitute the starting point of this paper.

3. Theoretical foundation

Based on the deficiencies of previous studies, the following part further analyzes the constructing process and the tolerance of counterexamples of GS. As mentioned in the introduction, the structure of generic sentences is described as "S + P", in which generic attribute "P" refers to certain characteristics of generic word "S". So "P" and "S" is closely related, and "P" can be activated by "S", from which we can see that the formation of generic sentences conforms to cognitive reference point relationship. Except that, the tolerance of counterexamples in generic sentences involves two categories of "S", one of which has the attribute of "P", while the other doesn't. For example, the sentence "Peacock has beautiful feathers" involves two categories of "peacock", one of which is the category of "beautiful feathers", while another is the category "without beautiful feathers". Therefore, categorization is the cognitive mechanism that concerns the constructing process of generic sentences. Based on this, this paper unfolds the discussion on the theoretical basis of cognitive reference point relationship and categorization.

3.1. Cognitive reference point relationship

"Cognitive Reference Point" was first proposed by Rosch (1975) in his research of prototypical category, further discussed and elaborated by Langacker (2008: 84).

According to Langacker (2008), the operation of cognitive reference point model in **Figure 1** can be divided in two phases: 1) since entities have different prominence levels, the relatively prominent entities are focused by conceptualizer as the reference point, through which conceptualizer establishes "mental contact" with other entities. After the reference point is invoked and becomes the focus of conceptualizer, it provides access to its dominion that contains all of its potential targets, and refers to one of the targets in the dominion; 2) as focus shifts to the target, the original reference point fades into background, and the target can activate its own dominion and then may be invoked

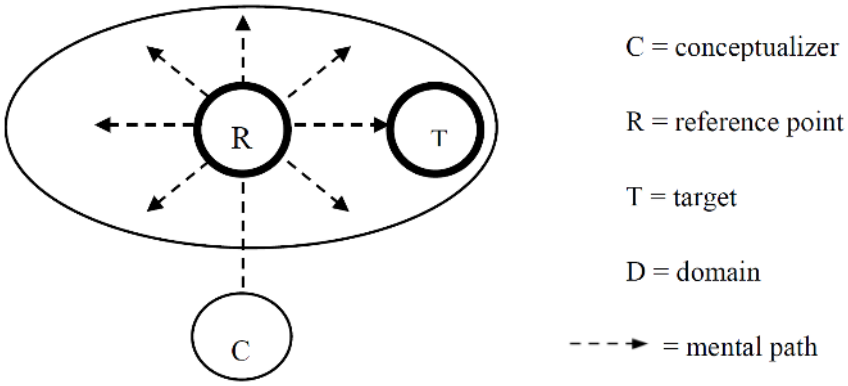


Figure 1. Cognitive Reference Point Model (Langacker, 2008: 84).

as the reference point to continue to reach another target. In a word, the mental path, through which the reference point refers to the specific target, is consciously and subjectively constructed by conceptualizer according to communicative purpose (Wei, 2008: 96).

3.2. Categorization

According to Jackendoff (1985: 77), "*an essential aspect of cognition is the ability to categorize: to judge that a particular thing is or is not an instance of a particular category.*" The word category comes from Greek and refers to predicates in propositions expressing judgments. Aristotle systematically studied and sorted the category system for the first time in *Organon • Category* (Wen and Jiang, 2001). Categorization is the basis for the construction of categories and the mental process by which human beings classify things (Ungerer and Schmid, 2005). Except that, "*the mechanism of categorization must be assigned to the level of conceptual structure, where all these types of information are available. In short, a categorization judgment is the outcome of the juxtaposition of two conceptual structures* (Jackendoff, 1985: 78)." Specifically, categorization is a mental process of categorizing external things based on subjective and objective interaction, and a rational activity that endows the world with structure based on the subjective concept and classification of objective things (Wang, 2001). The outcome of categorization is what we usually call concept, and language reflects the categorized conceptual structure of the external world (Geeraerts, 1997).

Therefore, based on cognitive reference point relationship and categorization, the motivation of IRPR-RC Model is further discussed in the following part, which tentatively explains the cognitive process and tolerance of counterexamples of GS.

4. Cognitive motivation of IRPR-RC Model

4.1. Interactive reference point relationship

This paper argues that it is interactive reference point relationship rather than reference point relationship mentioned above that can be used to explain the tolerance of counterexamples in generic sentences. The cognitive reference point relationship is a universal cognitive mechanism, which not only exists in language expression, but also in vision. For instance, Langacker (1991: 170) once used "night-time sky" to describe the relationship of visual reference point relationship, i.e., conceptualizer usually locates the position of other stars by taking bright and salient stars as the reference point, which applies the universal cognitive mechanism of visual reference point relationship. Wang (2002) believes that such visual reference point relationship is also common in daily life. For example, to find a place on the map, it is often necessary to find a large and famous place and use it as a reference point to localize other places. Whereas, our visual cognitive mechanism not only exists in the process of reference point pointing target, but also in target pointing back to reference point. For example, in order to remember the position of W, O, R and K when typing "work" on the keyboard, unfamiliar with the keyboard, beginners usually first take Q as the reference point and move their eyes to W, and then take W as a new reference point, moving their eyes from W back to Q to re-locate W.

Just as the visual cognitive reference point relationship is reflected in language, the visual interaction reference point relationship also exists in language. Langacker (1993: 16) believed that there was a "potential interaction" between the reference point and the target.

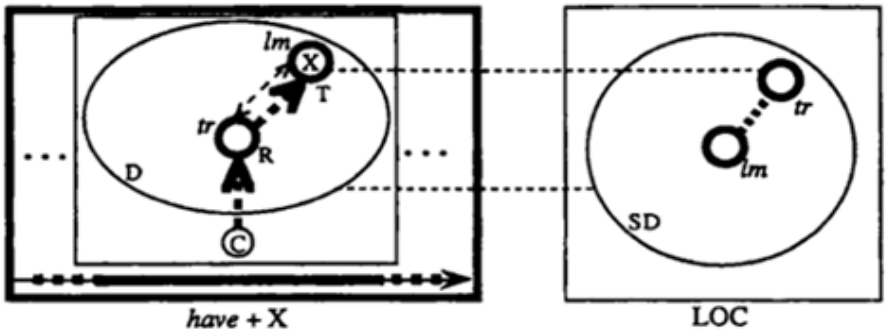


Figure 2. "*have* + X" Reference Point Relationship Model (Langacker, 1993: 16).

According to **Figure 2**, since the target in the possessive domain is equal to the trajector in the searching domain (Langacker, 1993), so "skunks" and "we" in (3a) can realize the potential interaction between trajector and landmark, which is the potential interaction between the cognitive reference point and the target. Besides, Langacker (2008: 79) also mentioned the conceptual anaphora in time dimension. In (4a), "quit my job" and "got married" are successively focused as reference point to refer to next target. But in (4b), due to the existence of "in reverse order", the original reference point relationship similar to (4a) is reversed. Starting from the last target "quit my job", the target refers back to the reference point one by one in the mental scanning process of the conceptualizer, forming the interactive relationship between reference point and target.

Other than that, Langacker (2008: 84) pointed out that in (5a), "*the reference point is the surgeon, and his wife is the target*". Conceptualizer first directs the attention to the reference point "the surgeon" for specific purpose of locating the target "his wife". Since "*it is no accident that the target 'his wife', contains a possessor pronoun referring back to the reference point*", thus the interaction between reference point and target is formed.

(3) a. We have a lot of skunks around here. (Langacker, 1993: 16)

(4) a. I quit my job, got married, and had a baby. (Langacker, 2008: 79)

b. I had a baby, got married, and quit my job—in reverse order, of course. (Langacker, 2008:79)

(5) a. Do you remember that surgeon we met at the party? His wife just filed for divorce. (Langacker, 2008: 84)

Moreover, the extension of the usage of reference point and its interactive relationship is a common phenomenon in Chinese (Liu, Qu, Chen, et al, 2021; Liu, Luo and Wu, 2021). Thus, interactive reference point relationship exists in language. This paper holds that the interactive relationship between the target P and the reference point S provides premise for re-categorization in generic sentences, which is discussed as follows.

4.2. Re-categorization

The interaction reference point relationship mentioned above provides premise for re-categorization. Liu and Li (2005) pointed out that in translation, a category in one language would not exactly correspond to that in another language because of different categories dimensions, so there would

be a phenomenon of crossing correspondence among different languages, which leads to re-categorization. However, it not only exists in the interaction among different languages, but also in the expression of same language. The members in one category need to be reclassified under specific circumstances (Wang, 2015). In generic sentence, the generic word S is first invoked as a reference point by conceptualizer in relevant dominion and then refers to the target P, which is the generic attribute. After P is invoked and focused as the target, it becomes a new reference point referring back to S, which forms the interactive reference point relationship, and the purpose of this anaphora is to achieve re-categorization. In the process of re-categorization, based on the characteristics of generic attribute, the conceptualizer selects a particular member of generic word's category, which usually is a sub-category of S. As an entity of S's actual reference, it contains the features of generic attribute, so the tolerance of counterexamples in generic sentences is formed. In conclusion, the interaction between the reference point and the target is the premise of re-categorization, and re-categorization is the source of tolerating counterexamples.

Based on the analysis above, this paper constructs an IRPR-RC Model to give a unified explanation to the main types of counterexample-tolerating GS listed by Xu (2010a). As shown in Figure 3, in generic sentences, the reference point in this model is generic word S, and the target is generic attribute P. The interaction between the reference point and the target provides premise for re-categorization, which forms the subcategory that conforms to the features of attribute word. In all, this model can be used to explain not only the construction process of counterexample-tolerating GS, but also the motivation behind the tolerance of counterexample in generic sentence.

5. Explanatory power of the IRPR-RC Model

The previous section discussed the motivation of the IRPR-RC Model. Now based on this model and its operation, this section offers a unified explanation to the construction of these three types of counterexample-tolerating GS classified by Xu (2010a), to further verify the explanatory power of the model.

The IRPR-RC Model can uniformly explain the main types of counterexample-tolerating GS. As shown in Figure 3, the model operates in three stages. 1) Conceptualizer uses a generic word as a

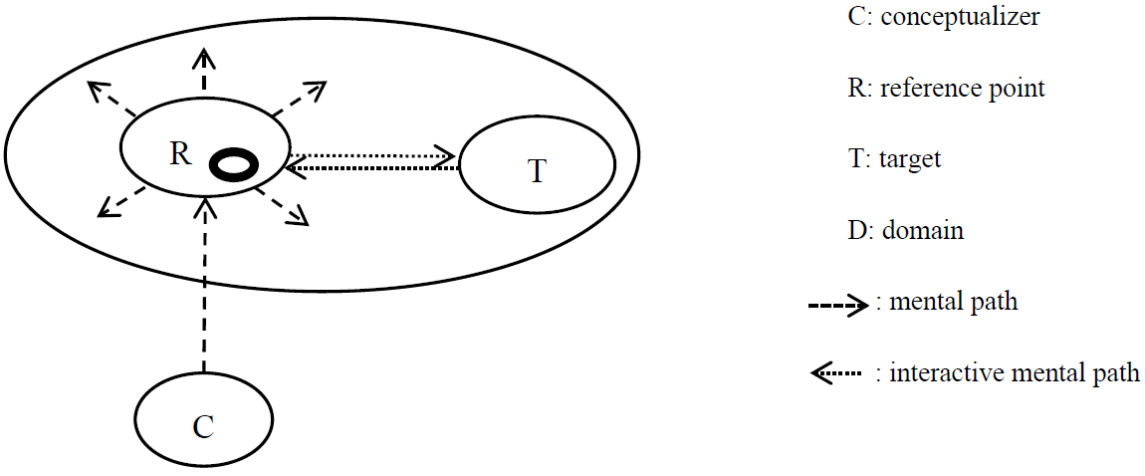


Figure 3. Interactive Reference Point-target Relationship and Re-categorization Model.

cognitive reference point, activating its cognitive dominion. For the communicative purposes, the conceptualizer mentally scans to the target (attribute P) via the reference point. 2) The target becomes salient as a new cognitive reference point, which refers back to the previous reference point (new target). The interaction between the reference point and the target is realized in this stage, which provides the premise for the re-categorization. Choosing one subcategory of the generic word and making it salient can also be viewed as attribute word coercing generic word, which is consistent with the process of lexical coercion (Wang, 2009). 3) The conceptualizer re-categorizes the new target (generic word). Thus, the generic word is divided into subcategory or subcategories that conform to the attribute of the generic word. A particular subcategory is the foreground in the expression, and the rest parts fade into background, thus forming the tolerance of counterexamples. Meanwhile, the counterexample-tolerating GS is constructed.

Based on the different operation characteristics of IRPR-RC Model in the different types of GS, the Model can be classified into three sub-types: Focusing Type, Imbedding Type and Repulsing Type. Relying on these sub-types of the model, this paper explains the main types of counterexample-tolerating GS one by one.

5.1. The explanation of law-like GS

The IRPR-RC Model in this type of GS is the Focusing Type, the characteristic of which is that after the target points back to the reference point (generic word), some members are brought into the focus of the attention, forming the foreground part of generic word, and becoming the word's subcategory bearing the features of the original target (attribute word's profile).

The operation of the model can be divided into three stages: in the first stage, the conceptualizer views the generic word as a cognitive reference point, activating its relevant cognitive dominion and scans to the target (attribute word's profile) in the dominion. In the second stage, when the target is salient, it becomes the new reference point pointing back to the generic word, which becomes salient again, establishing the interaction between the reference point and the target. In the third stage, the conceptualizer re-categorizes the generic word, picking a subcategory out of the category of generic word. And this subcategory bears characteristics expressed by the attribute word, and plays the role of foreground, while the rest part of the generic word's category fades into background, which produces the tolerance of counterexamples.

(1a) is explained as an example of the operation process of IRPR-RC Model in law-like GS via **Figure 4.** In (1a), the conceptualizer takes "Birds" as a reference point to activate its cognitive dominion. And the conceptualizer mentally scans to the target that can facilitate communication, i.e., "can fly", which then becomes the new reference point, and refers back to the generic word "Birds", which becomes salient again. Finally, the conceptualizer re-categorizes the category of "bird" by selecting a subcategory "the birds that can fly". This subcategory of the "bird" reflects the feature of the original target (attribute word's profile), and it is the foreground concept (as shown in bold ellipse) according to conceptualizer's construal. Thus, in this three-stage online cognitive process, this type of GS is constructed and the counterexample tolerance is formed.

5.2. The explanation of invisible GS

The IRPR-RC Model in this type of GS is the Imbedding Type, in which the re-categorization involves the imbedding of three categories.

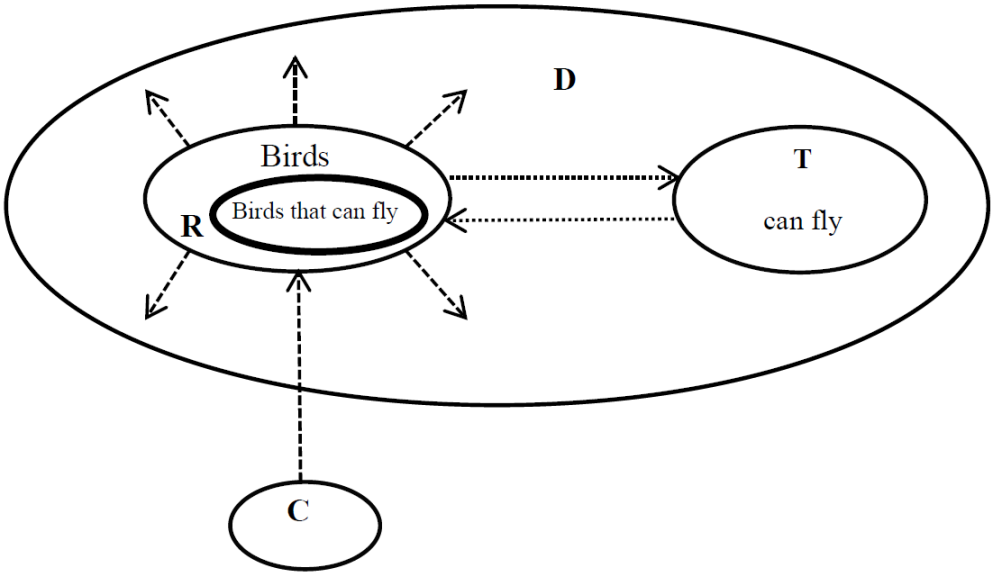


Figure 4. Focusing Type with *Birds fly* as an example.

This model operates as follows: first, with a generic word as the cognitive reference point, the conceptualizer activates its relevant cognitive dominion and mentally scans to target1, which is usually expressed directly. Afterwards, based on communicative purposes, target1 serves as a new reference point, pointing to target2, which is indirectly expressed. Next, as the new reference point, target2 in turn refers back to the generic word, forming the interaction between reference points and targets, which provides premise for re-categorization. Lastly, conceptualizer re-categorizes the generic word, in whose process, the conceptualizer divides the category of the generic word into two sub-categories: one category with target1's features and another with target2's features. This leads to a successively imbedded categories: the category to which the generic word refers, contains the subcategory with the feature of target1, which in turn contains the subcategory with the feature of target2.

Based on **Figure 5**, taking (2a) as an example, this section explains the operation of IRPR-RC Model in invisible GS. For certain communicative purposes, conceptualizer chooses "Elephants" as a reference point to activate its relevant cognitive dominion; meanwhile, conceptualizer mentally scans to target1 "ivories" which facilitates communication. Then target1 becomes a new reference point pointing to target2 "be killed for ivories of good quality" ("be killed" means "be haunted", and the whole sentence means that the elephants are killed because of their ivories, implying that only elephants with high-quality ivories will be killed. Elephants with broken tusks, or with newly grown ivories, etc., will not be killed for that reason, so target2 is expressed indirectly). Then target2 points back to the generic word "Elephants", which becomes salient again, forming the interaction between reference points and targets. Thus, based on target1 and target2, conceptualizer re-categorizes the category of "elephant" into two subcategories: the one with ivories and another with high-quality ivories. The "elephant" category includes elephants with ivories, and the latter includes elephants with high-quality ivories. This constitutes imbedding relationships in which the category "with high-quality ivories" is presented as foreground (as shown in bold ellipse in Figure 5), and with the other categories in the background, this leads to tolerating of counterexamples.

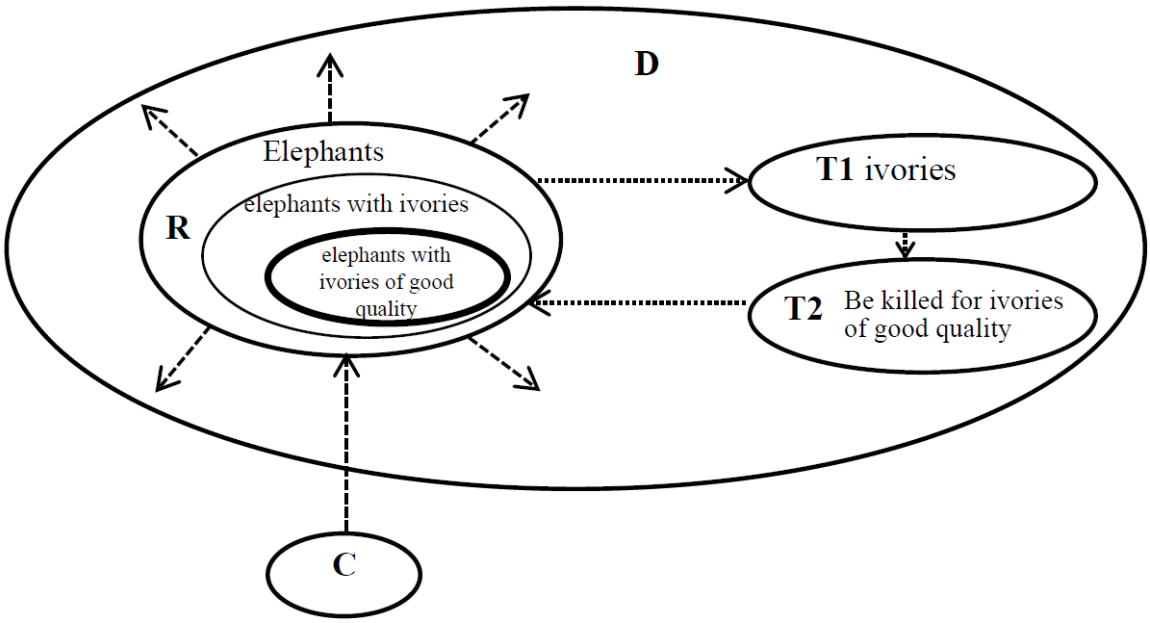


Figure 5. Embedding Type with *xiàng chǐ fén shēn* as an example.

5.3. The explanation of temporary GS

In the temporary GS, the IRPR-RC Model is the Repulsing Type, and the target2 is the salient counterexample. In this model, the category expressed by generic word is foreground concept, with a subcategory corresponding to target2 repulsed.

The operation mechanism of this type of model is similar to that of the Imbedding Model, with the main differences in the re-categorization stage. In this stage, the conceptualizer chooses the corresponding subcategory of target2, which is repulsed, as background concept, from the category of the generic word, with the remaining category content of the generic word construed as foreground concept, serving communicative purpose.

Taking (2b) as an example, the online cognitive process of it is shown in **Figure 6**. The operation is similar to that of Figure 5, but there are differences in re-categorization. At this stage, the conceptualizer re-categorizes the category referred by the generic word, selecting a subcategory corresponding to target2 "Kim working". This subcategory (marked by × as shown with dotted line ellipse in Figure 6) is repulsed from the category of the generic word. And the remaining category of the generic word is construed as foreground concept (as shown with the bold ellipse in Figure 6, and marked with √), with the subcategory construed as background for communicative purpose.

6. Conclusion

Why some Generic Sentences can tolerate counterexamples in contradictory relations? This question puts forward a challenging question for linguistics. The answer to the question is crucial for recognizing and characterizing the features and properties of things and making conventional statements of things (Xu, 2010b).

This paper constructed the IRPR-RC Model and claimed that the reason why some Generic Sen-

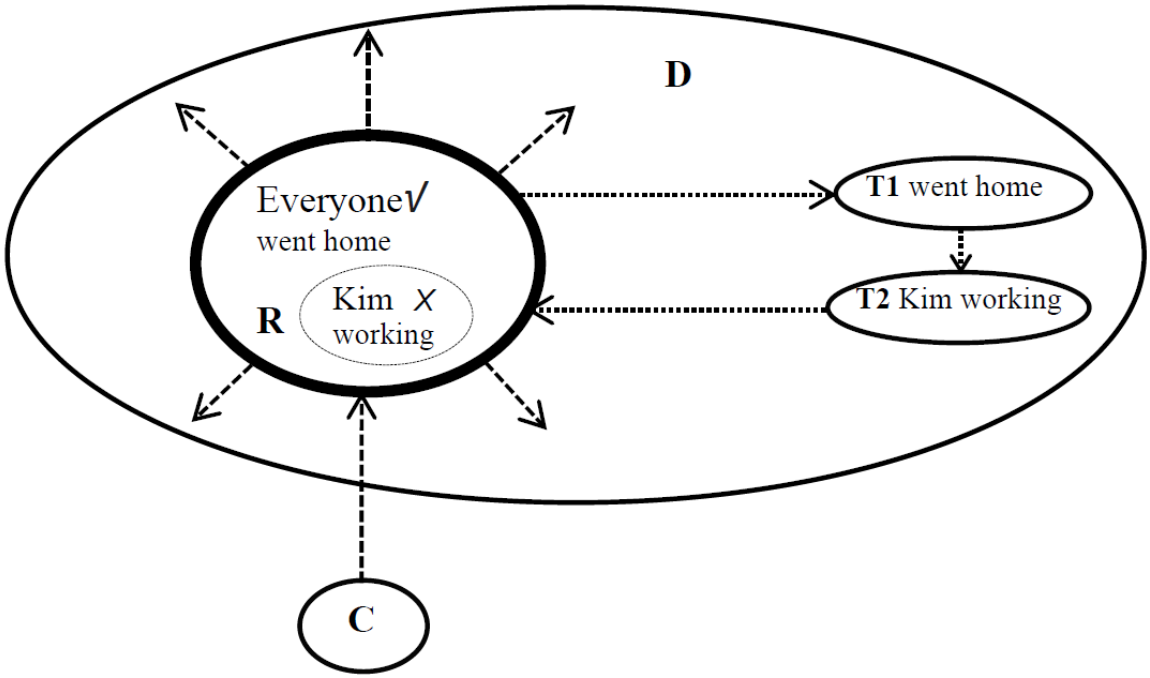


Figure 6. Repulsing Type with *Everyone went home, only Kim still working overtime* as an example.

tences can tolerate counterexamples lies in the interaction between reference point and target and re-categorization based on the interaction, which is also the online cognitive process of these Generic Sentences. And different types of Generic Sentences tolerating counterexamples reflect different types of IRPR-RC Model. The operation mechanism of these types of IRPR-RC Model offers the answer to Li's (2013: 6) question: "*What exactly cognitive mechanism works in Generic Sentences?*"

There are two trends in current researches related to Generic Sentences. 1) Analyzing the construction of generic sentences and the motivation behind its tolerance of counterexamples, but in isolated ways; the formation of Generic Sentences and the tolerating of counterexamples are studied separately. 2) Different types of Generic Sentences are analyzed respectively, and the explanations related to one or several types of generic sentences cannot be applied to other types. For the first trend, this paper argues that the motivation behind tolerating counterexamples must exist in the online cognitive process of Generic Sentences, it cannot exist in the pantheon outside that process. Therefore, this paper explains the online cognitive process of Generic Sentences and the phenomenon of counterexample tolerance in a unified framework, which is the first manifestation of the new approach in this paper. For the second trend, this paper holds that it fails to fulfill the generalization commitment of cognitive linguistics, which believes that different dimensions of human language share particular and universal foundational organizing principles. If the relevant series of phenomena at the syntactic dimensions (different types of Generic Sentences that can tolerate counterexamples) cannot yet be explained uniformly, how can we explain the cognitive mechanisms shared by different dimensions of language? This paper proposes the IRPR-RC Model, tentatively offering a unified explanation to the cognitive construction of Generic Sentences and trying to further realize the generalization commitment of cognitive linguistics, which is the second manifestation of the new approach in this paper.

Conflict of interest

No conflict of interest was reported by all authors.

Acknowledgements

This work is supported by the Social Science Foundation of Hubei Province (Grant number: 21ZD037).

References

Alexander N and Frank CK (2019) Generics designate kinds but not always essences. *Proceedings of the National Academy of Sciences* 116(41): 1–6.

Bu F (2012) *A Cognitive Approach to Chinese Generic Sentences*. Master's Thesis, Hunan University, China.

Carlson GN and Pelletier FJ (1995) *The Generic Book*. Chicago: University of Chicago Press.

Eckardt R (1999) Normal objects, normal worlds and the meaning of generic sentences. *Journal of Semantics* 16(3): 237–278.

Fu Z (2010) *A Conceptual Metonymic Approach to Generics.* Master's Thesis, Southwest University, China.

Fu Z (2017) The embodied basis and cognitive motivation of generic sentences. *Journal of Southwest University (Social Sciences Edition)* 43(6): 137–145, 196.

Gao Y (2013) *A Study of English Generics Based on Conceptual Metonymy*. Master's Thesis, Yan'an University, China.

Geeraerts D (1997) *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*. Oxford: Clarendon Press.

Jackendoff RS (1985) *Semantics and Cognition*. Cambridge: MIT Press.

Krifka M (1987) *An Outline of Genericity*. Germany: Seminar für natürliche-sprachliche Systeme, Tübingen University.

Langacker RW (1991) *Foundations of Cognitive Grammar: Descriptive Application (Vol.II)*. Stanford: Stanford University Press.

Langacker RW (1993) Reference-point construction. *Cognitive Linguistics* 4(1): 1–38.

Langacker RW (2008) *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University.

Lei Q (2019) Denotations and connotations of generic sentences: A prospective from an embodied-cognitive approach. *Foreign Languages in China* 16(6): 44–49.

Leslie SJ (2007) Generics and the structure of the mind. *Philosophical Perspectives* (21): 375–403.

Liu C (2010) A cognitive explanation of generic sentences: Mental model selection under the control of the generative whole. *Journal of PLA University of Foreign Languages* 33(3): 1–6, 127.

Li S (2012) On generic sentence and categorization in terms of cognitive structures. *Journal of Changsha University of Science and Technology (Social Science)* 27(3): 125–129.

Li T (2013) The generation of generic sentences from the philosophical perspective of mind. *Foreign Languages Research* (4): 5–9. DOI: 10.13978/j.cnki.wyyj.2013.04.007

Liao Q (2010) Generic sentence and paradox. *Foreign Language Teaching and Research* 42(6): 424–430, 480.

Liu W and Li H (2005) The cognitive characteristics of re-categorization in C-E translation. *Foreign Language Research* (4): 49–54, 80.

Liu Y, Luo J and Wu F (2021) Construction of cognitive model of syntactic-semantic mismatched resultative constructions: From the perspective of interaction construction grammar. *Foreign Language Education* 42(2): 60–66.

Liu Y, Qu Q, Chen Q, et al. (2021) Chinese quasi-attributive construction: From the perspective of cognitive refer-

ence point relationship and metonymy. *Journal of Foreign Languages* 44(2): 31–40.
Prasada S and Dillingham E (2009) Representation of principled connections: A window onto the formal aspect of common sense conceptions. *Cognitive Science* 33(3): 401–448.
Rosch E (1975) Cognitive reference points. *Cognitive Psychology* (7): 532–547.
Sandeep P, Sangeet K, Sarah-Jane L, et al. (2013) Conceptual distinctions amongst generics. *Cognition* 126(3): 405–422.
Sun X and Cheng X (2013) Generic sentences and its language application. *Foreign Languages Research* (1): 37–41, 112. DOI: 10.13978/j.cnki.wyyj.2013.01.012
Ungerer F and Schimd HJ (2005) *Introduction to Cognitive Linguistics*. Beijing: Foreign Language Teaching and Research.
Wang T (2015) The semantic analysis model of pair construction "A, bai B ne" based on "(De/Re)-categorization". *Language Teaching and Linguistic Studies* (6): 16–21. DOI: 10.13458/j.cnki.flatt.004199
Wang Y (2001) *Semantic Theory and Language Teaching*. Shanghai: Shanghai Foreign Language Education Press.
Wang Y (2002) The philosophical basis for cognitive linguistics: Embodied philosophy. *Foreign Language Teaching and Research* (2): 82–89, 160.
Wang Y (2009) Revision on construction coercion: Lexical coercion and inertia coercion. *Foreign Languages and Their Teaching* (12): 5–9.
Wei Z (2008) The cognitive reference point and pragmatic presuppositions. *Foreign Language Research* (3): 93–97. DOI: 10.16263/j.cnki.23-1071/h.2008.03.017
Wei Z (2012) A cognitive construal of the presupposition mechanism of generic sentences. *Foreign Language and Literature* 28(1): 69–73.
Wen X and Jiang X (2001) Categorization: Cognition in language. *Foreign Language Education* (4): 15–18.
Wu B (2010) On the kind-denoting function of generic sentences. *Foreign Language Teaching and Research* 42(2): 92–96, 160.
Wu B and Wei X (2012) A lexical pragmatic approach to generic sentences. *Foreign Language Education* 33(5): 36–40. DOI: 10.16362/j.cnki.cn61-1023/h.2012.05.012
Xu S (2010a) True or not: About generic sentence and its counterexample. *TCSOL Studies* (1): 47–51.
Xu S (2010b) Cognitive-pragmatic implications of the study of generic sentences. *Language Teaching and Linguistic Studies* 42(2): 83–91, 160.
Zou C and Zhang J (2011) Metonymy and generic. *Heilongjiang Science and Technology Information* (25): 228.
Zhou B (2004) On the essence of generic sentence and concept. *Journal of Peking University (Philosophy and Social Science)* (4): 20–29.

ORIGINAL ARTICLE

A dialogic view on construal: A study on the instantiations of wh-interrogative words in wh-dialogues

Guocai Zeng*

Sichuan University, Chengdu, China

Abstract: As the basic notion in cognitive grammar, construal is expounded from various aspects in terms of different conversational purposes in linguistic communication. Based on the well-discussed dimensions of construal from Langacker's view, this study investigates the dimensions of construal from a dialogic view, particularly in cases of English wh-dialogues, with the aim of deciphering how the dialogic focus, the wh-word positioned at the head of the wh-question, is cognitively grounded in the answer. According to the ways that such wh-words are grounded, the dynamic adjustment of dialogic focuses in wh-dialogues is then examined, hopefully to shed some light on the interpretation of the utterance meaning from a dialogic view.

Keywords: construal; wh-word; dialogic construction; grounding; utterance meaning

***Corresponding author:** Guocai Zeng, College of Foreign Languages and Cultures, Sichuan University, Chuanda Road, Shuangliu County, Chengdu 610207, Sichuan Province, China; 2370739862@qq.com

Received: April 15, 2022; **Accepted:** June 17, 2022; **Available online:** August 19, 2022

Citation: Zeng G (2022) A dialogic view on construal: A study on the instantiations of wh-interrogative words in wh-dialogues. *Forum for Linguistic Studies* 4(1): 50–65. DOI: 10.18063/fls.v4i1.1455

1. Interactive nature of construal operation

Construal is a key notion in cognitive grammar. In the view of Langacker (2007: 17), *construal* is our ability to conceive and portray the same situation in different ways. In linguistic communication, grammatical constructions essentially encode the ways interlocutors' interpretations of the shared focuses, and different linguistic expressions might work as descriptions for the same target discussed by speakers in conversation, as **Figure 1** shows.

According to Langacker (2008: 43), there could be four options of construal of the same conceptual content in Figure 1, with four different expressions designating the content distinctly.

Construal (1): *the glass with water in it* designates the container;

Construal (2): *the water in the glass* designates the liquid it contains;

Construal (3): *the glass is half-full* designates the relationship wherein the *volume* occupied by

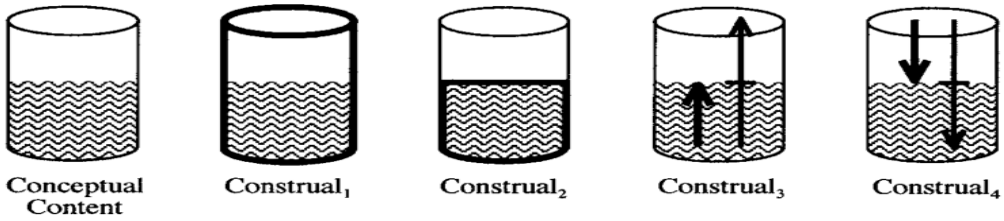


Figure 1. Construal of the same situation (quoted from Langacker, 2008: 44).

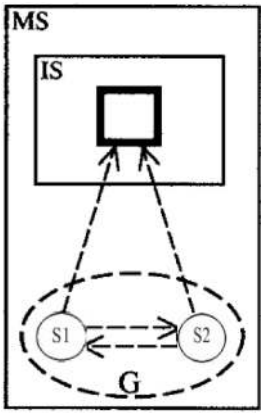


Figure 2. Interactions in the construal process. S1-Speaker 1; S2-Speaker 2; G-Ground; IS-Immediate Scope; MS-Maximal Scope. (adapted from Langacker, 2008: 261).

the *liquid* is just half of its potential volume;

Construal (4): *the glass is half-empty* designates the relationship wherein the *volume* occupied by the *void* is just half of *its potential volume*.

Upon reviewing the above four expressions for the same scene from a dialogic view, they basically indicate how a shared focus in conversation is construed by different speakers, suggesting interactions between speakers, the object being construed, the time and space, and other contextual factors, as shown in **Figure 2**.

To specify, in Figure 2 we can see in the process of construing something in the dialogue, there are interactions between speaker 1 and speaker 2 (the hearer), the object being construed in the immediate scope (IS) and the maximal scope (MS) of interlocutors' joint attention. The ground (G) indicates the time, space, and other elements of circumstance.

This study is to investigate the interactive nature of construal operation from a dialogic view and to discuss the dimensions of construal that function in interpreting the shared focuses in conversations with evidence shown in WH-dialogues. The structure of this study is outlined as follows: at the beginning, the interactive nature of construal operations in conversation is proposed, followed by a fine-grained discussion of how construal works in the grounding of dialogic focuses in English wh-dialogues. After that, types of dynamic adjustments of dialogic focuses in such conversations are investigated. At the end of this study, a brief summary of the dimensions of construal from a dialogic view is provided.

2. Dimensions of construal: A developmental view

From the view of Langacker (2007: 17), there are many dimensions of construal in linguistic communication. In general, the studies on construal in cognitive grammar until now could be classified into four major phrases.

In the early works on cognitive grammar contributed by Langacker (1987, 1991, 1993, 1999), construal is discerned in a broad sense from five aspects, namely the ***level of specificity*** (or ***schematicity***), ***prominence***, ***scope***, ***background***, and ***perspective***. Later, Langacker (2007: 17) particularly emphasizes three of them, which are the ***level of specificity***, ***prominence***, and ***perspective***. Along with in-depth explorations of natural languages in the framework of cognitive grammar, the core concept of construal is re-elaborated from four dimensions by Langacker (2008/2013: 55); they are ***specificity***, ***prominence***, ***perspective***, and ***focusing*** that includes ***selection***, ***foreground***, and ***background***. In his recent studies, Langacker (2015: 120, 2019: 141) re-examines the nature of construal and specifies it with five dimensions, namely ***perspective***, ***selection***, ***prominence***, ***dynamicity***, and ***imagination***.

With a closer look at the studies on construal in cognitive grammar, we can find that ***prominence***, ***perspective***, and ***specificity*** are the fundamental dimensions, whereas ***focusing***, ***dynamicity***, and ***selection*** are extended aspects of construal, in contrast to the earliest version of dimensions of construal (cf. Langacker, 1987, 1991, 1993, 1999). Langacker (2014, personal communication) also pointed out that construal basically consists of those factors, but the process of construal is not only affected by the factors listed above. To put it another way, there are different ways to categorize the language phenomenon of construal, and any grouping of the dimensions of construal is primarily for purposes of discussion (Langacher, 2014, personal communication).

Additionally, in different theoretical frameworks, other scholars examine how construal works in linguistic communication. According to Talmy (1988), construal is the imaging system, and it is discussed from four dimensions, namely structural schematization, deployment of perspective, distribution of attention, and force dynamics, whereas Croft and Cruse (2004) analyze construal based on the aspects of attention/salience, judgement/comparison, perspective/situatedness, and constitution/gestalt.

Nevertheless, when Langacker, Talmy, and Croft and Cruse figure out the dimensions of construal, they care more about how an object is construed by a speaker but are less concerned with the dynamic interactions between language, speakers, and dialogic situations. Although Verhagen (2005, 2007) proposes a dialogic account of construal and emphasizes the cognitive coordination between speakers, he does not deal much with the interactions between object(s) being construed, the interaction between speakers, or how the co-focused object(s) can be grounded in conversation.

With the developmental view and taking account of the interactive nature of the process of construal, this study especially proposes dimensions of construal from a dialogic view, as suggested in **Table 1**, to examine how a co-focused object is construed in dialogic interaction.

Table 1 not only summarizes the (shared) features of construal at different phrases of the studies of construal in cognitive grammar, but also proposes the dimensions of construal that might be effective to describe how a dialogic focus is interpreted based on the interactive nature of construal

Table 1. A developmental view on the dimensions of construal

1	2	3	4	5
Langacker (1987, 1991, 1993, 1999)	Langacker (2007: 17)	Langacker (2008/2013: 55)	Langacker (2015: 120, 2019: 141)	A dialogic view
prominence	prominence	prominence	prominence	***prominence***
perspective	perspective	perspective	perspective	***perspective***
the level of specificity	level of specificity	specificity		***specificity***
scope		focusing		***focusing***
background			dynamicity	***dynamicity***
			selection	***selection***
			imagination	

operation in conversation.

3. Dialogic interaction in a typical wh-dialogue[1]

Human beings live in the world in a dialogic way (Bakhtin, 1981). Collingwood (1940[1998]: 23) even asserts that every statement that anybody ever makes is made in answer to a question. Questioning-answerings are the common dialogic phenomena in human beings' linguistic communication. In this research, the category of a dialogue consisting of a WH-question and one of its answers (WH-dialogue for short) is specially employed to discuss the dimensions of construal from a dialogic view.

In terms of the grammatical pattern, a WH-dialogue is prototypically structured by a WH-question whose regular pattern is "*WH-interrogative word* (WH-word for short) + *auxiliary* + *remainder*?", and an adjacent utterance *X*, the answer[2] to the question. The syntactic pattern of a WH-dialogue[3] can then be described as the following:

A WH-dialogue — *WH-word + auxiliary + remainder ?*
An answer[X]

In conversation, a WH-dialogue is employed by a speaker to verify a known message or to garner unknown information.

While a speaker is construing something, s/he is experiencing mental contact with it, demonstrating the interaction between the subject (human being) and the object (something being construed). The dialogic interaction is the way that the subject's conceptualization of something in the process of construing is shared by the dialogic partner(s) (cf. Couper-Kuhlen and Selting, 2018; Dąbrowska, 2014; Fischer, 2015; Hancil, 2018; Hsieh and Su, 2019; Linell, 2009, 2017; Jaszczolt, 2016; Săftoiu,

[1.] A WH-dialogue hereafter in this study is the case of an English wh-dialogue.

[2.] In this study, an answer is syntactically defined as the adjacent utterance immediately following the WH-question in a WH-dialogue.

[3.] Cases where there are more than one WH-word heading a question and where WH-questions are embedded in other sentences do not fall into the scope of this research and will be addressed in another paper.

2019; Weigand, 2017; Zeng, 2021), whereas in dialogue, as Figure 2 demonstrates, the object being construed is jointly focused by the speaker and the hearer. Prototypically, in a WH-dialogue the questioner and answerer work together to negotiate the specified content of the WH-word. What the questioner does is to select a particular WH-word to linguistically encode the unknown information in communication and then place the WH-word at the beginning of the question to attract the hearer's attention. Meanwhile, a WH-question indicates the construal frame(s) working as the background or scope for the answerer to specify the conceptual content of the WH-word. In the ongoing dialoguing process, questioners also evaluate the qualities of the answers, to justify whether the utterances hearers offer can serve as the expected answers to WH-words or questions. In this sense, the interactional dimension of construal in a WH-dialogue is characteristic of the process of questioning-responding as well as the answer-evaluation, wherein interlocutors, utterances, and the given situation are all participants in the dialogic interaction.

4. Construal of the shared focus in WH-dialogue

From a developmental perspective as viewed in Table 1, this study proposes several dimensions to describe the process of construing the dialogic focus in English wh-dialogue. These dimensions, namely ***level of specificity***, ***dynamicity***, ***selection***, ***perspective***, ***prominence***, and ***focusing of attention***, can be well elaborated in the grounding of wh-words[4] leading questions in these wh-dialogues.

4.1. Level of specificity

In a wh-dialogue, the wh-word (e.g. what, when, who) placed at the beginning of a wh-question essentially indicates a certain category of element participating in structuring an event or a scene. For instance, "*what*" basically refers to *something* that participates in constructing an event or a scene, and "*where*" often denotes *some place* where an event occurs or a scene exists. In other words, in this type of dialogue the semantic content of the wh-word is not specific. Thus, the wh-word heading a wh-question is schematic in nature[5]. Thinking in this way, in terms of the semantic content, a wh-question is also a schematic frame, while the answer to a wh-question is prototypically more specific. From the cognitive view of the schema-instance principle (cf. Langacker, 1987: 373; Taylor, 2002: 125), the pair of a wh-question and the answer to it can therefore be viewed as a whole or a unit, thus acquiring the status of a *dialogic construction* with its own form and function, in accordance with Goldberg's (1995: 4, 2006: 3) definition of a construction (cf. Brône and Zima, 2014; Nikiforidou et al., 2014; Zeng, 2016). That is to say, the answer in a wh-dialogue is quintessentially an instance of a schematic frame that is linguistically encoded by a wh-question. Accordingly, both the wh-question and its answer are in fact the linguistic representations for the conceptualization of the same schema but with different degrees of specificity, as displayed in **Figure 3**.

Figure 3 demonstrates that the wh-word or the wh-question is a schema whose instance is its adjacent answer. Grounded in the schema-instance relation, the question and the answer in a wh-dialogue signify the different levels of schematicity of the same conceptual structure.

[4] In this study, a "wh-word" especially refers to the single question word positioned at the head of a wh-question.

[5] R.W. Langacker shares the same view in my personal communication with him by email on April 2, 2015: "Yes, apart from the difference between questioning and stating, a wh-word is schematic with respect to its possible answers".

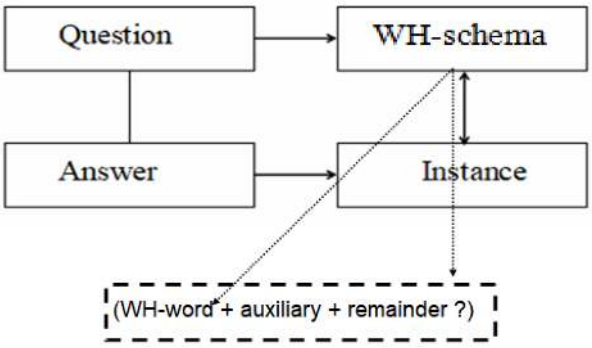


Figure 3. The schema-instance relation in a wh-dialogic construction.

4.2. Dynamicity

The schema-instance relation between a wh-question and its answer suggests the dynamicity of the grounding of wh-word heading the question. Grounding is the process through which an abstract *TYPE* concept is specified as a *CONCRETE* example in the ground that consists of such elements as the speech event, its participants (speaker and hearer), their interaction, and the immediate circumstances (notably, the time and place of speaking) (cf. Langacker, 2008: 259). In a wh-dialogue, the schematic wh-word fundamentally designates a type concept that is the participant to structure a wh-question which is also schematic in nature as a whole. Hence, to answer a wh-question is in essence the dynamic process in which the answerer searches, compares, and then selects a proper instance for the schematic wh-word, as diagramed in **Figure 4**.

Figure 4 shows that the wh-word heading a question uttered by speaker 1 in fact provides a domain of instantiation, while the answer offered by speaker 2[6] is one (group) of the possible members in this domain. The answer denoting the instantiation of a wh-question means that both speakers 1 and 2 have established mental contacts with the specific member of the type concept indicated by the wh-word, then signifying the successful grounding of the dialogic focus, whereas for the type of wh-dialogues with answers implying non-instantiation of wh-words, the wh-words or the default dialogic focuses are not eventually specified in communication.

4.3. Selection

In a wh-dialogue, the wh-question serves as the structural basis for the following adjacent utterance that is supposed to be an answer. In such a conversation, when constructing an answer, the hearer might selectively reproduce some or all of the linguistic resources including the words, grammatical patterns, prosodies, or functions of the linguistic signs that are previously employed in the question. The reproduced linguistic resources then bring about parallelism between the question and the answer. Consequently, the structural mappings from the answer to the question exhibit structural affinities, thereby producing dialogic resonance between the two utterances. The structural symmetries that emerge in a local wh-dialogue indicate that the basic way human beings produce language in conversations is to take language to make language (cf. Du Bois, 2014: 359).

4.4. Perspective

Normally, speakers with different life experiences might have similar or different interpretations

[6.] For this research, the questioner (S1) and the answerer (S2) are not the same person.

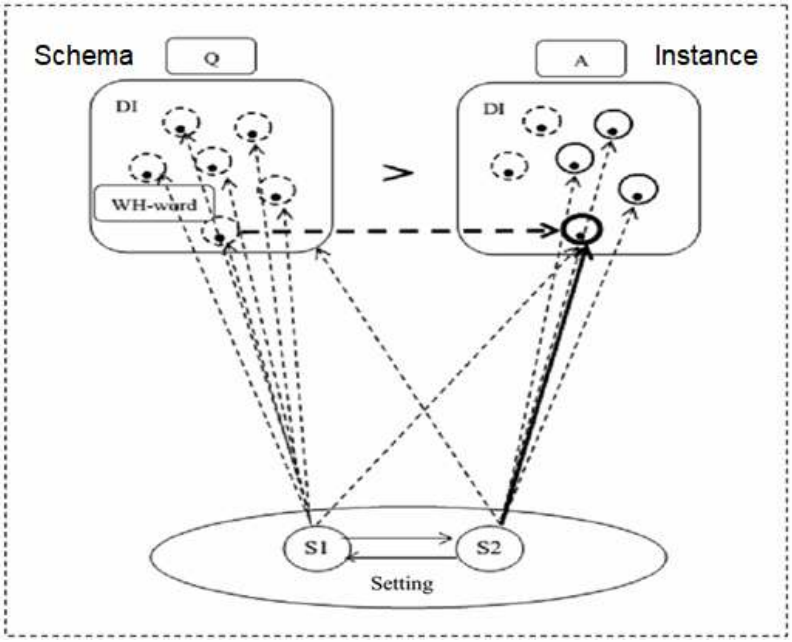


Figure 4. The grounding of schematic wh-word in a wh-dialogic construction.

S1-Speaker 1; S2-Speaker 2; Q-Question; A-Answer. (adapted from Zeng, 2019: 290).

concerning the same event or object being construed, implying that the aspect of perspective functions in the construal operation. It is often the case that in a wh-dialogue, concerning the schematic conceptual content encoded by the wh-word, the questioner and the respondent do not share the same understanding. In particular cases, answerers, who have their own interpretations of the objective reality, possibly take stances distinctive from those of questioners, hence elaborating the schematic wh-word in diverse ways.

4.5. Prominence

Structurally speaking, the head or the end of an utterance, or the stress in intonation could be the syntactic indication of the salient conceptual structure of an event or a scene in the speaker's mind. The most striking structural feature of a regular English wh-question is typically the position of the wh-word, which is at the head of the question utterance and supposed to draw most of the answerer's attention. From the cognitive linguistic view, a wh-word applied to initiate a question, such as *what*, *how*, *when* or *why,* defines a certain salient aspect of an event, viz *what* being the participant(s) in the event, *how* being the way, *when* being the time, and *where* being the place an event occurs. With regard to the cases where the wh-word in a wh-dialogue is successfully grounded (as shown in Figure 4), the wh-word is mentally salient in both speakers' minds, while for the cases with ungrounded wh-words, wh-words are not prominent in the mental world of the answerer.

4.6. Focusing of attention

In a wh-dialogue, the default dialogic focus is the wh-word heading the question. Since the interlocutors are likely to have different perspectives to instantiate the schematic content of the wh-word, the focal information in the question and answer could be matched or unmatched. For the former case, the questioner and the answerer share the dialogic focal information, which is prototypically abstract in the question and more specific in the answer. For the latter one, the focal structure of the

answer is not consistent with that of the question, suggesting that the wh-word is not grounded in the dialogue.

5. Dynamic focal adjustment in wh-dialogic constructions

As shown in Figure 2, the dialogic setting or ground basically consists of the speech event, its participants (speaker 1 and speaker 2) and their interactions, the event linguistically encoded by a wh-question, and the immediate circumstances (especially the time and place of speaking). In a given situation, to start a wh-dialogue, speaker 1 sets the dialogic focus as well as the background for construing it. Speaker 2, based on his or her life experience and knowledge, takes the same or different perspective towards the prominent conceptual content (viz. the wh-word) of the question. For speaker 2, s/he might directly or indirectly provide the answer to the dialogic focus, or s/he does not speak of anything related to the dialogic focus but introduces a new topic, or even shows negative attitude towards the talking and intends to end the conversation. As such, the types of answers in wh-dialogues naturally indicate the dynamic adjustment of dialogic focus in wh-dialogues.

5.1. Focal adjustment in a single wh-dialogue

In a single wh-dialogue, the answer could or could not be the instance of the wh-word positioned at the beginning of a wh-question, implying there are different types of focal adjustment in such a conversation.

5.1.1. Dialogic focus specified

In the dialogic interaction, when the answer contains the grammatical structure that semantically corresponds to the wh-word of the question, the dialogic focus is then specified. In regard to these cases, the schematic wh-word is directly specified by its instance (viz. the answer), as exemplified in the diagraph[7] of dialogue (1).

Diagraph of dialogue (1)[8]

Question	CHRIS-CUOMO-1-ABC#(Off-camera):	… **What**	do you want for Christmas	?
Answer	AUDIENCE-MEMBER-1#:	**Diamonds**		.

In this dialogue, the default dialogic focus is set as "***what***", which suggests a schematic category of noun indicating THING being wanted by the answerer for Christmas. The answer is structured by only a noun with its plural form "***Diamonds***", demonstrating the semantic mapping from the answer to the wh-word. To be precise, the answer here serves as a group of instances in the domain of instantiation. Hence, the schema-instance relation is overtly manifested between the wh-word and the answer, suggesting that the abstract dialogic focus is specified as concrete entities in this single wh-dialogue.

[7] "Diagraph" employed by Du Bois (2014) is used in this paper to indicate the dialogic resonance between utterances.

[8] All the dialogue examples are selected from COCA. The utterances before or after the single dialogue are omitted and indicated by three dots. The bold black structures in the dialogues marked in the diagraph show the syntactic and semantic correspondences between the question and the answer.

Dialogue (2) is another case displaying the dialogic focus specified as an instance indirectly, as revealed in its diagraph.

Dialogue (2)

Question	Mr-BUDZYN:	… **Why**	am I here	**?**
Answer	BRADLEY:	***If* you never hit him at all, *how is it that* you were** charged and convicted of second-degree murder		?

In this short conversation, the reason for the occurrence of the scene "I am here" is the default dialogic focus encoded by "***why***". The answer, however, is not uttered with the assertive or positive tone that demonstrates the specification of "***why***", but is structured with a rhetorical question, based on which speaker 1 infers the instance of the schematic "***why***", namely "***because you hit him***" and then "***you are charged and convicted of second-degree murder***". Regarding this type of dialogue, the answer could be taken as the extended member from the prototypical answer of the schema "***why***", whereas the prototypical member of the schema, also the potential direct answer, can be implied from the speaker 2's utterance(s), as **Figure 5** illustrates.

Figure 5 informs us that the wh-schema is instantiated by the extended member marked by the solid line with bi-directional arrows. The dotted lines with uni-directional arrows mean that the prototypical member of the wh-schema is emergent in the inference process based on the extended member. In such a conversation, the schema-instance relation is implicitly shown though, the dialogic focus is indeed specified.

5.1.2. Dialogic focus shifted

In terms of the case where the dialogic focus is shifted, speaker 2 does not offer any information that could be interpreted as the instantiation of the wh-word. The original dialogic focus established in the question is then shifted to a new one introduced by speaker 2, suggesting that there is no overt or covert schema-instance relation between the wh-question and its adjacent utterance. Dialogue (3) is an example.

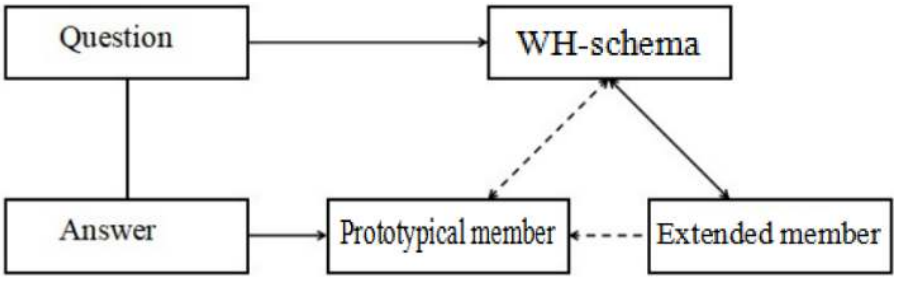


Figure 5. The relation between a wh-schema, a prototypical member, and an extended member.

Dialogue (3)

Question	NEAL-CONAN:	…	**How**	did	Taylor	defend	the indefensible	?
Answer	MATT-STEINGLASS:	**Well, the question is more**	**how**		is laywer, Courtenay Griffiths,	defended		.

This diagraph shows that the questioner inquires how the event "***Taylor defended the indefensible***" occurred, whereas speaker 2 does not provide the detail concerning "***the way***", but describes a new event "***his lawyer, Courtenay Griffiths, defended***". By doing so, speaker 2 sets a new dialogic focus, *the way the new event came into being*. Accordingly, no semantic correspondence is demonstrated between the utterance, the so-called answer, and the default dialogic focus "***how***" that has in fact been shifted to the salient structure of another event, namely a new "***how***". For this group of dialogues, the utterances by speaker 2 can be viewed as the instances of new schemas triggered in the dialogue, as seen in **Figure 6**.

Figure 6 shows that the question is followed by an utterance by speaker 2 and the two lines of utterances compose an adjacent pair only in form, as is designated by the solid line without bi-directional arrows. That is, what speaker 2 utters is not the instance of Schema A, viz. the wh-word or the wh-question, whose instance is in fact not provided in this dialogue. The dotted line with arrows implies the absent interactional relation between utterances. Nevertheless, speaker 2's utterance can be viewed as the instance of schema B, as shown by the solid line with bi-directional arrows. The new schema might function as the template in the anticipated talk turns.

5.1.3. Dialogic focus removed

Any aspect of an event or a scene could be made prominent and then linguistically encoded by a wh-word at the head of an English wh-question. In on-line linguistic communication, it might be the case that the dialogic focus is negated or doubted by the answerer, for whom the salient part of the questioned event or scene is not the right or proper one semantically marked by the wh-word at the beginning of the question, but something else. Dialogue (4) is one of the cases.

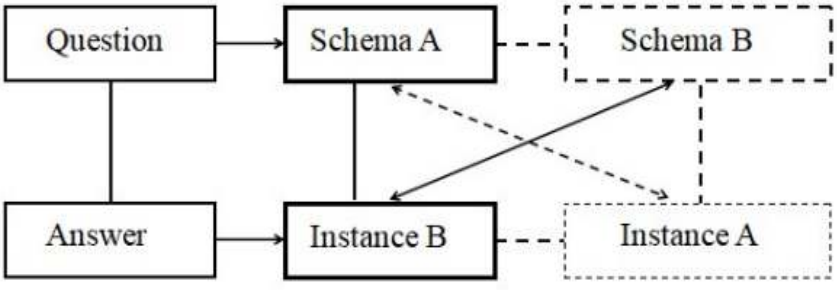


Figure 6. The new schema-instance relation triggered in a dialogic-focus-shifted dialogue.

Dialogue (4)

Question	KOPPEL:	… **Where**	do	you	expect this sort of	uprising,	if there is to be any,	against Saddam Hussein to originate	?
Answer	Vice Pres. QUAYLE:			I	am not saying that		there is going to be an		
						uprising			.

Obviously, in this dialogue, the structural mappings from the answer to the question show no indication of syntactic and semantic correspondences between the answer and the wh-word "***where***", but we can see a portion of the non-wh-word part of the question is re-produced in the answer. To be specific, what the questioner most cares about is ***the place*** *where the event relating to "uprising" will occur*, whereas the answerer does not share the same dialogic focus with the questioner but negates the occurrence of the event that is linguistically encoded by "uprising", hence removing the original dialogic focus, the "***where***", in the question.

5.1.4. Dialogic focus discontinued

In principle, the question in a wh-dialogic construction hints at the domain of instantiation of the dialogic focus. For pairs of a wh-question and an answer, there are examples wherein the answers only partially instantiate the schematic content of wh-words. To put it simply, at the end of a wh-dialogue, a wh-word is not fully instantiated, and more detailed information concerning the wh-word is expected to be offered in the ongoing conversation. In this account, the elaboration process of the dialogic focus is discontinued. Dialogue (5) is a case in point.

Dialogue (5)

Question	UNIDENTIFIED PERSON:	… **What**	are	the pros	and cons	of	going from one method to the other?	?
Answer	MR-HARDY:			**The pros**		of	being independent	.
			is					
		that you control your product from start to finish						

This diagraph demonstrates that the instance of the dialogic focus "***what***" should include two components, one of which is the specified ***pros*** and the other the detailed ***cons*** of "***going from one method to the other***". However, the instance of "***what***" indicated by the answer only elaborates "***the pros***", without providing any particular information on "***the cons***" that is also cared about by the questioner. Viewed in this way, the instantiation of the schematic "***what***" is merely half done or discontinued.

5.1.5. Dialogic focus suspended

Cases of dialogic focus suspended refer to those wh-dialogues in which both syntactic and semantic correspondences are exhibited between the question and the answer, but certain grammatical structure in the answer functions to negate the emergent schema-instance relation between the wh-word and the answer, implying that the answer *is not* or *does not contain* the qualified instance of the wh-word, the dialogic focus. The conversation ends but the dialogic focus is not successfully grounded in the dialogic interaction, as exemplified in dialogue (6).

Dialogue (6)

Question	Ms-LIEBERMAN:		... **Who**	would use that AI	?
Answer	ROKER:	**Not**	***me***		.

With regard to this dialogue, the personal pronoun "***me***" in the answer is in the domain of instantiation delimited by the dialogic focus "***who***", displaying the schema-instance relation between the question and the answer. Nonetheless, the marker of negation "***not***" in the answer negates the qualification of "***me***" as the instance of "***who***", suggesting that the dialogic focus "***who***" is ultimately ungrounded in the dialogic process.

Still notable is that the above mentioned wh-dialogues with dialogic focuses shifted, removed, discontinued, and suspended share the common feature that qualified instances of wh-words are not contained in answers, whereby they are essentially wh-dialogues characterized by ***non-instantiation of wh-words***. The potential schema-instance relation in these dialogues can be depicted in **Figure 7**.

Figure 7 suggests that a wh-question and its adjacent utterance constitute a single wh-dialogue merely at the syntactical level (represented by the solid line without arrows), whereas the part or the whole answer utterance is not semantically correspondent to the wh-word or the wh-question. Consequently, the schematic content of the wh-word is not specified in any part of the answer (indicated by the box with dotted lines), although there might be parallelism emergent between the utterances.

5.2. Focal adjustment in a series of wh-dialogues

Grounded in the immanent schema-instance relation between a wh-question and its answer, section 5.1. depicts the adjustment of dialogic focus in single wh-dialogues. For this sub-section, a series of wh-dialogues consisting of one question but with different answers are investigated, to un-

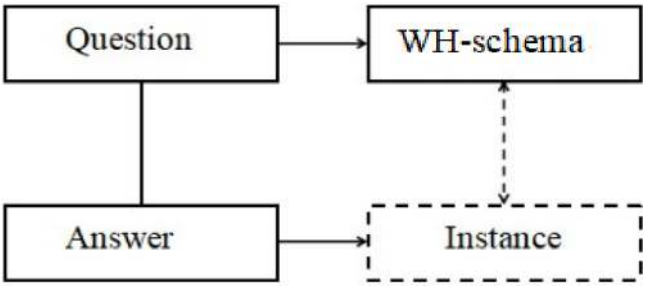
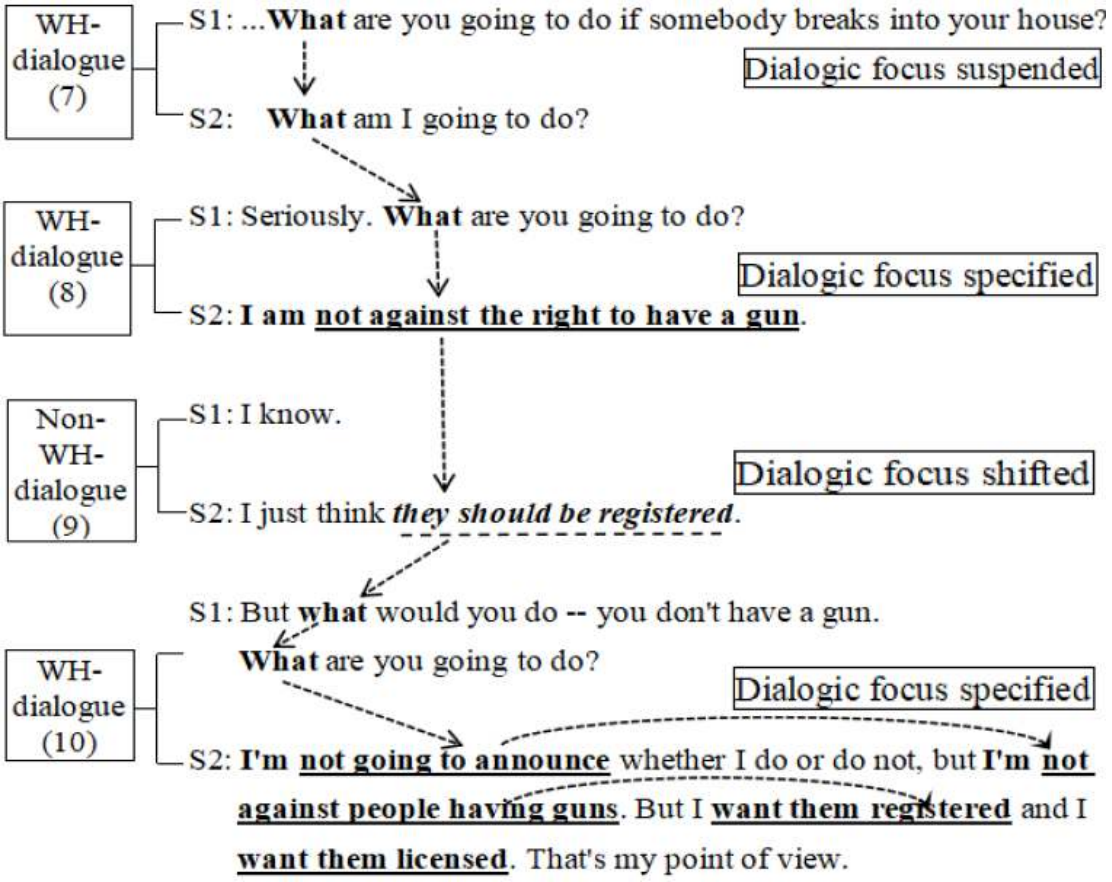


Figure 7. The potential schema-instance relation in wh-dialogues featuring non-instantiation of wh-words.

ravel the dynamic focal adjustment, as displayed in the discourse space consisting of wh-dialogues (7), (8), (10), and the non-wh-dialogue (9).



(S1-speaker 1; S2-speaker 2)

In this local discourse, S1 interacts with S2 to negotiate the specific instance of the dialogic focus encoded as "***what***" in the question "***what are you going to do***?". To start with, S1 desires to garner the information on the future event that S2 is presumably to describe on the condition that somebody breaks S2's house, with ***what*** working as the initial dialogic focus.

Strikingly, S2 does not elaborate the schema-instance relationship between the question and the answer in dialogue (7), but restates the schematic structure of the future event. We can see that the dialogic focus is unspecified and then suspended in this talk turn.

With the intention of cooperating with S2 for the successful grounding of "***what***" in the initial question, S1 iterates the abstract structure of the future event in which S2 is supposed to participate in. By doing so, the dialogic focus "***what***" is clarified for a second time. In response to S1's concern, S2 narrows down and at the same time defines the domain of instantiation of the dialogic focus "***what***", which includes any instance of a schematic event that is characteristic of "***being not against the right to have a gun***". In this process of question-response, namely dialogue (8), the dialogic focus "***what***" is more specific than that in dialogue (7).

As we can see, S2's response in dialogue (8) receives positive feedback from S1, indicated by S1's utterance "***I know***", which then motivates S2's further elaboration on the negated instance of

"***what***", by uttering "***they*** (the guns) ***should be registered***". The interaction between these two utterances denotes a non-wh-dialogue (9), and the dialogic focus is now shifted from S1's concern of the future event that S2 is engaged to the new event that S2 cares about and is related to the guns' registration.

While, in the ongoing dialogic interaction, without receiving positive answers from S2, S1 claims the dialogic focus "***what***" once again and the schematic future event that has been mentioned in dialogues (7) and (8). S1 also adds more background information, viz. "***you don't have a gun***" in the case that "***somebody breaks into your house***", to help S2 construe the situation-specific meaning of ***what***. What's more, S1 repeats the same question "***what are you going to do***?" for a third time, showing again S1's intention to cooperate with S2 for the successful grounding of the dialogic focus "***what***" that is originally made prominent in dialogues (7) and (8).

Following that, S2 constructs new utterances consisting of both negative and positive statements, which are the detailed conceptual content of "***what***". Subsequently, wh-dialogue (10) is structured. The expression "***this is my point of view***" by S2 implies that the dialogic focus "***what***" initiated in dialogue (7) is then definitely elaborated with specification in dialogue (10).

In brief, in this local discourse, the utterances by S2 display different degrees of specification of the dialogic focus "***what***", significantly demonstrating the dynamic focal adjustment in the wh-dialogues.

6. Conclusions

Based upon Langacker's analyses of the dimensions of construal operation at different phrases of his studies on cognitive grammar, this paper proposes a dialogic view on the dimensions of construal. These dimensions, including ***level of specificity***, ***dynamicity***, ***selection***, ***perspective***, ***focus of attention***, and ***prominence***, are elaborated in detail with the evidence shown in English wh-dialogues. Following this, the types of dynamic focal adjustment in wh-dialogues are investigated, and their classifications can be summarized in **Table 2**.

As Table 2 indicates, the ***specificity*** in construal operation in a dialogic situation is described in terms of the schema-instance relation, based on which English wh-dialogues might be roughly categorized as (1) dialogues with direct instantiation of the wh-word, (2) dialogues with indirect instantiation of the wh-word, and (3) dialogues with non-instantiation of the wh-word. The ***dynamicity*** aspect of construal refers to the grounding process of the schematic dialogic focus in situation-specific conversation, implying two groups of wh-dialogues, one with wh-words grounded, the other with ungrounded wh-words. In wh-dialogues, part or all of the linguistic resources in questions might be reproduced in the adjacent answers, suggesting explicit or implicit structural parallelism between questions and answers, exemplifying the dimension of ***selection*** in construal.

Other cases of wh-dialogues worthy of our attention are the dialogues where interlocutors might take the same, similar or different ***perspectives*** to interpret the dialogic focus, leading to the fact that there could be different answers to one wh-question, as evidenced in the local dialogues from (7) to (10), while ***focus of attention***, another dimension of construal in dialogic interaction, implies the central attention allocated by interlocutors in conversation. As such, there are cases of wh-dialogues with matched or unmatched dialogic focuses between questions and answers. As for ***prominence***, it

Table 2. A dialogic view on construal operation and the focal adjustment in wh-dialogues

specificity	schema-instance		schema-(instance)-new schema			
	direct instantiation	indirect instantiation	non-instantiation			
dynamicity	dialogic focus grounded		dialogic focus ungrounded			
selection	previous language resources reproduced/explicit or implicit parallelism displayed					
perspective	similar or the same		different			
focusing of attention	dialogic focus matched		dialogic focus unmatched			
prominence	wh-word salient in the mental worlds of both S1 and S2		wh-word non-salient in the mental world of S2			
the focal adjustment	focus specified		focus shifted	focus removed	focus discontinued	focus suspended

signifies the salient component of the conceptual structure(s) in the mental worlds of the speakers.

The dimensions stated above essentially reveal the interactive nature of construal operation in conversation, in particular indicating the interrelation among speakers, utterances, and the specific dialogic situation. Grounded on these dimensions of construal, the focal adjustment in wh-dialogues could be classified into five groups, namely cases of dialogic focus *specified*, dialogic focus *shifted*, dialogic focus *removed*, dialogic focus *discontinued*, and dialogic focus *suspended*. This study is hopefully to shed some light on the understanding of utterance meaning from a new perspective, that is, a dialogic view on the meaning co-construction in conversation.

Conflict of interest

The author declared no conflict of interest.

Acknowledgements

This work was supported by the National Social Science Fund of China (grant number: 18BYY076) and the fund from the China Scholarship Council (CSC grant number: 202006240343) for my research at the University of Cambridge.

References

Bakhtin MM (1981) *The Dialogic Imagination: Four Essays*. Holquist M (ed.). Emerson C and Holquist M (trans.). Austin and London: University of Texas Press.

Brône G and Zima E (2014) Towards a dialogic construction grammar: *Ad hoc* routines and resonance activation. *Cognitive Linguistics* 25(3): 457–495.

Collingwood RG (1998[1940]) *An Essay on Metaphysics*. Oxford: Clarendon Press.

Croft W and Cruse DA (2004) *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Couper-Kuhlen E and Selting M (2018) *Interactional Linguistics: An Introduction to Language in Social Interaction.* Cambridge: Cambridge University Press.

Dąbrowska E (2014) Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics* 25(4): 617–653.

Du Bois JW (2014) Towards a dialogic syntax. *Cognitive Linguistics* 25(3): 359–410.

Fischer K (2015) Conversation, construction grammar, and cognition. *Language and Cognition* 7(4): 563–588.

Goldberg AE (1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago and London: The University of Chicago Press.

Goldberg AE (2006) *Constructions at Work: The Nature of Generalization in Language*. Oxoford: OUP.

Hancil S (2018) Discourse coherence and intersubjectivity: The development of final but in dialogues. *Language Sciences* 68: 78–93.

Hsieh CYC and Su LIW (2019) Construction in conversation: An interactional construction grammar approach to the use of xiangshuo 'think' in spoken Taiwan Mandarin. *Review of Cognitive Linguistics* 17(1): 131–154.

Jaszczolt KM (2016) *Meaning in Linguistic Interaction: Semantics, Metasemantics, Philosophy of Language.* Oxford: Oxford University Press.

Langacker RW (1987) *Foundations of Cognitive Grammar vol. I: Theoretical Prerequisites.* Stanford, California: Stanford University Press.

Langacker RW (1991) *Foundations of Cognitive Grammar vol. II: Descriptive Application.* Stanford, California: Stanford University Press.

Langacker RW (1993) Universals of construal. P*roceedings of the Annual Meeting of the Berkeley Linguistics Society* 19(1): 447–463.

Langacker RW (1999) *Grammar and Conceptualization (Cognitive Linguistics Research)*. Berlin and New York: Mouton de Gruyter.

Langacker RW (2007) *Ten Lectures on Cognitive Grammar*. Beijing: Foreign Language Teaching and Research Press.

Langacker RW (2008/2013) *Cognitive Grammar: A Basic Introduction*. Oxford: OUP.

Langacker RW (2015) Construal. In: Dąbrowska E and Divjak D (eds.) *Handbook of Cognitive Linguistics*. Berlin and Boston: De Gruyter Mouton, pp.120–143.

Langacker RW (2019) Construal. In: Dąbrowska E and Divjak D (eds.) *Cognitive Linguistics: Foundations of Language*. Berlin/Boston: Walter de Gruyter GmbH, pp.140–166.

Linell P (2009) *Rethinking Language, Mind, and the World Dialogically: Interactional and Contextual Theories of Human Sense-Making.* Charlotte, NC: Information Age Publishing, INC.

Linell P (2017) Dialogue, dialogicality and interactivity. *Language and Dialogue* 7(3): 301–335.

Nikiforidou K, Marmaridou S, and Mikros GK (2014) What's in a dialogic construction? A constructional approach to polysemy and the grammar of challenge. *Cognitive Linguistics* 25(4): 655–699.

Săftoiu R (2019) The dialogic turn in language study: A review of *The Routledge handbook of language and dialogue** (edited by Edda Weigand 2017) and *From Pragmatics to Dialogue* (edited by Edda Weigand and Istvan Kecskes 2018). *Language and Dialogue* 9(3): 471–483.

Talmy L (1988) The relation of grammar to cognition. In: Rudzka-Ostyn B (ed.) *Topics in Cognitive Linguistics*. Amsterdam: John Benjamins, pp.165–205.

Taylor J (2002) *Cognitive Grammar.* Oxford: OUP.

Verhagen A (2005) *Constructions of Intersubjectivity: Discourse, Syntax and Cognition*. Oxford: Oxford University Press.

Verhagen A (2007) Construal and perspectivisation. In: Geeraerts D and Cuyckens H (eds.) *Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, pp.48–81.

Weigand E (ed.) (2017) *The Routledge Handbook of Language and Dialogue*. New York and London: Routledge.

Zeng G (2016) A cognitive approach to the event structures of wh-dialogic constructions. *Southern African Linguistics and Applied Language Studies* 34(4): 311–324.

Zeng G (2019) *Dialogic Construction Grammar: A Theoretical Frame and Its Applicatio*n. Chengdu: Sichuan University Press.

Zeng G (2021) Repetition in Mandarin-speaking children's dialogs: Its distribution and structural dimensions. *Linguistics Vanguard* 7(1): 1–14.

ORIGINAL ARTICLE

A social-cognitive research on metaphor and metonymy: Taking age appellations in Chinese as an example

Yan Zheng[1,2*]

[1] Sichuan International Studies University, Chongqing, China
[2] Lanzhou Jiaotong University, Gansu, China

Abstract: The process from individual language to communal language, from individual experience to communal embodiment, and from individual cognition to collective cognition is based on human embodiment and social interactions. Metaphor and metonymy are two basic cognitive approaches to the world. Thus, human cognition is metaphorical and metonymic in essence, and metaphonymy is embodied and social in nature. In the perspective of social embodied cognition, this paper, taking social metaphonymy as a theoretical foundation and the Chinese age appellations as the corpus, demonstrates that Chinese age appellation has eight metaphonymic categories, namely, physiology, gender, wisdom, accessories, raising, behavior, animal and plant, and number. The paper, by expounding the embodiment and metaphonymy, aims to reveal metaphonymic cognitive mechanism of Chinese age appellations.

Keywords: embodiment; sociality; metaphonymy; age appellation

***Corresponding author:** Yan Zheng, Sichuan International Studies University, Zhuangzhi Road 33, Shapingba District, Chongqing, China; School of Foreign Languages of Lanzhou Jiaotong University, Anning West Road 88, Lanzhou, Gansu Province, China; deborahzheng@mail.lzjtu.cn

Received: April 20, 2022; **Accepted:** June 11, 2022; **Available online:** September 16, 2022

Citation: Zheng Y (2022) A social-cognitive research on metaphor and metonymy: Taking age appellations in Chinese as an example. *Forum for Linguistic Studies* 4(1): 66–77. DOI: 10.18063/fls.v4i1.1456

1. Introduction

Human beings are social animals, and sociality is one of the essential attributes of human beings. The transition from individual language to communal language is based on this attribute which is also the process from individual experience to communal embodiment, from individual cognition to collective cognition. The prerequisite for language to undertake the communicative and cognitive functions is intersubjectivity (Zlatev, 2007; Cheng, 2009; Zhao, 2015) which means when performing cognitive processing on a specific object, humans must reach an agreement in a specific community to form "knowledge", which will then become the basis of human communication, shared by all members of the society. Metaphonymy is the basic means of human cognition. Therefore, human cognition is metaphorical and metonymical, and human sociality determines that metaphonymy is social in nature. Age appellation is an important research field of sociolinguistics. In the perspective

of social embodied cognition, this paper, taking the Chinese age appellations as the corpus (mainly from the National Language Commission Corpus of Ancient Chinese), demonstrates that Chinese age appellation has eight metaphonymic categories, namely, physiology, gender, wisdom, accessories, raising, behavior, animal and plant, and number, which expounds the embodiment of metaphonymy and reveals the metaphonymic cognitive mechanism of Chinese age appellations.

2. Sociality of metaphor and metonymy

Lakoff and Johnson (1980) subverted the traditional view of metaphor and metonymy pioneered by Aristotle in *The Metaphors We Live by*, arguing that metaphors are not just rhetorical devices, but more importantly, the basic cognitive style of human beings. Later, Lakoff and Johnson (1999) discussed the embodiment of metaphor and metonymy based on the system of experienced philosophy. Panther and Radden (2002: 1) believe that metonymy is more basic than metaphor to some extent. Although scholars still have different views on the relationship and cognitive mechanism of metaphonymy, at least the following four points of agreement have been reached:

(1) Metaphor and metonymy are the basic ways for human beings to understand the world;

(2) Metaphor and metonymy are the basic modes of human thinking;

(3) Metaphor and metonymy are the result of human experience and cognitive processing;

(4) Both metaphor and metonymy try to establish the connection or relevance between cognitive objects.

As far as the basic cognitive styles of human beings are concerned, metaphor and metonymy play important roles in human cognition and the external world (Lakoff and Johnson, 1980, 1999), which form the two basic cognitive ways for human beings to recognize themselves and the world. This paper holds that metaphor and metonymy have both connections and differences. Metaphor establishes the cognitive connections between the source domain and the target domain based on iconicity, while metonymy uses one cognitive domain to replace another based on relevance. For example, in terms of age appellations, the ancients used “cardamom” to refer to “13- to 14-year-old girls”. The basic cognitive methods and cognitive processing steps are as follows: firstly, in people’s cognitive experience, “cardamom” is very “young and tender”, while 13- to 14-year-old girls are also in the stage of “immature blooming season”. Using metaphorical mechanism and conceptual projection, the cognitive relationship between “cardamom” and “13- to 14-year-old girls” is established with their “immature” iconicity; in the second step, people use metonymy to refer to one part and another based on the relevance of the two subjects. In this case, “cardamom” is used to refer to “girls aged 13 to 14”.

In terms of the social attribute of cognition, we have to emphasize that human cognition of embodiment is not the cognition of individual members, but the collective cognition of people in a specific community. Therefore, as a social animal, human’s cognitive embodiment must be marked with social brand, that is to say, the embodiment is social. Specifically, human thinking is shaped by practical activities, consciousness, and the material environment. As a part of the whole, the three are contained within each other (Lantolf and Qin, 2018: 1). The formation of the conceptual mechanism of human metaphor and metonymy is the process by which human beings experience and

process social reality and reach consensus in the long-term productive labor (Lindblom, 2015; Wen, 2019). That is to say, the cognitive mechanism of metaphonymy is not an individual behavior but a collective behavior, which implies the sociality of metaphonymy. For example, the ancients believed that "boys under teething in eight months will grow permanent teeth when they are eight years old; girls under teething in seven months will grow permanent teeth when they are seven". Based on this cognitive experience, the ancients used the mechanism of metaphonymy in terms of the permanent teeth period to refer to children aged 7 to 8. In the long-term cognition of embodiment, based on the specific social reality, especially the application of these intelligent complex age appellations in scholars' literary works and statements, these expressions finally are spread, developed and accepted in Chinese language and Chinese culture community, which also reflects the adjustment of society to the cognitive experience of cognitive subjects and finally helps the formation of intersubjectivity which becomes the basis and common ground of interaction between humans. At the same time, this has made the complex age appellations and profound cognitive mechanism behind them an important content of Chinese traditional culture.

3. Cognitive mechanisms of metaphor and metonymy in age appellations

With the historical development, the ancients have continuously established and developed the Chinese age appellations, which is the manifestation of both ancients' wisdom and their cognition upon themselves and the outside world. Based on the cognitive analysis of the corpus in this paper, the age appellations in Chinese reflect that the ancients have adopted eight types of cognitive mechanism on metaphonymy, which is also an important part of Chinese traditional culture, in an attempt to recognize age based on the specific social context and cognitive embodiment. The following will expound on them respectively.

3.1. Metaphonymy of physiological attributes

Embodiment is the basic way for humans to perceive the world. Human embodiment is the main source and foundation of human knowledge (Lakoff and Johnson, 1999; Shapiro, 2019; Wang, 2014, 2015; Winter, 2019). The human being, as a complex biological system, is material. Metabolism is the natural law of human being. Being a complex biological system, the human body is composed of many organs. In the human life cycle, people will show some physiological attributes at a specific age. According to Shen (2009), from a cognitive perspective, when human beings use mechanism of metaphonymy, the most important attribute is often apparent and becomes the first choice to represent the whole. Therefore, in the long-term development of Chinese culture, age is often connected with physiological attributes by means of metaphor and metonymy, thus creating a rich representation system of physiological attributes of age. There are eight such expressions in the corpus of this study, as shown in **Table 1**.

Generally speaking, when girls are 7 years old and boys are 8 years old, their deciduous teeth fall off and permanent teeth begin to grow. The ancients called this stage *tiao* or *chen*. Based on this physiological attribute, the year of *tiao* or *chen* refers to children aged 7–8 years old, just as the examples (1) and (2) show. Conversely, the examples of age appellations in these literary works also promote the spread of these expression in Chinese and construct their status in age appellations and Chinese culture.

Table 1. A list of metaphonymy of physiological attributes in age appellations

Number	Appellation	Age	Example	Source
1	*chi zi* 赤子	Newborn	If you protect your people as if they were newborn babies, they will be happy and obedient.	*The Book of History · Kang Gao*
2	*dao* 悼	Seven years old	Children aged 7 can be called *dào*. Even if they committed a crime, there was no punishment, for they had poor knowledge of the law and did not break it intentionally.	*The Rites·Qu Li*
3	*tiao chen* 龆龀	Seven to eight years old	Boys under teething in eight months will grow permanent teeth when they are eight years old; girls under teething in seven months will grow permanent teeth at seven.	*Shuowen Jiezi*
4	*hui chi* 毁齿	Seven years old	Girls under teething in seven months will grow permanent teeth when they are seven years old.	*Records of the Grand Historian*
5	*ni* 齯	The aged	The aged and the young make a family rich.	*Complete Tang Poems*
6	*hua fa* 华发	The aged	The aged has treasurable experiences.	*Hou Han Shu*
7	*gu xi* 古稀	70 years old	It's common to wine on credit but rare for any to live to 70 years.	*Qu Jiang*
8	*mao die* 耄耋	80 to 90 years old	It is quite satisfying for anyone to live up to 80 to 90 years old.	*Strange Tales from a Lonely Studio*

(1) Boys under teething at eight months old will grow permanent teeth when they are eight. (*The Unauthorized Biography of Han Poem*)

(2) Boys who haven't lost deciduous teeth cannot be soldiers. (*Guan Zi*)

In addition, *chen* has the same meaning as "losing teeth" in Chinese. Other expressions related to this kind of expression are: *chen tiao*, which means baby teeth falling off and unbound hair, is used to refer to childhood; *shi chen* refers to a child of seven or eight years who has just lost his baby teeth (An, 1987: 74).

In contrast, when an old man's teeth are worn out and then he spits out fine teeth like children's, this period of old people can be addressed as *ni* (Ma, 2007: 54). The ancients used this metonymy mechanism of physiological characteristics to refer to the elderly. Also, it was taken as a symbol of longevity, as shown in example (3):

(3) The elderly and children play with each other to celebrate and the whole family prospers. (*Han Dyke Poetry*)

From the above examples, it can be concluded that the ancients' cognition of tooth metabolism is the same as our cognition on teeth today; through this metonymic mechanism, the ancients referred to the corresponding age with the physiological feature of tooth loss or rebirth, which not only made it an important part of the rich culture of age appellations in the long-term social development, but also demonstrated the wisdom of the ancients.

3.2. Gender metaphonymy

Gender refers to the difference between male and female. Normally, every species has two gen-

Table 2. A list of metaphonymy of gender attributes in age appellations

Number	Appellation	Age	Example	Source
1	*dou kou* 豆蔻	Girls aged 13 to 14	Young girls of 13 and 14 have a graceful posture and a light and beautiful demeanor, just like the cardamom flower budding in February.	*The Farewell*
2	*ji ji* 及笄	Girls aged 15	A girl should be an adult when she is fifteen.	*Book of Rites*
3	*hua xin* 花信	Women aged 24	When flowers bloom in March, famous flowers trade in the wind.	*Yan Fan Lu*
4	*dai nian/dai zi* 待年 / 待字	The age of 15 when a woman is to be married	The beginning of staying years is the age of going on a journey.	*Quan Tang Wen*
5	*sang mei* 搡梅	A woman who is ready to be married	Plums fall one after another, leaving 70% on the tree. Young man, please don't delay the good time.	*The Book of Songs*
			Plums fall one after another, leaving only 30% on the tree. The young man who wants me must not wait.	
			Plums fall to the ground one after another, and dustpans should be used to clean up. Young man who wants me, please speak now and don't hesitate.	
6	*wu xiang* 舞象	A man aged 15 to 20	Boys aged 15 to 20 can be called as *wu xiang* and are old enough to learn archery and horse riding.	*Book of Ri*tes
7	*ruo guan* 弱冠	A man aged 20	Boys aged 20 can be called adults.	*Book of Rites*·

ders, male and female. For humans, it's the same. In the complicated age appellations in ancient China, the expressions that distinguish male and female are also very rich. They are not only the basic embodiment and cognition of ancient Chinese ancestors, but also the manifestation of their wisdom. There are seven such expressions in the corpus of this paper, as shown in **Table 2**.

Based on social metaphonymy, according to the appellation of male and female age differences, this paper finds there are three obvious trends about male and female gender in ancient China. First, women's status is relatively low. It is generally believed that women at certain age are supposed to marry and stay home while men can work as officials. In example (4), if a woman cannot get married on time, the word "*dai nian*" is used to refer to her age, which highlights the social cognition that women must get married at an appropriate age. Example (5) holds that when a man reaches the age of 20, he must have the crown ceremony to show that he is an adult and can become the master of his own house and pursue fame and fortune.

(4) Xi said: I have a daughter who is not yet fifteen but she is smart enough. She can come to your house and wait to be married, if you like. (*Eastern Zhou Kingdoms*)

(5) By the age of 20, not only was I good at writing, but I also read widely. (*Book of Rites*)

Second, women are weak and men are masculine. For example, "cardamom" in example (6) highlights the beauty of women's softness, while "*shu fa*" in example (7) refers to a boy's hair tied in a bun at the age of 15 before the Qing Dynasty, indicating that he has entered the age of a child

and should be masculine.

(6) Young girls of 13 and 14 have a graceful posture and a light and beautiful demeanor, just like the cardamom flower budding in February. (*The Farewell*)

(7) When boys grow up, they will tie their hair in a bun. (*Book of Rites*)

Third, women are good at dressing, and men are good at martial arts. For example, *ji* (a hairpin) in example (8) originally refers to knotting hair but is worn with a hairpin. It is an iconic ornament of ancient women and is used to refer to the age of 15 years. In example (9), "dancing" is homophonic with fighting in Chinese, which means one can go to the battlefield. Therefore, *wu xiang*, literally "dancing elephant", is used to highlight that men are aggressive and good at martial arts.

(8) I have five daughters, all of whom are smart. First, I taught them the Confucian classics, and then poems and songs. Although they do not reach the period of wearing a hairpin (the age to get married), they all write good articles. (*Old Tang Book*)

(9) Boys who aged 15 to 20 can be called as *dancing elephants* (*wu xiang*) and are old enough to learn archery and horse riding. (*Book of Rites*)

3.3. Wisdom metaphonymy

In the long-term production practice of mankind, personal intelligence will increase with age; so will the knowledge. And people's understanding and perception of life will be gradually deepened and widened. The ancients often referred to age based on their perception, knowledge and wisdom of life, and the cognitive mechanism is a metonymic mechanism based on social embodiment. There are 10 corpora in this paper. See **Table 3** for details.

The history of human's social development is a history of human cognition of themselves. Based on long-term social practice, the ancient people recognized that each age stage should have a specific mission and have a special understanding of life. Human beings used the mechanism of metaphonymy to refer to a specific age with certain social mission and wisdom characteristics of a specific age stage. In example (10), *you xue* is used to refer to the age of 10, because the mission of this age is to acquire knowledge, or go out to learn from teachers, which indicates that people at this age should step into society and continue to learn and acquire knowledge; in example (11), *bu huo* means that at forty, people can be intelligent and broad-based, and can distinguish things clearly. According to Zheng Xuan's explanation, *zhi shi* example (12) refers to the retirement from the government and give the official work back to the king (see Ma, 2007: 57). Later, *zhi shi* was used as a substitute for the age of 70.

(10) A boy reaching the age of ten is called *you*, and it is time for him to go to school. (*Book of Rites*)

(11) I wasn't confused in case of trouble when I was 40. (*The Analects of Confucius*)

(12) By the age of 70, old and frail, it is time for officials to retire. (*Book of Rites*)

3.4. Metaphonymy of ornaments

During the development of human society, people's love of beauty has also been continuously

Table 3. A list of wisdom metaphonymy of age appellations

Number	Appellation	Age	Example	Source
1	*you xue* 幼学	Ten years old	A boy reaching the age of ten is called *you*, and it is time for him to go to school.	*Book of Rites*
2	*wai fu* 外傅	Ten years old	Boys can go for schooling at the age of ten.	*Preschool Education*
3	*hua nian* 华年	Days of youth	Why does the exquisite serpent have fifty strings? Every string and column remind me of my youth.	*The Zither*
4	*er li* 而立	30 years old	When I was 15, I set my mind on learning. When I was 30, I could do things independently.	*The Analects of Confucius*
5	*bu huo* 不惑	40 years old	I wasn't confused in case of trouble when I was 40.	*The Analects of Confucius*
6	*qiang zhuang* 强壮	40 years old	A man in his 40 is called *qiang* (strong), and it is time to become an official	*Book of Rit*es
7	*zhi fei* 知非	50 years old	In my fifty, I know the mistakes I have made in the previous forty-nine years.	*Huai Nan Tz*u
8	*zhi ming/zhi tian ming* 知命 / 知天命	50 years old	At the age of 50, I can know those who can't be controlled by human power are on God's orders.	*The Analects of Confucius*
9	*er shun* 耳顺	60 years old	At the age of 60, I can tolerate different opinions. At the age of 70, I can do whatever I want while not exceed the rules.	*The Analects of Confuci*us
10	*zhi shi/zhi zheng* 致事 / 致政	70 years old	Officials aged 70 are supposed to retire.	*Book of Rites*

Table 4. A list of metaphonymy of ornaments related to age appellations

Number	Appellation	Age	Example	Source
1	*chui tiao* 垂髫	3 to 9 years old	The elderly and children were all at ease and enjoy themselves.	*Peach Blossom Spring*
2	*ji ji* 及笄	Girls aged 15	Girls aged 15 are old enough to wear a *ji*.	*Book of Rites*
3	*shu fa* 束发	Boys aged 15	When boys become adults, they can go for the Great Learning.	*Book of Rites*
4	*zong jiao* 总角	Childhood	How happy we were, our hair in tufts, how fondly we talked and laughed, how solemnly we swore to be true!	*The Analects of Confuci*us
5	*ruo guan* 弱冠	Men of 20 years old	Boys aged 20 can be called *ruo* and it's time to be crowned as adults.	*Book of Rites*
6	*ai fu* 艾服	Men of 50 years old	Men of fifty years old is called *ai* and are supposed to participate in state affairs.	*Book of Rites*
7	*zhang xiang* 杖乡	Men of 60 years old	When one is getting old, s/he will turn to a walking stick for support.	*Reciprocal Poems in Jian'an*

strengthened. According to the existing archaeological findings, human beings have all kinds of ornaments in the early stage; and with the development of society, ornaments that are suitable for all

age stages have been formed. Therefore, people use the mechanism of metaphonymy to refer to age with ornaments. There are seven such expressions in author's corpus, as shown in **Table 4**.

According to the *Book of Rites*, accessories that people wear in specific situations are an important part of our traditional etiquette. Therefore, in the long-term social development, people continue to form some etiquette norms about accessories. For example, in example (13), *ji* refers to the accessories used to fix the knotting of hair. The ancients used *ji ji* (a hairpin) or *ji* to indicate that a woman has reached the age of marriage; in example (14), *ruo guan* refers to a man who is 20 years old, and *ruo* means "small" and "crown" refers to the "Crown Ceremony" which is connected with our traditional custom that ancient men have the crown ceremony to show their adulthood at the age of 20.

(13) Girls aged 15 are old enough to wear a *ji*. (*Book of Rites*)

(14) Boys aged 20 can be called *ruo* and it's their time to be crowned as adults. (*Book of Rites*)

In addition, looking up the existing corpus, there are as many as accessories for males compared to that for females. For example, *ruo guan*, *ai fu* and *zhang xiang* are typical male accessories, which also shows that men had a higher social and political status in the society at that time.

3.5. Raising metaphonymy

Lindblom (2015) believes that people's development is not sound at birth. In order to grow healthily, people must constantly interact with society to experience and understand the external world. In this process, from childhood to adulthood, the raising of parents or others is of the most importance. Therefore, in ancient Chinese, under the cognitive mechanism of metaphonymy, the ancients often used physiological or external characteristics related to raising to refer to age. There are five such expressions in the corpus of this paper. See **Table 5** for details.

According to the physiological rules, people need to be brought up by others from birth to about 3 years old and also in the last days of life, so the word "baby" in Chinese means a newborn child. As shown in example (15), when a baby is born, the mother holds the baby in arms to breast feed, and *ying* means breast, which can be used to replace the Chinese character *ying* (which means babies) (Chen, 1984: 74). When people get old, they often need to be taken care by others, and the elderly especially need more rest and quietness, as shown in example (16).

Table 5. A list of raising metaphonymy in age appellations

Number	Appellation	Age	Example	Source
1	*ying er* 婴儿	Newborn	A newborn baby can talk even if there is no teacher to teach him. That is because he is surrounded by talking people.	*Chuang Tzu*
2	*yi ni* 嬰婗	Newborn	When a baby is newly born, she initiates her babbling.	*Complete Poetry of the Tang*
3	*qiang bao* 襁褓	No more than 1 year old	People from all directions will come to serve you with their babies, leaving their farms behind.	*The Analects of Confucius*
4	*hai ti* 孩提	2 to 3 years old	Kids aged 2 to 3 years old.	*Mencius*
5	*qi yi* 期颐	100	Centenarians can't take care of themselves, so they need to be supported or taken care of by others.	*Book of Rites*

Table 6. A list of behavioral metaphonymy in age appellations

Number	Appellation	Age	Example	Source
1	*wu shao* 舞勺	13–15 years old	Boys began to learn music, recite poetry and dance at the age of 13.	*Book of Rit*es
2	*wu xiang* 舞象	15–20 years old	Boys over 15 years old can be called adults and can learn archery and horse riding.	*Preschool Education*

(15) Babies are in the year of babbling.

(16) Centenarians can't take care of themselves, so they need to be supported or taken care of by others. (*Book of Rites*)

3.6. Behavioral metaphonymy

People's biological and social attributes determine that they need to interact with the external world in production practice, so as to carry out cognitive processing on specific objects. Some specific behaviors constitute the specific behavioral characteristics of human beings at specific ages. Of course, these characteristics may be different due to different cultures. During the long-term development, ancient Chinese people associated specific behaviors with age, and then used cognitive mechanism of metaphonymy to refer to specific age. There are two such expressions in this paper, as shown in **Table 6**.

As shown in example (17), at the age of 13, children would learn dancing with minors, play with their peers and engage in social communication. The ancients used this behavior metonymy to refer to the age of 13; example (18) tells us that when children are 15 years old, they will dance with weapons to show that they can go to the battlefield. Therefore, *wu xiang* refers to the age of 15 to 20.

(17) Boys began to learn music, recite poetry and dance at the age of 13. (*Book of Rites*)

(18) Boys over 15 years old are old enough to learn dance with weapons. (*Preschool Education*)

3.7. Metaphonymy of animal and plant

Life on earth has experienced and is still undergoing a long process of evolution. Human evolution and survival are inseparable from the external environment, and the animals and plants around us are most closely related to human beings. Therefore, human beings use cognitive methods such as metaphor to project the physiological characteristics of human beings at a specific age and the characteristics of some animals and plants to form relevant metaphorical expressions which are then used to refer to a specific age of humans. There are seven such expressions in the corpus of this paper, as shown in **Table 7**.

In the cognition of the ancients, *huang kou* originally refers to the beak of young birds. Through the metaphorical cognitive mechanism, the iconicity between the beak of the young birds and the newborn baby is established, and the metonymy mechanism is used here to refer to babies. As shown in example (19), *huang kou* refers to the newborns. In addition, as shown in example (20), the ancients believed that in old age, the wrinkled skin of the old was like the markings on the mackerel which indicates that the old men's strength faded and the hue on their skin diminished, so

Table 7. A list of metaphonymy of animal and plant in age appellations

Number	Appellation	Age	Example	Source
1	*huang kou* 黄口	More than10 years old	In ancient times, the young will be left alive in wars between countries.	*Huai Nan Tzu*
2	*dou kou* 豆蔻	Girls aged 13 to 14	Young girls of 13 and 14 have a graceful posture and a light and beautiful demeanor, just like the cardamom flower budding in February.	*The Farewell*
3	*tong* 童	15 years old	A bare mountain raising no vegetation is called *tong*.	*Shimi*ng
4	*sang mei* 搡梅	A woman who is ready to be married	Plums fall one after another, leaving 70% on the tree. Young man, please don't delay the good time.	*The Analects of Confucius*
			Plums fall one after another, leaving only 30% on the tree. The young man who wants me must not wait until now.	
			Plums fall to the ground one after another, and dust-pans should be used to clean up. Young man who wants me, please speak quickly and don't hesitate.	
5	*tai bei* 鲐背	90 years old	The spots on the back of the old man are like mackerel patterns, which is a sign of longevity.	*Erh*ya
6	*huang gou* 黄耇	90 years old	People aged 90 are called *tai bei, huang gou*, *dong li*, *ni chi* or *mei shou*.	*Shiming*
7	*dong li* 冻梨	90 years old	People aged 90 are called *tai bei*, *huang gou*, *dong li*, *ni chi* or *mei shou*.	*Shiming*

they got a thin body and wrinkled skin. In addition, senile spots on the skin of the elderly shaped like the surface of frozen pears, so *dong li* (frozen pear) is used to refer to the elderly (Chen, 1984: 73), as shown in example (21).

(19) In ancient times, the young will be left alive in wars between countries. (*Huai Nan Tzu*)

(20) Ninety-year-old people are having fun hitting boards which are called *rang*. (*Weide Theory*)

(21) People aged 90 are called *tai bei*, *huang gou*, *dong li*, *ni chi* or *mei shou*. (*Shiming*)

3.8. Numeric metaphonymy

People's definition of age is inseparable from human's definition of time based on astronomical knowledge. Taking the earth's revolution around the sun as a year, the concept of human age was born. Of course, Chinese characters themselves condensed the Chinese nation's cognition of the world, so the ancients also tried to express age with the wisdom of Chinese characters which both demonstrated politeness in a specific context and highlighted the speaker's cultural level. Therefore, the ancients used some Chinese numbers to form metonymy which can be used to refer to age. There are seven items in the corpus of this paper, as shown in **Table 8**.

Ancient Chinese literati had the habit of dismantling characters. For example, literati disassembled the character *gua* for two and eight to mark the year. Therefore, as shown in example (22), the metonymy of *po gua* refers to a girl of 16 years old. In addition, in ancient times, the heavenly stems and earthly branches cooperated with each other as the chronological period. Every sixty years is called a *hua jia*, also known as a *jia zi*. Here, *hua* is used to describe that Chinese era was intricate

Table 8. A list of numeric metaphonymy in age appellation

Number	Appellation	Age	Example	Source
1	*chu du* 初度	1-year-old child	The birthday of a 1-year-old child is called "the first birthday".	*Preschool Education*
2	*po gua* 破瓜	Girls aged 16	Maybe at that time, the talented showgirls under 16 wore hair accessories; today when actors played as Dan (a young female role in Chinese Opera), they also wear hair accessories.	*Sheng'an Poetry*
3	*ban bai* 半百	50 years old	Xun Ouyang, one of the best calligrapher of Wang Xizhi's semi-cursive script, started learning calligraphy over 50 years old; even his works can be seen as perfect as god's works, it was not until he was 53 years old that he was acclaimed for his success in writing it.	*Miscellaneous Words of Mongolian History*
4	*da yan* 大衍	50 years old	*Da yan*, fifty years.	*Yi*
5	*hua jia* 花甲	60 years old	Since I had lived for more than 60 years, I will not regret it if I die now. It's just Lanying, my poor daughter, will suffer all the consequences.	*Ming Zhu Yuan*
6	*bai shou* 白寿	99 years old	Thousands of *bai shou* brought the post to the Criminal Department. Liu Ji asked for the original memorial, but they refused to give it, and Liu Ji also refused to issue it.	*Ming History*
7	*cha shou* 茶寿	108 years old	When one is in the period of *cha shou*, he has a really long longevity.	*Documents of History of Qing Dynasty*

and uneven. Thus, as shown in example (23), the metonymy of *Hua Jia* refers to the age of sixty. This is also the wisdom and embodiment of Chinese traditional culture (Lu, 2018).

(22) Haven't reached the age of 16, the girl had been enrolled into the mansion of marquess as a good singer and dancer. (Lu You, *Untitled*)

(23) In the year of 60, he stroked his white-colored beard. The world still goes round and round, and it circulates day after day, just like beads fiddled back and forth. (*Record Event of Tang Poetry*)

4. Conclusion

Based on the metaphonymical mechanism of social embodiment, this paper systematically analyzes the complex age appellation system in Chinese. From the cognitive ways of the eight types of age appellations described in this paper, the reason why these age appellations can become age references is that in human experience and cognitive activities, people established cognitive relevance based on specific context and cognitive perspective. Highlighting specific features and using metonymic mechanisms to refer to age have formed the wealth of age appellation expressions in ancient China, which strongly proves the view that sociality is the essential attribute of metaphor and metonymy. We can certainly see, from the source of the corpus, that the birth, development and dissemination of these age titles are also due to the literati, poets and their works, which causes these intelligent and rich expressions to appear, develop and continue. Therefore, literary works have the function of social construction. That is to say, some expressions in ancient times are still used by people today, but some are not. For example, the expressions in *the Analects* are more commonly

used in daily language expressions and widely spread in Chinese language and culture, while the expressions in some minority works have not been widely spread in Chinese traditional culture and are no longer commonly used. However, the wisdom reflected by them is still an important part of Chinese traditional culture.

Conflict of interest

The author declared no conflict of interest.

Acknowledgement

The research is sponsored by the Chinese National Fund Project for Philosophy and Social Sciences "A Study on Sociality of Language Based on Cognitive Sociolinguistics" (18XYY001).

References

An Y (1987) An etymological studies on ancient age appellations. *Journal of Pictorial* 3: 74–76.

Chen F (1984) A research on ancient age appellations. *Literature and History* 10: 73–74.

Cheng X (2009) Construction of subjectivity and intersubjectivity from the perspective of the philosophy of language. *Foreign Language Research* 1: 9–15.

Lindblom J (2015) *Embodied Social Cognition*. New York/London: Springer.

Lakoff G and Johnson M (1980) *Metaphors We Live by*. Chicago: Chicago University Press.

Lakoff G and Johnson M (1999) *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lantolf J and Qin L (2018) Social-cultural theory: Philosophical foundations, disciplinary property and research methodology. *Foreign Language and Their Teaching* 1: 1–18.

Lu S (2018) *The History of Chinese Culture*. Beijing: Beijing Daily Press.

Ma L (2007) An etymological studies on age appellations in Chinese. *Journal of Zhejiang Education College* 5: 53–58.

Panther K and Radden G (2002) *Metonymy in Language and Thought*. Amsterdam/Philadelphia: Benjamins.

Shapiro L (2019) *Embodied Cognition*. London/New York: Routledge Taylor and Francis Group.

Shen J (2009) Subjectivity of Chinese and the teaching of Chinese grammar. *Chinese Language Learning* 2: 3–12.

Wang Y (2014) Embodied-cognitive linguistics in the viewpoint of postmodernism. *Foreign Language and Literature* 6: 61–67.

Wang Y (2015) The theoretical exploration and practical value of embodied-cognitive monism: A new thought on philosophy of mind. *Foreign Language of China* 2: 24–31.

Wen X (2019) Sociocognitive linguistics based on social cognition. *Modern Foreign Languages* 3: 293–304.

Winter B (2019) *Sensory Linguistics: Language, Perception and Metaphor*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Zhao Y (2015) A study on cognitive sociolinguistics under the post-modernistic view. *Foreign Language Research* 4: 65–70.

Zlatev J (2007) Embodiment, language, and mimesis. In: Geeraerts D, Dirven R and Taylor JR. (eds.) *Body, Language and Mind (Volume 1): Embodiment*. Berlin/New York: Mouton de Gruyter, pp.297–337.

ORIGINAL ARTICLE

The developmental trajectory of mandarin Chinese-speaking children's pure metonymy comprehension ability

Songqiao Xie*

University of Cambridge, Cambridge, UK

Abstract: This empirical study investigates Chinese children's developmental trajectory of pure metonymy comprehension. In the light of the experiment design in Jiang's (2019) and Köder and Falkum's (2020) studies, the present study, adopting a quantitative approach, employed a modified behavioral experiment and an eye-tracking experiment. Drawing on the experimental data, the study finds that: a) children's metonymy comprehension performance showed a tendency towards the U-shape in the behavioral experiment tasks; b) children's target (metonymy) fixation proportion, however, developed with age in the eye-tracking tasks; c) children's metonymy comprehension not only developed with age but also showed different features in different difficulty levels of metonymies. Thus, this study explains the U-shape by arguing that age-4 and -5 children's pure metonymy comprehension ability can be masked not only by a literal preference reported in Köder and Falkum's (2020) study but also by the high randomness of task results of the age-3 participants and the high level of difficulty of culture-related metonymies. Moreover, the study also argues that year six is a crucial stage for children's metonymy comprehension development, which provides implications for children's early figurative language education.

Keywords: children's metonymy; pure comprehension; developmental trajectory; eye-tracking experiment; behavioral experiment

***Corresponding author:** Songqiao Xie, University of Cambridge, Cambridge; postal code: CB2 8PH, UK; sx264@cam.ac.uk

Received: April 22, 2022; **Accepted:** June 20, 2022; **Available online:** August 19, 2022

Citation: Xie S (2022) The developmental trajectory of mandarin Chinese-speaking children's pure metonymy comprehension ability. *Forum for Linguistic Studies* 4(1): 78–99. DOI: 10.18063/fls.v4i1.1457

1. Introduction

In recent years, the growing interest in figurative language use or use of expressions that require ad hoc concept constructions (Carston, 2002), doubled with the increasing interest in cognitive pragmatics, has given rise to a rich body of studies that informed the fields of human figurative language processing and production.

However, the existing figurative language studies have persistently focused on a narrow range of figurative language devices, among which the most frequently studied is metaphor (Rundblad and

Annaz, 2010), resulting in an inadequate concern for metonymy, a figurative device which expresses a real-world contiguity by referring to a target via a salient property. Metonymy is also claimed to rely on a cognitive pragmatic mechanism (Falkum, 2019) different from that of metaphor. While some studies were conducted on the cognitive mechanism of metonymy, few have approached this issue from a developmental perspective.

Even within the scarcity of literature on the trajectories of children's figurative language ability development, an imbalance of research efforts can be observed between children's metaphor development and their metonymy development. Existing studies (Özcaliskan, 2005; Peng and Zhang, 2009; Liu and Mi, 2008; Rundblad and Annaz, 2010) have found and confirmed children's early (at the age of four) ability, which improves chronologically with age, of understanding metaphorical uses of language, and have also realized that different categories of metaphors may be acquired at different stages (Zhou, 2003). However, when it comes to metonymy, children's metonymy development remained an uncharted territory until the recent decade.

Rundblad and Annaz (2010), Van Herwegen et al. (2013), Falkum et al. (2017) and Jiang (2019) have made valuable research efforts which paint a polychrome picture of this field. It is polychrome because, on the one hand, their studies serve as an intriguing guidance to a wide range of topics regarding children's metonymy, from metonymy comprehension to metonymy production (Falkum et al., 2017; Jiang, 2019), and from chronological development (Van Herwegen et al., 2013) to U-shaped development (Falkum et al., 2017); on the other hand, extant research also leaves contradictions (Van Herwegen et al., 2013; Falkum et al., 2017) and questions open to discussion. For example, one pending issue is how to, at the cognitive level, refine the existing experiments to further distinguish between children's metonymy comprehension and metonymy production. How to provide a more convincing and solid explanation for the U-shape observed in the previous research is another problem unsolved. Furthermore, whether the acquisition of different categories of metonymy corresponds with different age stages is also an interesting topic which few studies have explored.

In the light of the questions identified above, the present study aims to investigate Chinese children's pure metonymy comprehension development, employing eye-tracking and improved picture-selection activities as experiment instruments.

Köder and Falkum (2020) and Jiang (2019) managed to differentiate between children's metonymy production and comprehension, and Köder and Falkum (2020) concluded that "eye-tracking is a 'purer' and cognitively less demanding measure of figurative language comprehension." So, the present study employs an eye-tracking experiment, and also, with an intention to improve Jiang's (2019) experiment method (picture-selection, which touches upon the "production" level), the author of the present study will conduct the picture-selection activity followed by a retrospective interview with the children subjects, to minimize the influence of speaker intention in the instant context.

Since the subjects in the present study are Chinese children, this study will also serve as an attempt to test the translingual applicability of the U-shape observed in the previous studies, which were conducted in English and Norwegian contexts. Not confined to this, considering that the discussion (Köder and Falkum, 2020) of the U-shaped development of children's metonymy was backed up with solid evidence, this study attempts to provide a new perspective for the explanation of the U-shape, with the combined evidence from the eye-tracking and modified behavioral experi-

ment data.

Also, given that some metonymies are lexicalized (Hilpert, 2006), especially the culturally related ones, while some are ad-hoc metonymies whose understanding requires instant construction of relations, the present study will also probe into, though not so deeply, the question of the acquisition order of different categories of metonymies. To achieve this, my experiment will include a comparison between three levels of metonymies assigned with different difficulties.

2. Research on metonymy

In our daily communication, figurative language plays a ubiquitous part, making it possible for speakers and listeners to go beyond conventional lexical meanings and construct novel context-dependent meanings. One case of this is metonymy, a trope (Corbett and Connors, 1999) or a figurative device that expresses a real-world or mental contiguity by referring to a target via a salient property (Koder and Falkum, 2020). Refer to the examples in (i), (ii) and (iii):

(i). The piano is in a bad mood. (Panther and Thornburg, 2005)

(ii). The sax has the flu today and he will not be able to play tonight. (Zhang and Lu, 2010)

(iii). The moustache sits down first. (Falkum et al., 2017)

In (i) and (ii), the pianist and the saxophonist are referred to via "piano" and "sax" respectively, the names of the instrument they play. In (iii), the "moustache" refers to a person wearing a noticeable moustache, which possibly differentiates him or her from others in the context. The piano-pianist, sax-saxophonist, and moustache-moustache wearer are in a contiguity relation, either conceptually or physically.

2.1. Metonymy from cognitive perspective

In figurative language literature, in spite of the burgeoning research interest in the cognitive process of figurative language, extant studies have primarily centered on a narrow range of figurative devices, mainly metaphor (Rundblad and Annaz, 2010), unfortunately resulting in an inadequate attention towards metonymy.

It was only after the end of the 1990s that metonymy started to become a research focus among cognitive linguists (Langacker, 1993; Radden and Kövecses, 1999; Panther and Radden, 1999; Barcelona, 2000; Alač and Coulson, 2004), whose works mainly focus on the mental conceptualization strategies and cognitive mechanisms of metonymy and interfaces and demarcations between metaphor and metonymy. Langacker (1993) categorized metonymy as a "reference point phenomenon", based on which Alač and Coulson (2004) described the cognitive process of metonymy as a process where an item with relatively higher salience serves as the reference point, rendering other less salient items mentally accessible. Building on Lakoff and Johnson (1980), who proposed the idea of ICM (Idealized Cognitive Model) and argued that the human conceptual system is figurative in nature, Radden and Kövecses (1999) reemphasized cognitive nature of metonymy and defined metonymy as a cognitive process that functions within one ICM. All of these research efforts have deepened the understanding of the nature of metonymy and established a research trend and guidance for further cognitive linguistic studies on metonymy.

2.2. Research on children's metonymy

Since associative relations, which render the comprehension and production of figurative language, are frequently exploited in our everyday communication, it is important to study when people acquire this ability and how this ability develops from childhood to adulthood. This problem has been studied since the 1960s, when researchers (Asch and Nerlove, 1960; Gardner, 1974; Mendelsohn et al., 1984) tended to address this issue from a semantic perspective. However, in the literature on figurative language cognition, it was not until the 2000s that children's uses of figurative language started interested researchers (Zhou, 2003; Özcaliskan, 2005; Peng and Zhang, 2009; Liu and Mi, 2008; Rundblad and Annaz, 2010) working on language acquisition and cognitive development. Liu and Mi's (2008) research found that children of four are equipped with the ability to understand metaphorical utterances in neutral context and are also able to make abstract reasoning about the metaphorical mappings; Rundblad and Annaz's (2010) result confirmed that the ability to comprehend metaphor resides in children early on and improves steadily throughout childhood to adulthood. There is growing consensus that children have an early competence of and affinity with figurative language.

As can be concluded from the existing studies on children's figurative language comprehension, much has been said about when and how children begin to cope with the uses of metaphor, while, to the best of my knowledge, less is known about children's ability of metonymy. Only a handful of research has investigated children's development in comprehension strategy and production purpose of metonymic expressions (Rundblad and Annaz, 2010; Van Herwegen et al., 2013; Falkum et al., 2017; Köder and Falkum, 2020). Though few, these research efforts have opened up a new area for the study of children's metonymy comprehension. Rundblad and Annaz (2010) and Van Herwegen et al. (2013) investigated the trajectory of development of metonymy comprehension from year 5 and 6, and discussed the result comparatively with the result of metaphor, and finally found in children an early metonymy comprehension competence which could chronologically improve with age. Falkum et al. (2017) conducted a research studying children's metonymy comprehension and production, and found a U-shape in metonymy comprehension competence variation, with 3-year-olds doing surprisingly better than 4- and 5-year-old children; Jiang (2019) found the similar result

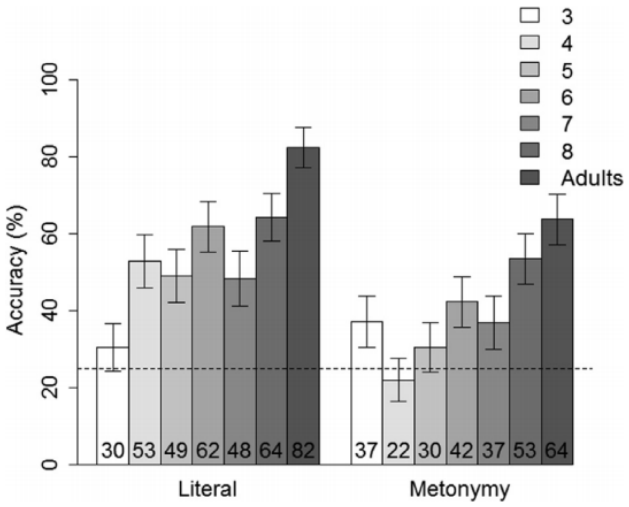


Figure 1. Percentage of correct picture choice in the literal and metonymy conditions (Köder and Falkum, 2020).

in the referential game she designed for the children tested. Köder and Falkum's (2020) conducted an empirical study on young children's (aged 3–8) comprehension of novel metonymy and they obtained both a similar and a different result from the two sections of the experiment they conducted, thus claiming that eye-tracking is "a purer and cognitively less demanding measure of figurative language comprehension" (Köder and Falkum, 2020).

2.3. Questions for further investigation

While providing valuable insights into children's abilities and affinity with metonymy, the extant studies are not flawless, thus leaving several contradictions and questions open to further investigation.

One thing problematic is a lack of effort to further distinguish between metonymy comprehension and production, especially in the experiment methods of previous research. Children's ability of metonymy consists of metonymy comprehension and metonymy production, as most of the studies on children's metonymy has differentiated between the two levels, focusing on either production on a pragmatic level or comprehension on a cognitive level. The researchers (Falkum et al., 2017; Jiang, 2019) have also designed and conducted their experiments accordingly. In Falkum et al.'s (2017) study and Jiang's (2019), researchers employed picture selection as their experiment instrument to examine and collect data concerning children's metonymy understanding, while using other methods to examine metonymy production. This seems to have made clear the boundary between the cognitive level and pragmatic level; however, the method itself is still problematic, as picture-selection inevitably involves interactions between the examiners and the children participants, whose perception of the examiners' speaker's intention would, to a large extent, affect their choices of the pictures. In other words, picture-selection as a method, though aiming to examine metonymy comprehension, still touches upon the production level, thus being a less pure method for metonymy cognition. Consequently, the present study, building on a reflection of the existing studies (Van Herwegen et al., 2013; Falkum et al., 2017; Jiang, 2019; Köder and Falkum, 2020), plans to employ picture-selection + interview and eye-tracking experiments, which was claimed to be a purer and less demanding measure of figurative language comprehension (Köder and Falkum, 2020), to collect data of children's pure metonymy comprehension.

Another pending issue is about the discussion of the U-shape developmental curve, which has been observed in several cognitive domains (Yan and Huang, 2009), not confined to the metonymy comprehension. One contradiction is that while Falkum et al. (2017) found a U-shape in her study, Van Herwegen et al.'s (2013) result found a linear steady improvement of children's metonymy comprehension. It might be due to a relatively small number of subjects in Van Herwegen et al.'s (2013) study, which might have concealed the U-shape. However, in Köder and Falkum's (2020) study, the researcher provided alternative explanations, claiming that the seemingly decreased competence of age-4 and age-5 children could be masked by a literal preference at that age, which means around the age of 4 and 5, children undergo a period when they show a general preference for literal interpretations and a rising awareness of speaker's intentions. In other words, it is possible that age-4 and -5 children are equipped with a better comprehension competence of metonymy; however, this competence is hidden under the cloak, and thus not manifested in the previous studies. Considering that U-shape is still a puzzling aspect of children's metonymy and that the previous study (Köder and Falkum, 2020) did not provide a thorough discussion over it, the present study

aims to provide some possible insights into this topic.

There is also a question that resides within "metonymy" itself, concerning the different categories of metonymic expressions. Few of extant studies on children's metonymy comprehension paid attention to the possible difference of comprehension order resulted by different kinds of metonymies. In Zhou's (2003) study on children's developmental trend of metaphorical ability of time, the researcher realized that children's comprehension ability of different categories of metaphorical expressions may be different with their age. For example, children's comprehension of personification of time and space-time metaphors appears earlier than other categories of metaphors (Zhou, 2003). Accordingly, we tempt to ask a similar question about metonymy: are different categories of metonymies acquired by children in different time order? Which categories of metonymy are acquired earlier? Given that the time is limited, the present study does not address this issue in detail; however, the present study, in order to avoid stereotyping metonymy as "all metonymies", sets the operational concept of metonymy by dividing metonymic expressions into different difficulty levels. In the behavioral experiment, the stimuli contain three difficulty levels of metonymies: the lowest difficulty-level of metonymy in this study is concerned with the relations of human-body-related physical contiguity, since children are early familiar with animated and human-like properties (Piaget, 1964), for example, the red cap and the cap wearer; the medium level is concerned with the relations of non-body-related contiguity; the highest level is added with cultural-related elements, for example, "lose your bowl" and "lose your job" (in the Chinese context).

As for the stimuli in the eye-tracking experiment, the present study modifies Peirsman and Geeraerts' metonymy classification model and constructs a three-dimensional model where metonymies, especially for children, can be classified into different types and difficulty levels.

Peirsman and Geeraerts (2006) provided a classification model for metonymies, where a large number of typical metonymies are differentiated and categorized in terms of the type of contiguity they are motivated by. Their model is three-dimensional (**Figure 2**).

As positive shifts are made on each dimension line (domain, boundedness or strength of contact), the difficulty of processing and understanding the certain metonymies increases. For example, along

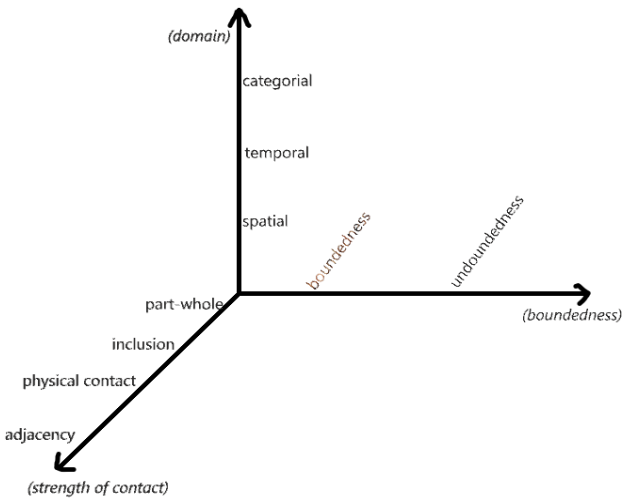


Figure 2. Metonymical patterns in Peirsman and Geeraerts (2006) (adapted).

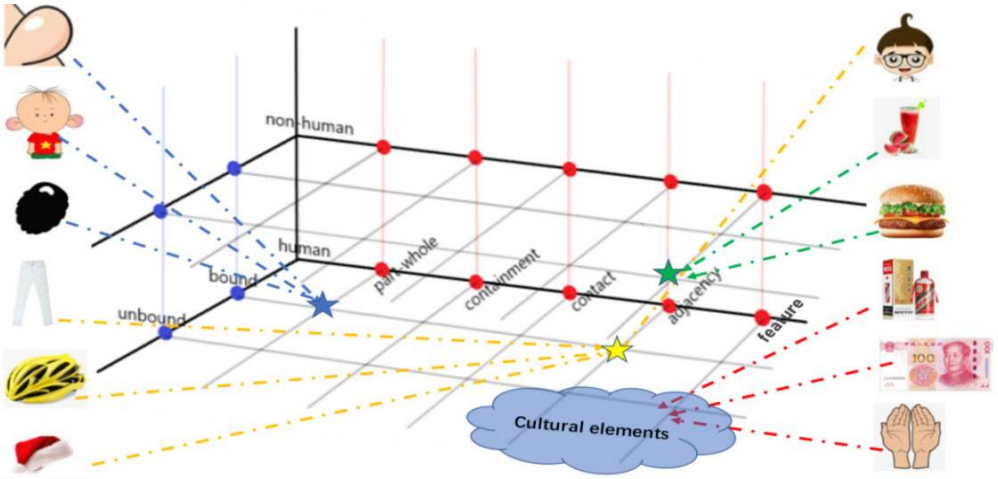


Figure 3. A revised model of metonymy taxonomy.

the line of strength of contact, the contact between the beard and the beard wearer (a part-whole contiguity) is stronger than the contact between the clothes and the clothes wearer (an adjacency contiguity), and thus the difficulty of using beard to refer to a man who wears a beard is lower than the difficulty of using a pair of shoes to refer to the person in them.

The present study, on the basis of Peirsman and Geeraerts' model, has made realistic adjustments in setting the difficulty of the stimuli in the experiment. Since children are early familiarized with animated and human-like properties (Piaget, 1964), human-body-related metonymies are considered less difficult than non-body-related metonymies, which is, in the present study, the primary standard for metonymy difficulty classification. Then, under this primary standard, the classification is subject to the above-mentioned model.

As a result, in **Figure 3**, the three dimensions are changed, compared with Peirsman and Geeraerts' model, into "human or not", "boundedness" and "strength of contact". Human related metonymies are attributed to the first layer, namely the "human platform", and are regarded less difficult to comprehend for children than the non-human-related metonymies, which belong to the second layer. For example, the clothes-wearer metonymy is easier than the bottle-water metonymy, although it reverses the rank between containment contiguity and adjacent contiguity.

The U-shape developmental curve in children's metonymy comprehension lacks test in the Chinese speaking context. Falkum et al.'s (2017) experiment exploring children's acquisition of metonymy was conducted in English, while the experiment of Köder and Falkum's (2020) another study was conducted in Norwegian, which marks an inter-language attempt. However, in view of the fact that English has, in history, received heavy influence from Scandinavian languages (Lutz, 2017), whose descendants include Norwegian, I argue that this attempt or extension is not far enough. To further test the universality of the conclusion made by the study, the present study is conducted to investigate the metonymy comprehension ability of children whose mother tongue is Chinese, a language that does not belong to the Indo-European language family.

Aware of the problems identified above, the author of the present study intends to address the following research questions:

- Can the U-shape be replicated in the Chinese speaking context?
- How can the explanation of the U-shape be further improved concerning the metonymy competence of children of different age groups?
- Can higher difficulty levels of metonymies mask children's pure metonymy comprehension?

3. Methodology

The present study employs a quantitative approach, conducting a three-section behavioral experiment (picture-selection plus retrospective interview + story Q&A) and an eye-tracking laboratory experiment. The behavioral experiment in this study partly follows and modifies the behavioral experiment designs in previous studies (Jiang, 2019; Falkum, 2020) by adding retrospective interview sections and adding stimuli of different difficulty levels. Eye-tracking is increasingly applied as an experiment instrument in the fields of human cognitive studies (Dong, 2014) and the method itself is supported by the Eye-brain Unity Hypothesis, which states that there is no appreciable lag between what is fixed and what is processed (Just and Carpenter, 1980). Thus, the present study employs an eye-tracking experiment to explore children's inner comprehension process of metonymy expressions.

3.1. Behavioral experiment

In the behavioral experiment, which is divided into three sections (each section contains 5 questions), we employed two experiment methods, including picture-selection plus retrospective interview (for section 1) and story Q&A (for sections 2 and 3). The three sections of the experiment are assigned with different levels of difficulty, aiming to investigate the children's metonymy comprehension development in different types or difficulty-levels of metonymies. All the participants are asked to complete all the three sections. In this experiment, the experimenters collect the data of error rate in all the three sections, the data of reaction time in sections 1 and 2.

3.1.1. Participants

We invited 32 children, all of whom are screened to exclude health and intellectual impairments and speak Chinese as their mother tongue, to participate in our behavioral experiment. The age of the children ranges from 3-year to 8-year, and all the participants were distributed into three age groups: group 1 (3 years old), group 2 (4–5 years old) and group 3 (over 6 years old). The context of this experiment is an international kindergarten located in Gaoxin District, Zhengzhou. Written parental consent was obtained prior to the experiment.

Prior to the experiment, we conducted a pilot study with two 3-year-old children to make sure that children older than 3-year can understand the words uttered by the experimenters and can recognize the different components that appear in the pictures.

3.1.2. Design

This was a 3 (age) × 2 (gender) × 3 (task) in-subject design, with age and gender as between-groups factors and difficulty level as an in-group factor.

The three independent variables are: (1) age (year 3, year 4–5, year 6–8); (2) gender (female vs.

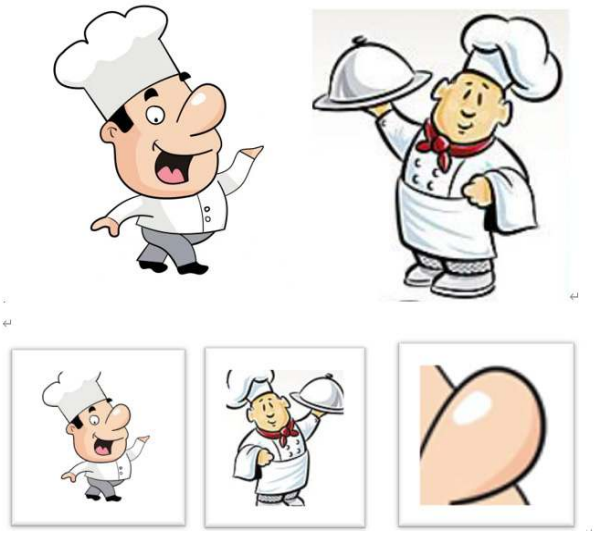


Figure 4. Example stimulus.

male); (3) task (picture selection + retrospective interview of human body-related metonymy, story Q&A of non-body related metonymy, story Q&A of culture-related metonymy).

There were also three dependent variables: error rate, experiment score, reaction time.

3.1.3. Stimuli and procedure

In the first section, aiming to test children's metonymy comprehension competence concerning the relations of physical human-body-related contiguity, the experimenters would show the participants 5 sets of pictures. Each set consists of 4 pictures: one background picture and three option pictures. The background picture gives participants a general idea of "what" is going on with "whom". The three option pictures contain one literal picture (the object, for example the "big nose" itself), one metonymy picture (the right answer) and the irrelevant picture (the other person from the background picture).

The participants are tested one by one. At each time, one participant is shown one set of the pictures while being told a story which narrates the scene and gives the instruction to choose one picture. Each story is composed with one context utterance ("These two cooks work in one restaurant…"), which helps the children to establish a focus on both human-components in the background picture, and one target utterance, which helps the participants to differentiate between the literal and metonymic pictures. Each target utterance contains a referential expression of the metonymic target ("the big nose…") and a describing expression ("cooks the better") which helps with the instruction.

One example of the story is shown below. (Since the whole experiment is conducted in Chinese, the lines below are the translated version of the story told in the experiment.)

"These two cooks work in one restaurant. However, most of the customers of the restaurant comment that the big nose cooks better meal. So, please tell me who cooks the better meal?"

After the instruction is given, the reaction timing starts, and stops once the participant gives his or her choice. Then, after the participant's choice is given, one follow-up question ("Could you tell

me why you choose this?") is asked to collect data concerning the retrospective comprehension process of the participant.

In the second section, the experimenter tells the participant short stories (5 in total) concerning non-body-related metonymies. Each story contains the background information of the target and the metonymic expression. Following the story, two choices are given to the participant who is instructed to pick one according to the context of the story. The reaction of each participant is timed.

One example is shown below.

"Lu Xun is a famous thinker and writer in China. As a writer, he has many famous works, however, many people think that Lu Xun is really hard to read. So, please tell me what is hard to read, Lu Xun the person himself or the works he wrote?"

In the third section, the experimenter tells the participant short stories (5 in total) concerning metonymies with cultural-related elements. Each story contains the background information of the target and the metonymic expression; however, this time no choices would be given to the participant—the experimenter would ask directly for an explanation of the meaning of the metonymic expression. In this section, the experimenter does not time the reaction of each participant, considering that the difficulty of the metonymies in this section is relatively higher than the previous, and that, therefore, "no response" from the participant may occur frequently.

One example is shown below.

"Uncle Wang has messed things up in his work. He does not dare tell this to his leader because he is afraid of 'losing his bowl' (understandable within the Chinese cultural background). So, please tell me what 'losing his bowl' means here?"

3.2. Lab experiment

This eye-tracking experiment was conducted in the cognitive laboratory of Shanghai International Studies University. The participants' eye movements were tracked by Eyelink-1000 plus, an eye-tracking device with a sampling rate of 1,000 hz. The whole experiment procedure was designed and programmed on Experiment Builder, an experiment design software intended for Eyelink devices. Data collected in this experiment were used to analyze children's eye movements during the process of hearing the audio story while looking at the picture stimuli on the display screen. After collection, the data were generated on Data Viewer, and analyzed on Python.

3.2.1. Participants

Six children ranged from 3 to 8 years old were invited and tested in this experiment. **Table 1** lists the mean age, number and gender distribution in the different age groups. All participants were native Chinese speakers screened to exclude health, vision and intellectual impairments. Children were recruited from kindergartens and schools in Shanghai. Written parental consent was obtained prior to the experiment. Each participant was tested individually in the eye-tracking lab.

3.2.2. Design

This was a 3 (age) × 2 (gender) × 2 (difficulty level) in-subject design, with age and gender as between-groups factors and difficulty level as an in-group factor.

Table 1. Participants

Age group	Mean age	Number	Gender (f/m)
Year 3	3.46	2	2/0
Year 4–5	5.04	2	1/1
Year 6–8	7.34	2	2/0

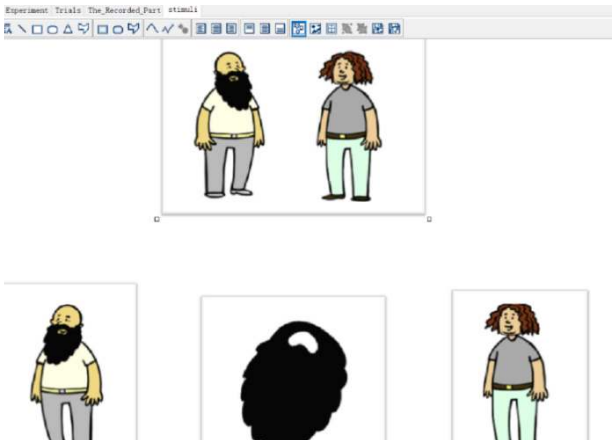


Figure 5. Example stimulus.

The three independent variables are: (1) age (year 3, year 4–5, year 6-8); (2) gender (female vs. male); (3) difficulty level (lower difficulty levels vs. higher difficulty levels).

The eye movement measure is children's fixation duration on the target areas.

3.2.3. Stimuli and procedure

The stimuli contain 12 sets of pictures, the first two sets being training stimuli. The rest 10 sets of picture stimuli are divided into four difficulty levels (two of lower levels and two of higher levels), following the classification provided in §2.4 (Figure 3). Every participant was asked to complete all the 12 trials.

Each set of pictures contains four pictures placed on four areas of the screen. The largest picture on the upper half of the screen is the background picture, which helps the audio listener to form the context of the story; the three pictures on the lower half are option pictures: one literal picture (the object, for example the "big beard" itself), one metonymy picture (the right answer) and the irrelevant picture (the other person from the background picture). The participants are asked to choose one, by pointing with finger, according to the story, which lasts around 20,000 ms, they hear.

Each audio stimulus, similar to the stories in the behavioral experiment, consists of one context utterance and one target utterance. Each target utterance contains a referential expression of the metonymic target and a describing expression, which helps to instruct the participants to make the choice.

The set of pictures are shown to the participant 4,000 ms earlier than the audio is displayed, which allows the participants enough time to recognize and be familiar with the elements in the pictures.

Eye-movements of participants are tracked and recorded during the whole process of each task. Viewing was binocular, but only the movement of the left eye was tracked and recorded[1].

Participants were positioned for individual comfort in an adjustable chair with a chin rest to minimize head movements for an increase in measurement accuracy.

4. Findings and discussion

4.1. Behavioral experiment results

To examine the age differences in metonymy comprehension and the influence of the task differences on performance, we used Excel to analyze the data collected from the behavioral experiment, treating "age" and "task" as independent variables, while "error rate", "experiment score" and "reaction time" as dependent variables.

Concerning the "error rate", all the wrong responses in picture-selection, including choosing the literal pictures and the irrelevant pictures, were all coded as "error". Concerning the "experiment score", all the right answers were coded as "1"; all the literal choices in picture-selection, all the irrelevant responses in retrospective explanation and all the wrong answers in other tasks were coded as "0"; all irrelevant choices in picture-selection were coded as "X" (not shown in the tables); all the half-relevant responses in retrospective explanation were coded as "0.5".

Concerning the "reaction time", considering the relatively high difficulty of task three and the ages of the children, we did not time the reaction of the participants in the third task, the story Q&A on metonymies with cultural elements.

4.1.1. Error rate

Table 2 presents the error rate data of the children's performance in the 3 experimental tasks. One-way analysis of variance (ANOVA) between groups is carried out on children of different ages in different experimental tasks. In the picture-selection task (the retrospective explanation task is not included), we find no difference between the performance of the age-3 group and the age-4-to-5 group ($F = 0.306$, $P = 0.587$); in comparison to the age-4-to-5 group, age-6-to-8 group demonstrate a lower error rate, however, this difference is not statistically significant ($F = 2.673$, $P = 0.119$). In the story Q&A concerning non-body related metonymies, no difference is found between the performance of the groups of lower ages (age-3 and age-4-to-5) ($F = 0$, $P = 1$); however, the error rate of the age-6-to-8 group shows a significant drop ($M = 0.12$, $F = 4.314$, $P = 0.052$). In the third task (story Q&A concerning metonymies with cultural elements), significant difference can only be detected between age-3 group and age-4-to-5 group ($F = 5.018$, $P = 0.038$).

4.1.2. Experiment score

We analyzed the data of experiment scores by conducting one-way ANOVA between groups. **Table 3** mainly shows how the mean scores of picture-explanations and the mean total scores vary between the three groups.

A significant trend can be found in picture explanation that children's ability to explain their

[1] Kliegl et al. (2006) has shown high correlations between right-eye and left-eye fixation durations ($r = 0.98$).

Table 2. Error rate of 3 age groups in 3 tasks

Age group	Picture selection		Non-body Q&A		Cultural-related Q&A	
	M	SD	M	SD	M	SD
Year 3	0.24	0.2459	0.3	0.2539	0.86	0.1350
Year 4–5	0.18	0.2394	0.3	0.2357	0.64	0.2797
Year 6–8	0.04	0.1265	0.12	0.1398	0.52	0.2700

Table 3. Score of picture explanation and total score

Age group	Picture explanation		Total score	
	M	SD	M	SD
Year 3	1.6	1.1972	9.6	2.2828
Year 4–5	2.4	1.3703	11.8	3.2335
Year 6–8	4.5	0.7454	16.1	2.3428

Table 4. Reaction time of different groups in tasks 1 & 2

Age group	Picture-selection		Non-body related Q&A	
	M (ms)	SD	M (ms)	SD
Year 3	3,559	3,183.6	2,203	1,627.4
Year 4–5	2,517	1,720.2	1,995	431.8
Year 6–8	2,101	2,479.1	1,239	620.5

choice of metonymic pictures improves with age. This improvement is most obvious between the group of year 4–5 and the group of year 6–8 ($F = 18.123$, $P = 0.000$).

In the analysis of the data "total score", which consists of four parts (picture-selection score, picture-explanation score, task 2 score and task 3 score), we find an improvement, not very significant though, from the group of year 3 to the group of year 4–5 ($F = 3.089$, $P = 0.096$). However, from year 4–5 to year 6–8, a significant development can be observed in children's metonymy comprehension ($F = 11.597$, $P = 0.003$).

4.1.3. Reaction time

Table 4 presents the one-way ANOVA result of the reaction time data of the three age groups in picture-selection task and task 2 (story Q&A concerning non-body related metonymies). The participants' reaction time in picture-selection task does not have significant difference, either between year 3 and year 4–5 ($F = 0.829$, $P = 0.375$), or between year 4–5 and year 6–8 ($F = 0.109$, $P = 0.668$). In task 2, the age-6-to-8 group shows a significant improvement in reaction rate compared to the group of year 4–5 ($F = 10.002$, $P = 0.005$).

4.2. Eye-tracking results

Heatmaps in §4.2.1 are generated by Data Viewer (64 bit 4.1.1) and the figures in §4.2.2 are generated by Numerical Python (NUMPY) packages in Python.

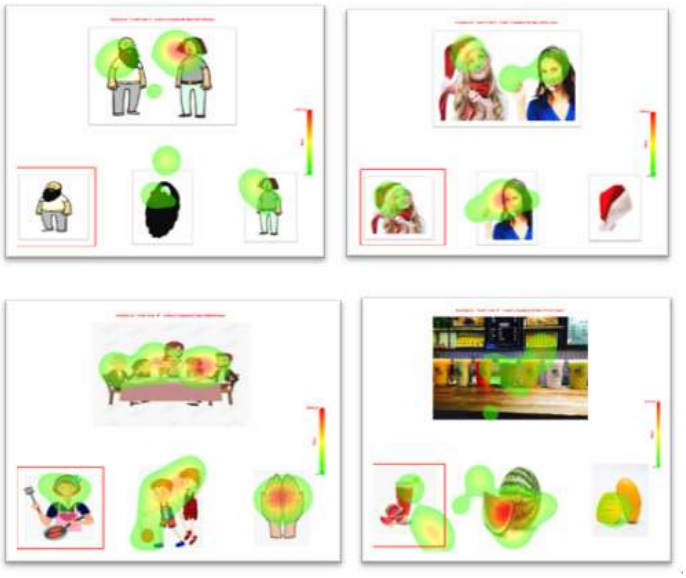


Figure 6. 3-year-olds' fixation area.

Figure 7. 4-to-5-year-olds' fixation area.

4.2.1. Gaze heatmap

The four heatmap examples in **Figure 6** show the 3-year-olds' fixation area during the 10,000 ms after the target utterance is heard. The four heatmaps cover all the four difficulty levels (from low to high clockwise) according to Figure 3. The pictures surrounded by red squares are the metonymic pictures, namely the target areas.

As can be concluded from the heatmaps, the fixation of age-3 participants on the different areas of the pictures are, to a large extent, at chance—there is no obvious preference between the three option pictures, and there is also no noticeable competition between the target (metonymic) picture and the literal picture.

The four heatmap examples in **Figure 7** show the 4-to-5-year-olds' fixation area during the

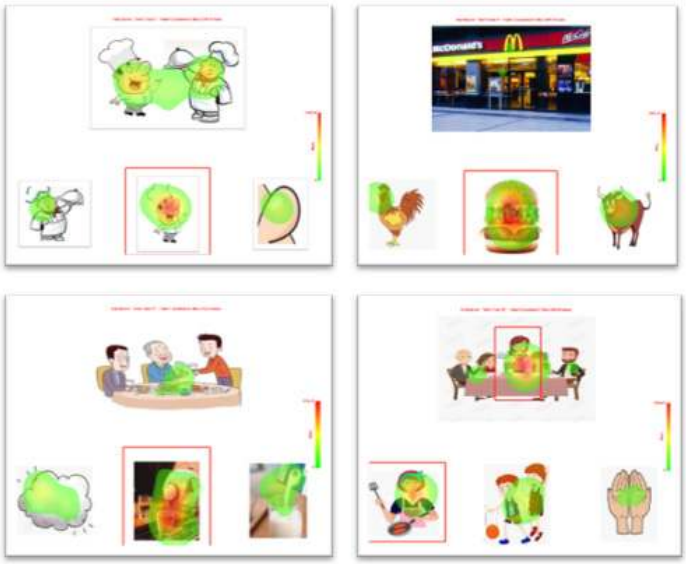


Figure 8. Over-6-year children's fixation area.

10,000 ms after the target utterance is heard. As can be concluded from the four heatmaps, the age-4 and age-5 children show a high preference to the literal pictures, whose fixation competes with fixation on the target pictures.

Concerning the pictures of lower difficulty levels of metonymies (the two in the first line), children of age-4 and -5 fix their eyes on irrelevant pictures significantly less, compared with the random fixation of 3-year-olds. However, as for higher difficulty levels, the gaze heatmaps of age-4-to-5 children are still largely disordered and random.

The four heatmap examples in **Figure 8** show the over-6-year children's fixation area during the 10,000 ms after the target utterance is heard. As can be concluded from the heatmaps, over-6-year participants show a significantly different and improved fixation heatmap compared with the other two age groups.

Compared with the heatmaps of age-4-to-5 children, over-age-6 participants' heatmaps in low difficulty level pictures show a clear preference towards target pictures and thus a less competition between target pictures and literal pictures. Furthermore, concerning the metonymies of higher difficulty levels (the second line), the heatmaps of over-6-year children show a lower randomness and the target areas are more heated.

4.2.2. Fixation proportion

Figure 9 presents the proportion of fixation on the target areas (in the trials containing lower difficulty levels of metonymies) during the 10,000 ms after the audio target utterance is played. The three lines represent the fixation performance of the three age groups respectively.

Fixation is detected and recorded by the eye-tracker when an eye movement lasts over 50 ms within the range from 0.5° to 2.0° (EyeLink Data Viewer 4.1.1, 2019). Thus, target fixations are hereby defined as the fixations whose coordinates fall within the area of target metonymy pictures.

Since the target area in each trial covers approximately 15% of all the picture areas, the fixation

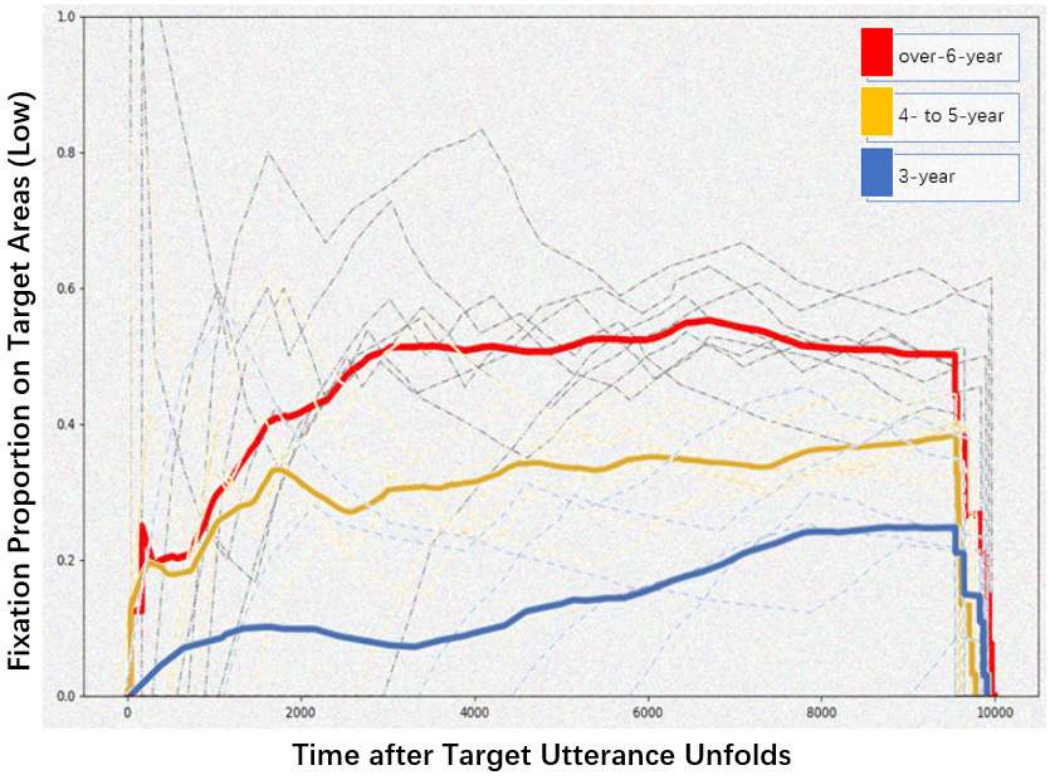


Figure 9. Fixation proportion on target areas (low difficulty levels).

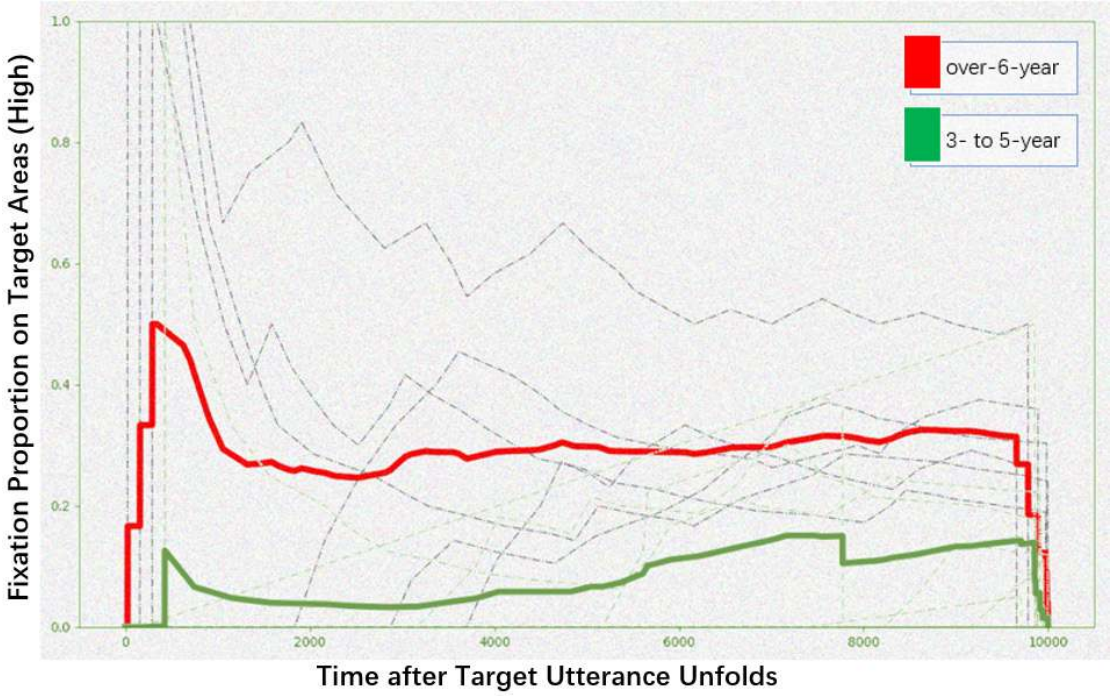


Figure 10. Fixation proportion on target areas (high difficulty levels).

of the 3-year-olds, though with a rise after the target utterance unfolds, can be regarded as largely at random. For the older age groups, the proportion of target fixations witnesses a rise as the audio target utterance is played. However, the target fixation proportion of over-6-year participants shows an

even more significant rise and levels out at around 0.55 as time passes.

Figure 10 presents the proportion of fixation on the target areas (in the trials containing higher difficulty levels of metonymies) during the 10,000 ms after the audio target utterance is played. The red line represents the fixation performance of the over-6-year participants and the green line, the performance of the other two age groups.

When being exposed to metonymic expressions of higher difficulty levels, both the age-3 children and the age-4-to-5 children produced a relatively random fixation pattern between different areas, with the target fixation consistently lower than 0.2 along the 10,000 ms.

However, target fixation proportion of the oldest group shows a different and improved feature. As can be concluded from Figure 10, their target fixation proportion, though being lower compared with Figure 9, is significantly above random, which means that their comprehension ability of higher difficulty levels of metonymies has developed at this stage.

Between Figure 9 and Figure 10, there is another comparison noted between lower and higher difficulty levels of metonymies. In different age groups, children's target fixation proportion in trials of higher difficulty levels is noticeably lower than that in trials of lower difficulty levels, except for the 3-year-olds whose fixation appears largely random in each level.

4.3. Discussion

4.3.1. U-shape tendency in picture selection

A U-shape demonstrating the trajectory of children's metonymy comprehension development was found by Falkum's (2017) study and an explanation for this phenomenon was provided in her following study (Falkum, 2020) that the inferior performance of group of age-4-to-5 can be explained as the result of a literal preference of the children at year 4 and 5.

However, the data collected and analyzed in the present study indicate some different characteristics, compared with the U-shape, concerning children's metonymy comprehension development.

The present study has not found a direct U-shape in children's performance in the picture selection task. However, the author has found a tendency towards the U-shape. Combining the data in Table 2, which show limited difference between age-3 and age-4-to-5 in picture selection, with the data in Table 3, which show a significance difference between the two groups in picture explanation, we can find that children's ability of picture explanation, compared with the accuracy of picture selection, marks a more noticeable improvement with age. Although the present study has not found a U-shape, the author can still argue that the method of picture selection, to a large extent, masks children's ability of pure metonymy comprehension, especially at the stage of year 4 and 5, and that the method of retrospective interview can alleviate this masking influence. In addition, the author could also argue that the alleviation of the masking influence is also due to the fact that the participants knew in advance that they would be asked to explain their choices of pictures, so that they were influenced less by the literal preference in the process of selection.

Furthermore, in the process of the behavioral experiment, the experimenter noticed some phenomena which may stand for the age-4 and -5's "literal preference". For example, in the second

task, when asked "What was boiling, the stadium itself or the spectators?[2]", one 4-year-old participant replied: "One player robbed the ball off the defender and ran, and the opposing players wanted to stop him and make his team lose the game." This is an irrelevant response to the question; however, it denotes a literal preference of the child—though possibly unable to understand the question, he still wanted to say something to show that he is in the conversation. Another example is that when asked "what do people prefer, the food they make or the name McDonald's itself?[3]", one 4-year-old children answered: "My mother took me to Pizza-hut." The literal preference of the 4-to-5-year-olds may be a significant contributor to the U-shape; however, it is not the only factor.

Another factor that may contribute to the U-shape in previous studies and the tendency to U-shape in the present study is a fake superiority of the performance of the 3-year-olds. In the data collection process of picture selection, the experimenter coded irrelevant choices as "X"s; however, in the data analysis process, these "X"s, together with literal choices, were all entered as wrong answers, which might have been a problem. The author has found that the wrong answers of 3-year-olds contain far more "X"s than other groups', indicating that the 3-year-olds made their choices largely at chance. Accordingly, it is reasonable that we suspect that the right answers of the 3-year-olds also have a high randomness. In this way, the positive performance of the 3-year-olds might have been faked up by this randomness, and thus the U-shape or the tendency to the U-shape can be observed.

4.3.2. U-shape explained by eye-tracking

Slight difference was detected between the 3-year group and the 4- to 5-year group; however, in the eye-tracking experiment, both the heatmaps and the fixation proportion indicate that, from age-3 to age-4-to-5, there is a significant development in children's metonymy comprehension ability.

The "fake superiority" and the high randomness of the 3-year-olds' metonymy comprehension can also be explained and thus supported by the data of the eye-tracking experiment. The heatmaps of the 3-year-olds show that the participants' fixation is highly random between different areas. Although, during the process of the eye-tracking experiment, the experimenter asked the children to choose one picture according to the story they heard and some of them gave correct feedbacks by picking the target pictures, it is still reasonable to argue that their answers were, to a large extent, by chance, which can be explained by their "disordered" fixation heatmaps.

Compared with the heatmaps of the 3-year-olds, age-4 and age-5 participants' heatmaps show that the irrelevant picture areas are significantly less heated than the target areas and the literal areas, which means that they no more look at the pictures at random and that their metonymy comprehension has started to develop, though possibly masked by a literal preference.

4.3.3. Age-6 as a crucial stage

The present study, supported by the data collected, argues that age-6 might be a crucial stage for children's metonymy comprehension development.

[2] The whole story and the question were: "There is a football game between China and Japan. In the last minute, the Chinese team scored a winning goal, and the whole stadium was boiling (a metonymy that only makes sense in the Chinese context). So, please tell me, what was boiling, the stadium itself or the spectators?"

[3] The whole story and the question were: "Mcdonald's and Dicos are two famous fast-food companies. Between them, more people prefer Mcdonald's. So, please tell me, what do people prefer, the food Mcdonald's make or the name Mcdonald's itself?"

The error rate data, experiment score data, and the reaction time data all indicate a significant improvement of the oldest group in metonymy comprehension competence, compared with other groups. In detail, during the picture explanation task, nearly all the participants of the oldest group could clearly point out the relation of contiguity between part and whole, for example, "you are asking about the beard, not the beard itself! Beard is not human being!". More surprisingly, we attempted to ask the participants of the oldest group to explain their answer in task 2, and some of them replied "The kettle cannot be inside itself. Only the water can, so it is the water that is boiling." and "Van Gogh would die if you put him on the wall![4]", which accurately tell the difference between metonymic meaning and literal meaning.

The eye-tracking heatmaps can also support the argument. Compared with the data of the other two age groups, the data of over-age-6 participants in the eye-tracking experiment generate a more ordered heatmap, in which target pictures received significantly more fixation than other areas. Even concerning metonymies of higher difficulty levels, the heatmaps of the oldest group also show a clear difference and regularity compared with the other two groups.

The factors contributing to the cruciality of the stage are manifold. Children at the age of four are able to construct the relation of similarity between concepts, thus able to comprehend metaphorical expressions (Zheng, 2008). Children's comprehension of metonymy, as has been found in the present study and previous studies (Falkum, 2017; Jiang, 2019; Köder and Falkum, 2020), develops earlier from the age of three. However, due to the poor vocabulary and a lack of social-cultural experience, the metonymy comprehension ability at the age of three is a rather incomplete one. It is after the kindergarten stage (6 years old) that children's vocabulary gets enriched due to primary school education or preschool education (for children ready to attend school in the next year) where children receive more systematic trainings on language. This, to a large extent, activates children's contiguity-comprehending ability, which already exists at the age of 3, and it also helps this metonymy comprehension mechanism to function between a wider range of concepts.

The cruciality of this stage may be also due to the procedural features of children's cognition development. Children's cognition of language, which develops correlatedly with their cognition of the physical objects (Liu and Hong, 2000), shows a tendency from concrete to abstract and goes through a mixed trajectory of continuity and discontinuity (Hudspeth and Pribram, 1990). "Discontinuity" here means that although children's language cognition improves with age, there are still clear phases and stages. Children at the age of 6 finish the stage of literal language development and enters a phase during which children's abstract thinking develops rapidly, which may help to explain the superior performance of the oldest group, and also may shed light on language education for preschoolers.

4.3.4. Comprehension development in different metonymies

Previous studies on children's metonymy comprehension development either investigated children's metonymy competence of one certain type of metonymy (Köder and Falkum, 2020), or stereotyped different types of metonymies as "metonymy" as a whole when designing their experiments (Falkum, 2017; Jiang, 2019). However, against this stereotype, the present study has found that children's competence of comprehending metonymies of different difficulty-levels shows different

[4.] These two examples are about the stories and questions in task 2.

developmental features.

From metonymies concerning human-body part-whole relations, to metonymies of non-body relations, and finally to the cultural-related metonymies, children's error rate, according to Table 2, rises with the increasing difficulty of the metonymies. Furthermore, the third task was where all the age groups answered over half of the questions wrong (error rate: year 6–8, 0.52; year 4–5, 0.64; year 3, 0.86), which forms a contrast with the error rates in other tasks.

Furthermore, heatmaps and figures generated from the eye-tracking data also tell the difference of children's comprehension development in different difficulty levels of metonymies. As for lower difficulty levels, participants' target fixation proportion demonstrates a steady improvement with age from age-3 to over-age-6; however, higher difficulty levels of stimuli give rise to a lower target fixation proportion of all age groups as compared to their performance in less difficult trials, even with the age-3 and age-4-to-5 children giving random fixations in higher difficulty levels of trials.

This result indicates that, even though there is development with age in children's cultural-related metonymy comprehension, children's metonymy comprehension competence not only improves with age, but also can be different between different types of metonymies. As the present study aims to investigate children's pure comprehension development of metonymies, the author hereby argues that children's inferior performance in the third tasks does not imply a low metonymic ability of children, but suggests that children's metonymy comprehension ability can be heavily masked by cultural elements.

5. Conclusion

The goal of the present study was to investigate the developmental trajectory of children's pure metonymy comprehension ability. The focus of this study was on children's metonymy comprehension development with age and children's metonymy comprehension development in different types or difficulty levels of metonymic expressions. Contributive points mainly include the modified behavioral + eye-tracking experiment method and the anti-stereotypical exploration into different types of metonymies when studying children's comprehension.

Data from the modified multi-task behavioral experiment, together with the eye-tracking data, provided innovative insights into how to further explain the U-shape developmental feature as reported in previous research. Concerning the U-shape tendency, the author argues that behavioral experiment design, the literal preference of the age-4 and -5 children, and the randomness of age-3 children's feedbacks jointly contribute to the U-shape tendency observed in the behavioral experiment. The findings of this study also suggested that, when studying children's metonymy comprehension development, the concept of metonymy could not be generalized, since the present study found that the different types and difficulty levels of metonymic expressions may render different stages of comprehension development. Although there are masking effects from cultural elements in higher difficulty levels of metonymic expressions, when children enter the stage of six years old, their metonymy comprehension ability, in both lower and higher difficult levels, shows an obvious improvement compared with other younger age groups, which suggests the cruciality of the stage of age six.

For future studies on children's figurative language comprehension development, this empirical

study would suggest that researchers avoid stereotypical operation of concepts, since different types of a certain figurative device may differ in difficulty, and thus in developmental stages. Furthermore, as the present study has not yet made a thorough enquiry into how different types of metonymies can influence the developmental trajectory of children's metonymy comprehension, the author recommends conducting more detailed investigation into different types of metonymies. Furthermore, researchers who are interested in children's metonymy comprehension are advised to set a wider range of variables in both behavioral and eye-tracking experiments, including, if possible, family backgrounds, personality, and (in eye-tracking) saccade events analysis.

Furthermore, although this study provides insights and implications for further studies, it comes with certain limitations. Due to the pandemic and the relatively short time span of the research, the author did not manage to recruit as many children as expected to participate in the experiments, which might have affected the experiment results. In the next stage, the author still has an ambition to expand or do more on the basis of the present results. The author plans to further improve the experiment instruments and hopefully, to invite more children to participate.

Conflict of interest

No conflict of interest was declared by the author.

References

Alač M and Coulson S (2004) The man, the key, or the car: Who or what is parked out back. *Cognitive Science Online* 2: 21–34.

Asch SE and Nerlove H (1960) The development of double function terms in children: An exploratory investigation. In: Kaplan B and Wapner S (eds.) *Perspectives in Psychological Theory: Essays in Honor of Heinz Werner.* New York: International Universities Press, pp.47–60.

Barcelona A (2000) On the plausibility of claiming a metonymic motivation for conceptual metaphor. In: Barcelona A (ed.) *Metaphor and Metonymy at the Crossroads: A Cognitive Perspective*. Berlin: Mouton de Gruyter, pp.31–58.

Carston R (2002) *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Corbett EPJ and Connors RJ (1999) *Classical Rhetoric for the Modern Students.* New York: Oxford University Press.

Dong W (2014) *Metaphor Cognition and Its Educational Application in Children Aged from 3 to 6.* PhD Thesis, Zhejiang University, China.

EyeLink Data Viewer 4.1.1 [Computer software] (2019) Mississauga, Ontario, Canada: SR Research Ltd.

Falkum IL (2019) *Metaphor and Metonymy in Acquisition: A Relevance-theoretic Perspective. Relevance, Pragmatics and Interpretation.* Cambridge: Cambridge University Press.

Falkum IL, Recasens M, and Clark EV (2017) "The moustache sits down first": On the acquisition of metonymy. *Journal of Child Language* 44(1): 87–119.

Gardner H (1974) Metaphors and modalities: How children project polar adjectives onto diverse domains. *Child Development* 84–91.

Hilpert M (2006) *Keep an Eye on the Data: Metonymies and Their Patterns.* New York: Mouton de Gruyter.

Hudspeth WJ and Pribram KH (1990) Stages of brain and cognitive maturation. *Journal of Educational Psychology* 82(4): 881–884.

Jiang X (2019) An empirical study of preschoolers' metonymic ability development. *Modern Foreign Languages* (4): 487–500.

Just MA and Carpenter PA (1980) A theory of reading: From eye fixations to comprehension. *Psychological Review* 87(4): 329–354.

Kliegl R, Nuthmann A, and Engbert R (2006) Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General* 135(1): 12–35.

Köder F and Falkum IL (2020) Children's metonymy comprehension: Evidence from eye-tracking and picture selection. *Journal of Pragmatics* 156: 191–205.

Lakoff G and Johnson M (1980) *Metaphor We Live by*. Chicago: University of Chicago Press.

Langacker RW (1993) Reference-point constructions. *Cognitive Linguistics (Includes Cognitive Linguistic Bibliography)* 4(1): 1–38.

Liu D and Hong X (2000) Study major factors influencing early-childhood language acquisition. *Journal of Chongqing University* 6(1): 103–106.

Liu Z and Mi X (2008) A study of preschool children's comprehension of metaphor. *Journal of Sichuan International Studies University* 6: 40–44.

Lutz A (2017) Norse loans in middle English and their influence on late medieval London English. *Anglia* 135(2): 317–357.

Mendelsohn E, Robinson S, Gardner H, et al. (1984) Are preschoolers' renamings intentional category violations? *Developmental Psychology* 20(2): 187.

Panther KU and Radden G (eds.) (1999) *Metonymy in Language and Thought* (Vol. 4). Amsterdam/Philadelphia: John Benjamins Publishing.

Panther KU and Thornburg L (2005) The role of conceptual metonymy in meaning construction. In: Ruiz de Mendoze Ibanez FJ and Pena Cervel MS (eds.) *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*. Berlin: Walter de Gruyter GmbH & Co., KG, pp.353–386.

Peirsman Y and Geeraerts D (2006) Metonymy as a prototypical category. *Cognitive Linguistics* 17(3): 269–316.

Peng X and Zhang S (2009) Trope ability development: An empirical study of preschool children. *Foreign Language Teaching and Research* 5: 359–364.

Piaget J (1964) Part I: Cognitive development in children—Piaget development and learning. *Journal of Research in Science Teaching* 40.

Radden G and Kövecses Z (1999) Towards a theory of metonymy. *Metonymy in Language and Thought* 4: 17–60.

Rundblad G and Annaz D (2010) Development of metaphor and metonymy comprehension: Receptive vocabulary and conceptual knowledge. *British Journal of Developmental Psychology* 28(3): 547–563.

Özcaliskan S (2005) On learning to draw the distinction between physical and metaphorical motion: Is metaphor an early emerging cognitive and linguistic capacity? *Journal of Child Language* 32(2): 291–318.

Van Herwegen J, Dimitriou D, and Rundblad G (2013) Development of novel metaphor and metonymy comprehension in typically developing children and Williams syndrome. *Research in Developmental Disabilities* 34(4): 1300–1311.

Yan C and Huang M. (2009) On children's U-shaped development and its inspirations on children's development assessment. *Early Childhood Education. Educational Sciences General* 33: 1–4.

Zhang H and Lu W (2010) *Cognitive Metonymy.* Shanghai: Shanghai Foreign Language Education Press.

Zheng L (2008) *A Study of the Rhetoric Features of Preschool Children's Language*. PhD Thesis, Nanjing Normal University, China.

Zhou R (2003) A preliminary study of children's developmental trend of metaphorical ability of time. *Modern Foreign Languages* 3: 222–231.

ORIGINAL ARTICLE

A semantic contrastive study of Chinese and English verb "跑/run" from the perspective of cognitive semantics

Peiyao Fang, Qi Liu*

School of Foreign Languages, Hangzhou Dianzi University, Hangzhou, China

Abstract: Based on corpus data, this paper finds that the semantics of "跑/run" have similarities, which lie in the consistency of the central meanings and common usages. The differences are: 1) the meanings of "run" are significantly more than those of "跑", resulting in a great divergence in the numbers of meanings; 2) compared with "跑", the implications of "run" display a higher level of specificity.

On the basis of cognitive theories, the paper has yielded the following findings. The semantic overlaps of "跑/run" root in their identical prototypical meanings. The semantic differences are generated by the following factors: 1) the image-schema distribution of "run" is wider with more abstract representations than that of "跑", due to which many unique semantic items come into being; 2) semantic systems of "跑/run" adopt different metaphorical methods, contributing to more comprehensive cognitive domains of the semantic mapping and closer semantic interrelations of "run" than those of "跑".

Keywords: contrastive study; cognitive semantics; polysemy; verb

***Corresponding author:** Qi Liu , School of Foreign Languages, Hangzhou Dianzi University, Hangzhou, China; liuqi@hdu.edu.cn

Received: April 23, 2022; **Accepted:** June 14, 2022; **Available online:** August 20, 2022

Citation: Fang P and Liu Q (2022) A semantic contrastive study of Chinese and English verb "跑/run" from the perspective of cognitive semantics. *Forum for Linguistic Studies* 4(1): 100–115. DOI: 10.18063/fls.v4i1.1458

1. Introduction

From the 1970s till now, cognitive linguists have investigated the relationship between language and cognition from various perspectives and levels and revealed the cognitive basis of speech generation and understanding (Huang, 2012). Though scholars have been constantly infusing brand new vitality into cognitive linguistics, such as making contrast from syntactical (Lv, 1999; Huber, 2017; Deng, 2021), typological (Talmy, 1991; Zhang, 2010; Liu, 2019; Aktan, 2020; Bai, 2022), pragmatical and semantical perspectives (Nida, 1975; Gries, 2006; Glynn, 2012; Shuai, 2021) by adopting quantitative and qualitative approaches as well as synchronic, diachronic, and dialectical methods, there is still room left for the studies of polysemous words and semantic, cognitive contrast studies

of synonymous verbs in Chinese and English.

As Chinese-English commonly used verbs, "跑/run" have a wide range of application, when matched with different meaning items, new meanings emerge. Significant as the research might be, little study conducted ever centered on the words' original and metaphorical meanings. In an effort to identify the real usages of "跑/run" and to find out the semantic similarities and differences while explaining the underlying mechanism from the perspective of cognitive linguistics, the paper makes several hypothesis: 1) the most frequently used meanings of "跑/run" are the same; 2) the amount of meanings of "跑/run" is relatively equal; 3) the specificity of the words demonstrates discrepancies to some extent. To serve the purposes of the analysis, the study material and method is presented first, then the results of semantic contrast are addressed. Afterwards, cognitive theories are incorporated to discuss cognitive factors that result in the semantic similarities and differences. The last part is the conclusion and suggestions for further studies.

2. Material and methods

To make a detailed and complete semantic description, the data used in this paper are all real usages drawn from the corpora BCC and COCA for these colloquial expressions constitute an indispensable part of modern vocabular and play an important role in meaning classification (Wang, 2016), Ant Conc (3.5.7) is applied as the analysis software.

In this study, 1,200 pieces of Chinese and English sentences are drawn respectively from BCC and COCA. The quantity of data is based on the statement of Xu (2007: 68) that at least 500 pieces of corpus material should be collected to analyze a linguistic phenomenon, the more, the better. Besides, it has been attested by Mahpeykar and Tyler (2015) that a total number of 800 tokens of verb-particle is valid for identifying the central sense and the extended senses. The semantic analyses of "跑/run" are conducted by following the procedures below.

Firstly, a total number of 1,200 sentences of "跑/run" from BCC and COCA are selected without restrictions on sources, and the first 1,200 pieces of data are downloaded in the txt format.

Secondly, unsuitable ones where "跑/run" are used as nouns as deleted based on *Xinhua Dictionary* (Zhongguo shehui kexue yuanyu yanjiusuo cidian bianjishi, 2011), *A Century Chinese-English Dictionary*, *Oxford Advanced Learner's Dictionary (9th edition)* (Hornby, 2015), *Collins Cobuild Advanced Learner's Dictionary (9th edition)*, dwindling the valid sentences to 1,134 for Chinese and 1,136 for English.

Thirdly, the data is processed in Ant Conc to reveal their collocations and the meanings of the collocations of "跑" are evaluated as follows: 1) to leap forward alternatively using two or four legs; 2) to leak; 3) to rush about to get things done; 4) to operate vehicles; 5) to escape; 6) to head for a place; 7) to achieve a result; 8) to deviate. The meanings drawn from the collocations of "run" are: 1) to go faster than walking; 2) to escape; 3) to operate; 4) to manage; 5) to be or near a level; 6) to meet someone unexpectedly; 7) to drive; 8) to campaign; 9) to use up; 10) to become; 11) to leave for a place; 12) to compete; 13) to make a machine function.

Finally, the sentences of "跑" are inspected one by one in complete contexts based on several Chinese dictionaries as *Xinhua Dictionary* (Zhongguo shehui kexue yuanyu yanjiusuo cidian bianji-

shi, 2011), *Comprehensive Dictionary of Regional Dialects of Modern Chinese* (Li, 2002), *Modern Chinese Dictionary (Version 7)* (Zhongguo shehui kexue yuanyu yanjiusuo cidian bianjishi, 2016) and the corpus data of "run" is closely analyzed in succession based on *Oxford Advanced Learner's Dictionary (9th edition)* (Albert, 2015), *Collins Cobuild Advanced Learner's Dictionary (9th edition)*.

3. Results

The top 8 frequently used meanings out of a total number of 21 senses of "跑"are shown in **Figure 1**.

The top 10 frequently used meanings out of a total number of 46 senses of "run" are shown in **Figure 2.**

No matter which language is studied, it is necessary to obtain a profound understanding by comparing it with other languages (Shen, 2001). As the meanings of the two verbs have been laid out above, a contrastive analysis is needed to further highlight the differences between "跑" and "run". Comparing the two tables of the meanings of "跑/run", we have the following findings.

3.1. The identical meanings

Firstly, the similarity of "跑/run" lies in that the meanings with the highest frequency in two languages are identical, several more meanings overlap to some extent.

Hence, hypothesis one is supported. Corpus data analysis shows that "fast pedestrian motion" in "跑/run" is the most frequently used term, exhibiting its central role (Durkin and Manning, 1989: 2).

Therefore, the two verbs both embody the self-motion feature as the core, which paves way for other overlapping meanings of "跑/run" including "to escape, to operate, to drive, to observe, to get sth. done, to leak". Of those meanings, the shared agents are human, animal, machinery, vehicle, gas or liquid, computer system, eyes, and purposes in common are "to exercise or compete, to get sth.

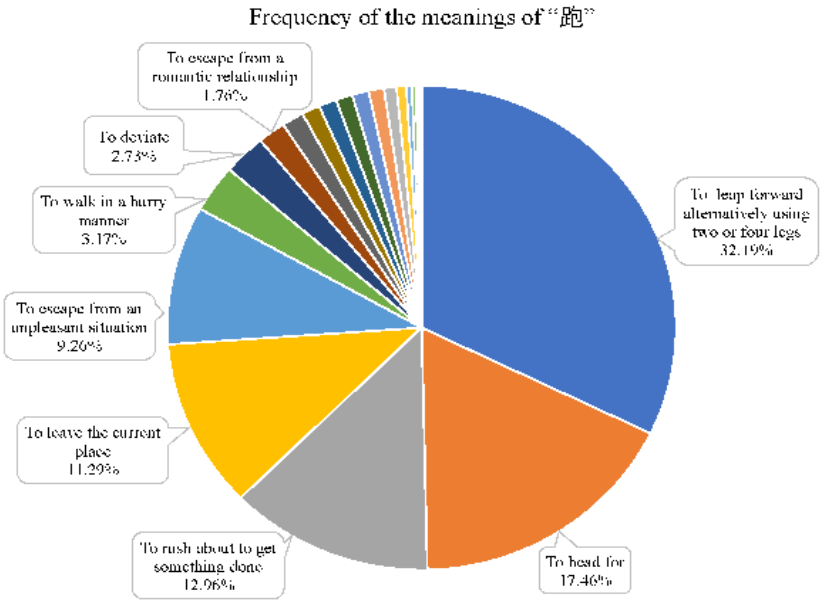


Figure 1. The frequency of the meanings of "跑".

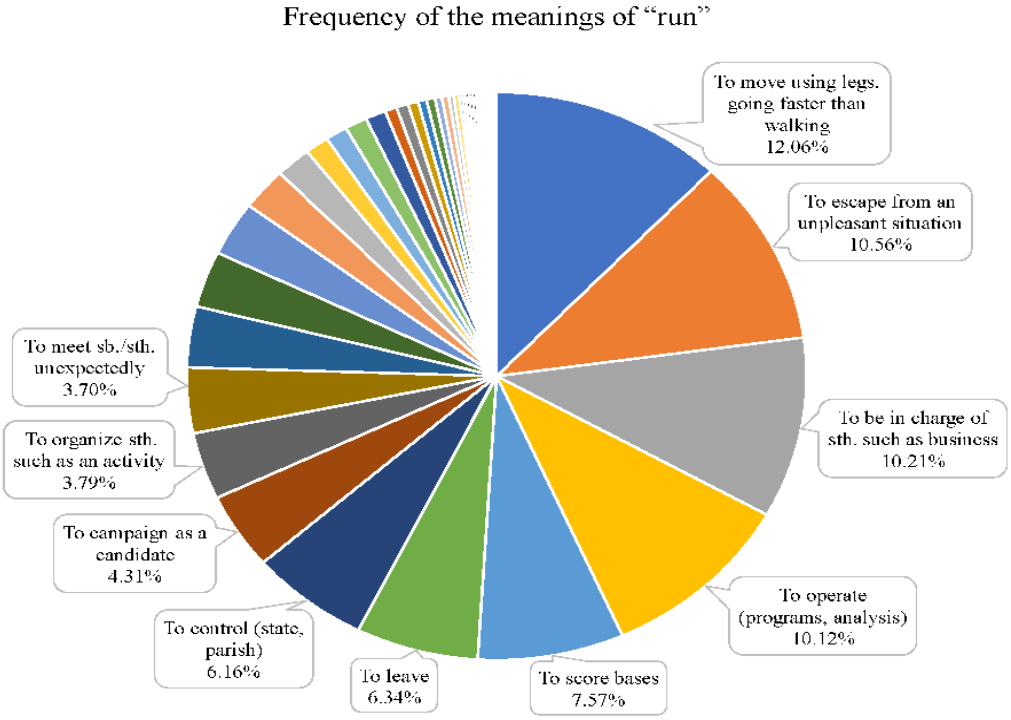


Figure 2. The frequency of the meanings of "run".

Table 1. The meaning with the highest frequency

	"跑"	**"run"**
Amount	365	137
Frequency	32.19%	12.06%

Table 2. The items of meanings of " 跑 /run"

	"跑"	**"run"**
The number of items	21	46

done, to start a new relationship, to observe in a close way, to gain information, to function or support to function". We can identify that the meanings all originate from the common experience of human beings, reflecting the embodiment essence of languages.

3.2. The disparity of amount of meanings

Secondly, one prominent difference is that the meanings of "跑" are far less than those in English, unlike the second hypothesis.

Many collocations of "run" display a high degree of abstraction, which are absent in Chinese, leading to the gap between the numbers of the meanings. Being a self-motion verb as "跑/run", the shared semantic features are "fast, pedestrian, motional, directional, continuous and consistent function as a whole, self-productive"; however, in different contexts, the characteristics are not all included, but the abstraction of one or more features.

To understand a word, one can start with its collocation. The words after "run" are business, or-

ganization, activities, parish, world, society, race, errands, blockade, human organs, commercials, people, number, aid, riot, story, shows in gallery, antibiotics, lines, views, or words, demonstrating the feature of state-highlighted and process-centered, that's why "run" develops state-highlighted meanings as "in charge of, manage, to break through" and process-centered meanings such as "to formulate, to mention to others, to experience", which are all absent in Chinese. In contrast, the unique objects of Chinese are fewer, including agent of motion, people to be contacted, result, role, or profession and are basically emotion-stressed, motivation-driven and result-centered, as a result, the meanings contained by "跑" as "to cost, travel to get things done, make a living" have no equivalent expression related to "run" in English.

3.3. The different level of specificity of meanings

Thirdly, another difference is that the usages of "run" display a higher level of specificity than those of "跑", conforming to the third hypothesis. The higher level of semantic specificity of "run" is achieved by detailed descriptions of its components including context, collocations such as agents, objects and sentence pattern; besides, causative and abstract meaning extension also effect to pave the way for meaning distinction.

Towards the same semantic core, the items of "run" are more specific, as shown in **Table 3**.

To begin with, the collocation of "run" is more restricted than that of "跑", leading to a higher level of specificity of "run" . This can be viewed from the semantic core "to play the key role". According to the classification of Gries (2006), among the most frequent usages of transitive senses of "run" are "to organize, to operate, to control" and "to be in charge of". All these four meanings have one common feature in essence, that they all involve human to "play the key role". Incorporating the meaning as "to operate program, computer system or analysis" as the usages in Chinese, "run" further illuminates its objects in detail so as to express in a logically and grammatically suitable manner. Likewise, from the semantic core of "to cover", the agents of "to disappear" of "跑" are generally human or objects, but for "run", the agents further includes play and event, which are all components absent in Chinese, therefore, we can tell that when the agent is clarified, the semantic orientation tends to be clearer.

Table 3. The different levels of specificity of "跑/run"

The meanings / Semantic core	"跑"	"run"
To play the key role	To operate	To organize (activity);
		To be in charge of (business);
		To operate (program, analysis);
		To control (country, parish);
To cover	To travel a long distance	To extend from one place to another;
		(A play, event) to continue for a period;
		To provide;
To function	To function	To function;
		To be made to function using a particular source of energy;

Afterwards, in terms of the same semantic core, "run" has more diversified meanings though applying analogy between spatial and temporal, concrete and abstract motion event leads to more complex semantic features of "run".

Concerning the transition of spatial motion to temporal one of "run", the tendency can be viewed from "to cover". For "跑", the verb is used to indicate the movement of an object, followed by distance to describe a process, which can be understood as "to travel a long distance", as shown in the 1st example.

1. 以前一副轮胎顶多只能跑一万三千公里，现在不超载，气压足，又控制速度，可以跑二万五千公里。

   (In the past, a pair of tires could only ***run 13,000 kilometers*** at most, but now if the vehicle does not overload, the tire pressure is sufficient, and the speed is controlled, it can run 25,000 kilometers.)

While for "run", when the emphasis is laid upon the agent, the semantic meaning of movement in the spatial scope is "to extend from one place to another" as example 2 clarifies, representing an image-schema transformation.

2. The chapel was the first building put up, and all the pipes ***run from*** it. Most plumbing problems will start here.

Under the circumstances above, both "跑" and "run" present their traits as a motion verb, reflecting the process of movement, but for "run", the meanings are more explicit. Just as Zhang (2021) stated, English and Chinese adopt different strategies for encoding spatial relationships. Apparently, the spatial motion can turn to the temporal domain, as in "continue for a period", the extension in space can also be understood as the continuation in time in example 3.

3. The weekly two-hour workshops ***run 10 a.m. till noon*** through Aug. 9, Thursdays at Robie House.

And it's also noticeable that the usages of "run" go beyond the concrete domain to the abstract domain, and the trend also reveal itself in the application of the meaning "to cover", from "to extend" to "to provide". Based on encyclopedia knowledge, when an object extends, it covers the ground underneath. In a similar way, when the object to be covered is regions by "aid", which seems to lead to nonsense, only when understood as "provide" can "run" be correctly interpreted as in example 4.

4. She couldn't get out of China, but she still risked getting infected to ***run aid*** to isolated regions.

Moreover, through causative extension, "run" in enabled to derive new meanings, as the most straightforward and most productive method (Levin, 1993: 31), it further clarifies the semantic component and benefits the determination of meanings, for the similarities between related ones. For example, to express the meaning "to function", the agents are generally machinery or transportation vehicles, while for "run", considering that equipment functions by the support of gears or other parts, special energy or resource is involved to sustain a continuous movement, therefore, "run" also has the meaning "be made to function by using a particular source or energy". Through semantic ex-

tension, the context of semantic generation is described in more detail, which, in turn, improves the accuracy of language descriptions.

4. General discussion

In the preceding part, the study on semantic contrast between the two words has been carried out. And this section attempts to elaborate on the reasons for similarities and differences from three aspects.

4.1. Prototype category and meaning formation

According to prototype category theory, among the different meanings of a polysemous word, one (in rare cases, more than one) should be dominant, whose status is determined in terms of such attributes as derivatives, frequency of usage and centrality.

Defining the prototypical meanings of "跑/run" could cast a clear insight into the study. The reason is that the confirmation of the prototypical meaning is closely relate to the formation of a word's meaning. This term is also often regarded as a typical member of word's meanings. Tyler and Evans (2001) put forward corresponding criteria for judging the typicality of a meaning, such as the earliest formed one, the centrality of semantic network, and the asymmetric judgment of semantic praise or derogation or iconicity. Dirven and Verspoor (1998: 31) pointed out that there are three ways to determine the core or typical meanings of polysemous words:

1) Empirical method: when speaking of a word, the first meaning that comes to mind.

2) Statistical methods: the most frequently used term in polysemous words.

3) The term that serves as the basis for the extension of other ones.

The prototype meaning of the two, "fast pedestrian motion", is found in this part following a set of rigorous conditions, for the intrinsic nature of the two verbs is a contributing factor to their semantic overlap, which can be well clarified by their prototypical meanings.

Combined with relevant criteria and corpus data, the generation of prototypical meanings can be divided into the following aspects:

To begin with, "fast pedestrian motion" is the most frequent meaning used in early language acquisition. Theakston et al. (2001) reached this conclusion after studying the data of Manchester Children's Elemental Corpus. Aktan (2020) later mentioned 112 children aged 5 or 7 years old that the earliest learned phrasal verbs include "run-away", in which "run" means to leave quickly, which is consistent with the former conclusion. In Chinese, Cheng (2009) did a follow-up survey on a 3-year-old child and tracked daily communication and made several tables. The conclusion presents that "跑" often appears in front of the complement displaying trend such as "来, 去" and is interpreted as fast motion. Compared with adults who can not only use the verb and tendency complement as a collocation, but its metaphorical meaning, children acquire a smaller amount of meanings, but it is sufficient to prove the semantic characteristics of verbs acquired in early childhood.

Secondly, according to an etymological dictionary based on historical text analysis, that is, a corpus-based approach. To be precise, the origins and interactions of semantics and phonology are

complex and ambiguous. Even in this way, the diachronic semantics of "run" included the former stated prototypical meaning. For the diachronic development of "跑", it is often thought to derive from " 走 ", which means fast motion in ancient Chinese. According to The Scripta Sinica Database, "跑" initially appeared in Guan Hanqing Opera Collection in the Yuan dynasty, denoting the chase between the roles, consistent with the present meaning.

Thirdly, like many other English verbs, "run" is zero-derived when used as a noun, and its collocations are largely fixed. Besides, it does occasionally appear in proper nouns without many changes in its main meaning. For example, the movie title "Run Lola Run" and a certain event "the Sardine Run", all have the feature of "fast walk". Meanwhile, "跑" also presents similar characteristics in such words as "领跑" and "跑位", still, they expose features of fast motion. According to the explanation of Lakoff (1987) of the relationship between semantic prototypicality and markedness in the elaboration of metaphor, the meaning with the least formal constraints can be considered as unmarked and therefore, the prototype one. Specifically, the meaning "fast pedestrian motion" is followed by the most varied prepositional phrases. This meaning has the highest number of semantic tag attributes. It also suggests that most of the semantic variation exhibited by the word comes from this formal and semantic feature, which has the strongest word formation, which in turn, strongly supports its unmarked essence.

4.2. Image-schema and meaning expansion

Initially, image-schema theory is proposed by Lakoff and Johnson (1980) while describing the conceptual metaphor, which is "the dynamic structure of perceptual interaction and continuous reappearance, which gives our experience coherence and structure".

The difference in the numbers of meanings of "跑/run" can be explained by image-schema analysis. With "run" covering all the major types of image schemas and "跑" only covering a part of them, it is revealed that the verb "run" highlights its high level of abstraction and temporal nature while "跑" presents its spatial traits.

The eight types of image schemas divided by Lakoff (1987: 267) are container schema, source-path-goal schema, center-periphery schema, link schema, front-back schema, up-down schema, part-whole schema, linear order schema.

The distributions of image schemas of "跑/run" are shown in **Table 4**.

Table 4. The image schemas of "跑/run"

Image schemas	Sum of "跑"	Sum of "run"
Source-path-goal schema	17	32
Container schema	2	1
Link schema	1	4
Center-periphery schema	0	1
Front-back schema	1[1]	4
Part-whole schema	2	0
Linear order schema	0	2
Up-down schema	0	2

[1] The meaning "to operate" in Chinese occurs in several domains, therefore, the total number exceeds the item of "跑".

Several conclusions can be drawn from the table above:

First and foremost, one of the discrepancies of "跑/run" is that though both "跑" and "run" concentrate in large numbers in the source-path-goal domain but the specific meanings differ. Though verifying the common bodily experience of human, the meanings of the two words are not identical. And it is revealed that Chinese emphasizes spatiality, while the meanings of "run" are rather abstract.

The constituent elements of the schema are starting point, ending point, path, and direction. For "跑", meanings that focus on spatial movement fall into this category, such as "to leave for, to get something done, to travel a long distance". More accurately, the motion verb puts more emphasis on space in the Chinese language, denoting the source, goal, process and becomes more generalized, leading to the result of action. For "run", the meanings that refer to abstraction motion can also be categorized here, such as "to operate, to become different in a particular way, to look through, to caress" for the elements are all included in these meanings but shift abstract and causative domains.

Secondly, one unique point of "跑" is that only the meaning "to operate" in Chinese covers 3 domains, no similar case is found in English as in "source-path-goal, front-back, part-whole" schemas due to the agents of the verb phrase can be a complete vehicle, such as a car, or a train used in the sentence "跑不上点" mentioned above, since the front of human body faces the direction of body's movement, being delay carries the meaning of failing to arrive on time, which is, behind the schedule. The meaning can also be applied in the part-whole schema owing to the agent "tires". As a part of a vehicle, tires can only function when supported by engines and other automobile parts; nevertheless, the agent can be used independently in the phrase "a pair of tires could only run 13,000 kilometers at most", while the part-whole schema involves an integral whole, a part and connection type, the meaning also falls in this category.

Thirdly, the distribution of the image schemas is complementary in terms of center-periphery schema, part-whole schema, linear order schema and up-down schema and the meanings of the overlapping schemas are not identical. In the following is a detailed explanation.

The container schema, which foregrounds the visual, auditory, and other sensory impressions left by the agent (Langacker, 2009: 117–118), can be used metaphorically in many non-spatial domains, in most cases, the agents of "跑" are living entities. But in the example above, "跑" can mean "to be away from home", with the essence emphasizing the behavior of "to leave the life partner". The reason lies in that a married couple is often regarded as a whole, therefore, companionship is "located" as "internal", and the behavior of leaving each other is to reach the external, as the promise of marriage is characterized by the container schema and the whole-part schema, since the behavior of leaning the former lover is expressed by the verb phrase "run off".

In contrast, container schema occurs more frequently in English as the occasional encounter of unexpected things and people, normally carried by the set phrase "run into". The former can be understood as being caught in a tricky situation, and the latter is more vivid as knocking into others. The meaning of disappearance in Chinese is personal characteristic, and it expands to the pass of time in English. According to the definition of the source-path-goal schema, the elements include a starting point, an end point, a path, and a direction, most of the meanings of spatial motion conform to this schema. In metaphors, the end point is often interpreted as the goal, which will be analyzed

in more detail later when we talk about metaphors, but here it is worth explaining why the meanings of "to manage, to execute/operate" are placed under this schema.

Under the link schema, here are four meanings, including two entities and connection relation, specifically, "to disappear" for Chinese. In the instance "the sale is not a deal, the righteousness is lacked", the benevolence and consideration between cooperation partners serve as a bridge to seal a deal; when the two sides fail to reach an agreement, the connection is broken. In English, the four meanings under this schema are "to accept and adopt", "to hang out", "to confuse" and "to coexist", carried by the verb phrase as "run with". The general meaning is to juxtapose two things, or to compare them, or to emphasize the connection between them, while this classification does not appear in Chinese.

The center-periphery schema only exists in English, where "run" is used as "to spread". The schema involves an entity, a center, and a periphery, which are all included in the meaning of "to spread". Among the examples that imply this meaning, one agent is "disease", and the spread of disease involves a center and affected areas, the latter is influenced and determined by the former one.

On the one hand, "to execute/operate" means to start a machine or software application that can then operate on its own or under the constant control of a human being. But on the other hand, "to manage" usually contains more abstract meanings, such as the management of the institutions, organizations. Though sometimes it's hard to decide to how much extent human involvement is in an event, in essence, "run" all points to the motion from a particular starting point to an expected direction. It might be a continuous operation for software, steady advancement for business. In addition, meanings in Chinese and English concentrate in this schema. Most meanings in Chinese involve specific movement, but in English, many meanings are changed to abstract directions, such as management and program operation mentioned above, while representing many meanings such as continuity, extension, and coverage in the temporal span.

The front-back schema is derived from the embodied human experience, where the front of the human body faces the direction of movement and is also metaphorically used for the direction of time. In English, the emphasis is laid upon different states of the moving agent in a successive time period, such as in the instance "the sky is running dark", referring to exterior changes in a continuous process.

Only meanings of "run" occur in the linear order schema and up-down schema. Linear schemas are generally used to represent temporal order. In English, it's exemplified by two meanings "continue for a period" and "time to pass by". No relevant corpus data can be found in Chinese for only the meaning of "run" incorporate the abstract domain. Likewise, the up-down schema only includes English language material, and is often followed by "over, up". Representative meanings are "cause to increase" and "to be near or at a particular level", all exactly suiting the metaphor "up is more" and having a certain spatial orientation, either going up or down. Of course, they are also applied in an abstract domain, such as "run up" when prices rise, which still has a logical connection between up and down.

4.3. Metaphor and meaning interrelation

Metaphor is a cross-domain mapping based on the principle of similarity, reflecting the correlation between two concepts or elements in the same cognitive domain. Sweetser (1990) pointed out

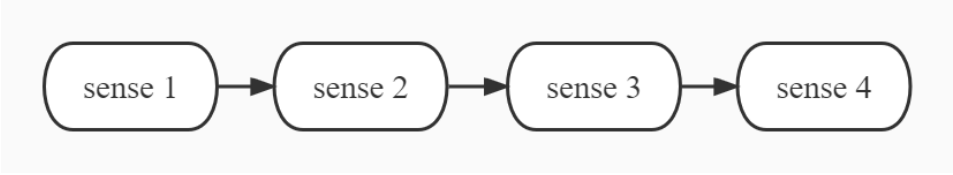


Figure 3. The method of meaning expansion—the meaning chain.

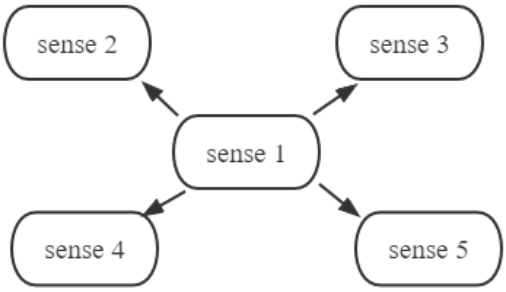


Figure 4. The method of meaning expansion—the ripple type.

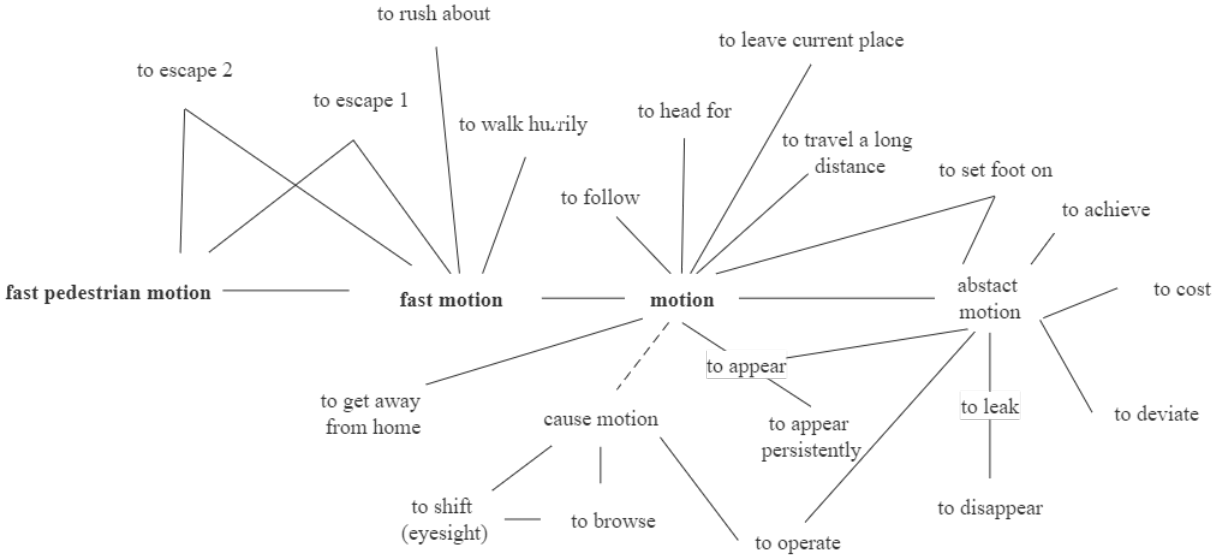


Figure 5. The radial network of "跑".

that the use of metaphor is the result of the continuous expansion and extension of meaning.

The differences of level of specificity of "跑/run" can be well explained by metaphor applied in a close meaning extension. Since metaphor is the mapping or projection between conceptual domains (Sun, 2010), the application of metaphor and cognitive domain concerning "跑/run" can be illustrated in radial networks of meaning interrelation of the two.

There're two types of meaning extension methods. The first one is the "meaning chain", in which meanings only share some semantic features with their immediate neighbors because of their iconicity in a certain part. The concept was proposed by Taylor (1995) when he analyzed seven example sentences including "climb".

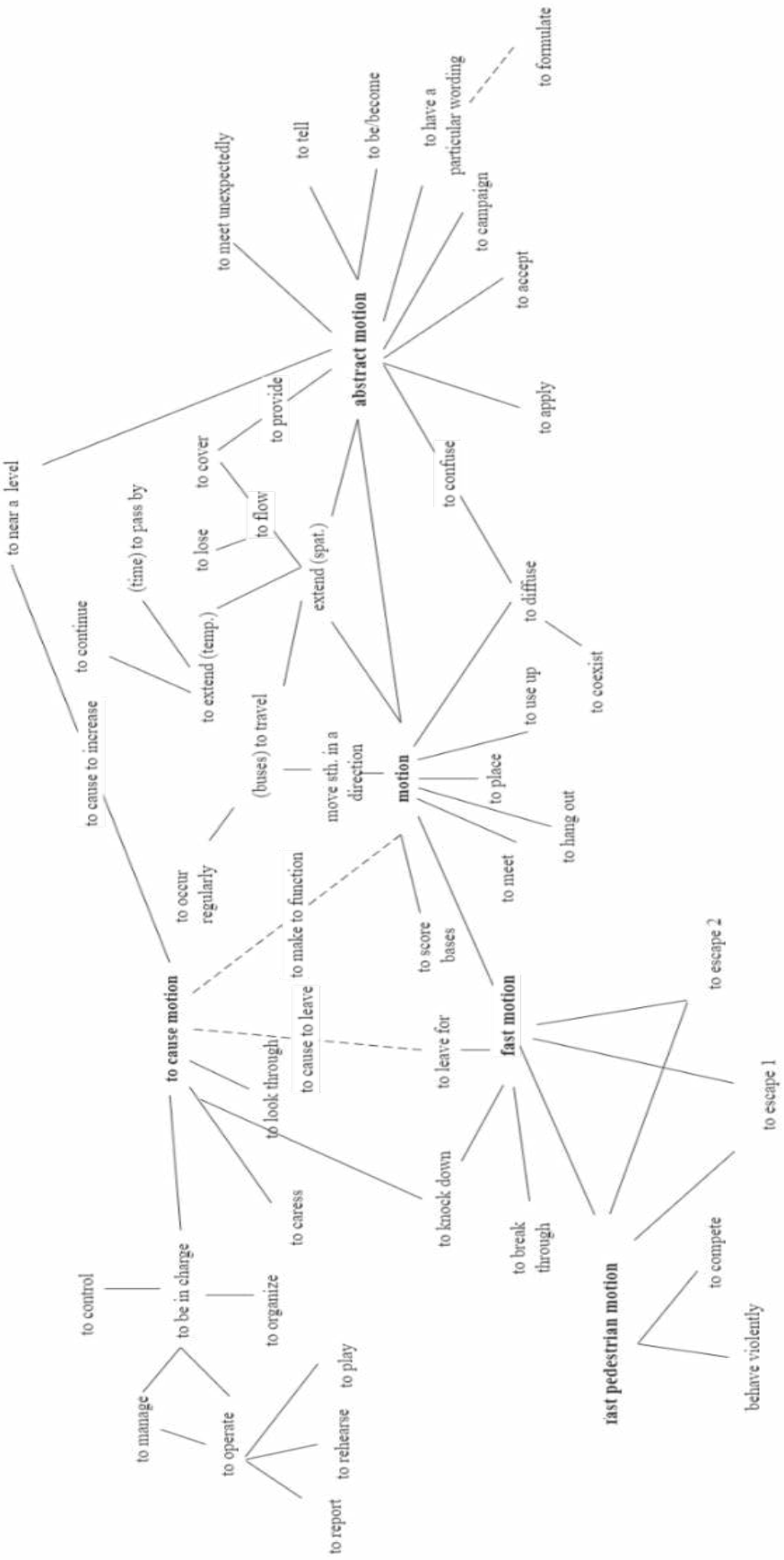


Figure 6. The radial network of "run".

The second one is the "ripple type", depicting the existence of a prototypical meaning in polysemous words, while other meanings derive from it in other directions, as shown in the **Figure 4**. Sweetser (1990: 9) put it forward and argued that if a word used to mean A, now it means B, then B is not produced out of blue, but there would be a stage when the word means both A and B, eventually, A was lost.

In fact, semantic extension is often a combination of the above two, as shown in the radial networks of "跑/run" below.

Dotted lines are drawn to show the causation links between motion and causative motion, which is what Levin (1993: 311) has termed as "induced action alternation". The prototypical meaning of "跑/run" is termed "fast pedestrian motion" for consistency. Meanings are shortened to five letters at most for readability. "To escape from an unpleasant situation" and "to escape from a romantic relationship" are referred to as "escape 1" and "escape 2", "spatially" and "temporally" are shortened as "spat." and "temp." in the radial network of "run".

Figures 5 and **6** visually present the interrelations of different meanings, and several conclusions can be drawn:

To begin with, in radial networks, there is a prototypical sense from which other partially similar meanings are derived, the latter may become the source of semantic expansion, and then the other meanings are separated. As it can be clearly seen in the figures, the semantic extension of "跑/run" combines the above two meaning extension methods together.

The closely related semantics come from a few senses with the strong power of word formation. From "fast pedestrian motion" to "fast motion", "motion", "abstract motion" and "causative motion", in this process of semantic divergence, the similarity between meanings is getting lower and lower, but there is still some internal connection, that is, family resemblance. The subsequent reason lies in the continuous improvement and development of people's cognitive ability, new meanings continue to extend from the prototypical meaning, and the extended ones can further become the source meaning, based on which expands the iconographic "family" system.

Take "to escape" and "to leave a place for something", as an example. They both take root in "fast motion", and are characterized by "quickness". Even in this way, they are applicable in different situations with reference to the distinguishing characteristics described above. The former often appears in adverse and critical situations, while the latter appears to be in a hurrying state without any external threat. However, the agent of "escape" is highly coincident with the prototype meaning, so the two are connected in the network diagram.

Secondly, the semantic meanings of "跑" are mainly mapped from the domain of physical attributes to psychological factors, and the attributes of familiar things are mapped to abstract things by analogy. Concerning "跑", meanings are relatively scattered while the derived meanings are evenly distributed. In terms of virtual displacement, relatively more meanings overlap with English ones. While for "run", the scope of agents, the paths are more abstract, extending from spatial to the temporal range. Its category expansion undergoes changes from [+life], [+controllable], [+displacement] of prototypical metaphorical extension agents to [-life], [-control], [+ displacement] role of the main body; finally, it extends to the virtual displacement subjects with the features [-active], [-controllable] and [-displacement]. The process reflects that English lays emphasis on verb's timeliness (Wang,

2007).

Lastly, the meanings of "run" are of a higher cognitive experience. They are developed on the basis of mental cognition of humans and are accumulated based on social experience with prominent cognitive traits serving as the connections between entities and psychological or social states.

The conclusion is also consistent with Lakoff's (1987: 248), who distinguished between two types of mental scanning: sequential scans and overall scans. Sequential scanning is a continuous recording of actions or events that occurred in different periods, highlighting a complex temporal relationship just like the meaning expansion trend exhibited by "run". While "跑" focuses on the simultaneous scan of all the components in the scene, consequently, the emphasis is on the result and reflects a static state.

5. Conclusion

This paper, dictionaries-based and corpus-driven, attempts to make a contrast of verb "跑/run" on the basis of analysis of real usages from corpora data and bring out the underlying mechanism of meaning extension and interrelations between Chinese and English from the perspective of cognitive theories.

According to the semantic contrast analysis, the present thesis yields the following major findings: first and foremost, no matter in dictionary entries or from actual usage, the meanings placed on top of the two verbs are the same. Besides, the gap between the number of senses is quite noticeable, and meanings of "run" are far more than those of "跑". Thirdly, English presents a higher level of specificity than Chinese, the reason can attribute to the detailed descriptions of its components through profile shift and causative meaning extension, paving the way for meaning distinction.

The investigation to reveal the cognitive causality behind semantic similarities and differences of the words draws conclusion as follows:

Firstly, the most basic meaning of "跑/run", fast pedestrian motion, turns out to be the prototypical meaning, from which diverse senses come into being.

Moreover, regarding the image schema, the distribution of the two words varies. The occurrences of meanings in the source-path-goal image schema holds the highest rate, but the meanings categorized here are rather different, in the meantime, the distribution of other meanings is complementary.

Furthermore, according to the contrast of the radial analysis, the overwhelming meanings of "run" distribute massively and are connected logically. Covering several cognitive domains into an abstract dimension, "run" displays a higher level of cognitive processing while the semantic meanings of "跑" cover a basic cognitive domain and are mainly mapped from the domain of physical attributes to psychological factors by analogy.

Though considerable effort has been made to enrich the present study, deficiencies and limitations are still inevitable.

To begin with, the paper focuses on semantic contrast studies, excluding the influence of syntax or other grammatical factors, but the syntactic components might interact with semantic factors,

which may play a prominent role in words' interpretation.

Besides, more inclusive, and comprehensive language data and more accurate analysis tools will undoubtedly lead to more enlightening conclusions.

The further studies can be carried out from the following aspects: firstly, diversified perspectives are to contribute to the research's comprehensiveness such as syntactical, morphological, and pragmatical approaches. Additionally, the expansion of data samples and tools for semantic analysis can improve the accuracy of the research.

Conflict of interest

The authors declared no potential conflict of interest.

References

Aktan A (2020) Effects of second language on motion event lexicalization: Comparison of bilingual and monolingual children's frog story narratives. *Journal of Language and Linguistic Studies* 16(3): 1127–1145.

Austin J (1961) The meaning of a word. In: Urmson JO and Warnock GJ (eds.) *Philosophical Papers*. Oxford: Clarendon, pp.39–54.

Bai X (2022) Xuni weiyi shijian gainian jiangou zhong de zhuanyu he yinyu [Metonymy and metaphor in conceptual construction of virtual displacement event]. *Foreign Languages Research* 39(1): 34–38.

Cheng L (2009) *Hanyu Ertong Quxiang Dongci Xide Ge'an Yanjiu* [*A Case Study of Chinese Children's Directional Verb Acquisition*]. Master's Thesis, Capital Normal University, China.

Collins H (2018) *Collins COBUILD Advanced Learner's Dictionary* (9th edn.). New York: Harper Collins Publishers.

Croft W and Cruse DA (2004) *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cruse DA (2009) *Lexical Semantic*s. Cambridge: Cambridge University Press.

Deng Y (2021) [A cognitive contrastive study of English and Chinese supernormal V-O constructions: Taking "eat/have" and "chi" constructions as examples]. *Chinese Language* 18(3): 31–39.

Dirven R and Verspoor M (1998) *Cognitive Exploration of Language and Linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Durkin K and Maaning J (1989) Polysemy and the subjective lexicon: Semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research* 18: 577–621.

Geeraerts D (1997) *Diachronic Prototype Semantics: A Contribution to Historical Lexicology.* Oxford: Oxford University Press.

Glynn D (2012) The many uses of run: Corpus-based methods and socio-cognitive semantics. In: Glynn D and Robinson JA (eds.) *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Amsterdam: John Benjamins, pp.17–144.

Gries ST (2006) Corpus-based methods and cognitive semantics: The many senses of to run. In: Gries ST and Stefanowitsch A (eds.) *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*. Berlin: Mouton de Gruyter, pp.57–91.

Hornby AS (2015) *Oxford Advanced Learner's Dictionary* (9th edn.). London: Oxford University Press.

Huang J (2012) Guo wai renzhi yuyanxue yanjiu de zui xin dongtai [The latest development of cognitive linguistics research abroad]. *Modern Foreign Languages* 35(1): 87–94, 110.

Huber J (2017) *Motion and the English Verb: A Diachronic Study*. Oxford: Oxford University Press.

Lakoff G and Johnson M (1980) *Metaphors We Live by*. Chicago: Chicago University Press.

Lakoff G (1987) *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: Chicago

University Press.
Langacker RW (2009) *Investigations in Cognitive Grammar.* Berlin, New York: De Gruyter Mouton.
Levin B (1993) *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
Li R (2002) *Fangyan Dacidian* [*The Dialect Dictionary*]. Nanjing: Jiangsu Jiaoyu Chubanshe, pp.4324–4330.
Liu Q (2018) Yinghan fangshi weiyi dongci chayi ji dui yingyu cihui jiaoxue qishi [Differences between English and Chinese motion verbs and enlightenments to English vocabulary teaching]. *Hangzhou Dianzi Keji Daxue Xuebao (Shehui Kexue Ban)* 14(2): 74–78.
Lv S (1999) *Xiandai Hanyu Babai Ci (Zeng Ding Ben)* [*Eight Hundred Words in Modern Chinese, the Updated Version*]. Beijing: The Commercial Press, pp.425–426.
Mahpeykar N and Tyler A (2015) A principled cognitive linguistics account of English phrasal verbs with up and out. *Language and Cognition* 7(1): 1–35.
Nida E (1975) *Componential Analysis of Meaning: An Introduction to Semantic Structures*. Mouton: The Hague, pp.18–138.
Shen J (2001) [A survey of studies on subjectivity and subjectivisation]. *Foreign Language Teaching and Research* (4): 268–275, 320.
Shuai Z (2021) Cong "fangshi" dao "jieguo" de yuyi yanbian ji lilun sikao—Yi "song, shuai, diu, zou, pao" wei li [The semantic evolution from "method" to "result" and its theoretical thinking—Take "送、摔、丢、走、跑" as an example]. *Studies of the Chinese Language* (6): 733–746, 768.
Sun G (2010) Yinyu renzhiguan yu yinghan yinyu shiyi [The concept of metaphoric cognition and the interpretation and translation of English and Chinese metaphors]. *Henan Daxue Xuebao* 50(6): 141–145.
Sweetser E (1990) *From Etymology to Pragmatics.* Cambridge: Cambridge University Press.
Talmy L (1991) Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society* **17: 480–519.**
Taylor J (1995) *Linguistic Categorization: Prototype in Linguistic Theory*. Oxford: Clarendson Press.
Theakston AL, Lieven EVM, Pine JM, et al. (2001) The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language* 28(1): 127–152.
Tyler A and Evans V (2001) Reconsidering prepositional polysemy networks: The case of over. *Language* 77(4): 724–765.
Wang N (2016) Dan yinjie dongci fei zhudong yongfa yu xiaoji jieguo [Involuntary usage and negative consequential meaning of monosyllabic verbs]. *Zhongguo Yuwen* (2): 228–240, 256.
Wang W (2007) Yinyu xing ciyi de shengcheng he yanbian [The generation and evolution of metaphorical meanings]. *Foreign Languages and Their Teaching* (4): 13–17.
Xu J (2007) *Yuyan Guihua yu Yuyan Jiaoyu* [*Language Planning and Language Education*]. Shanghai: Xuelin Press, pp.68.
Zhang J (2010) Zailun yinghan Shuang bingyu goushi [On the English-Chinese double object constructions]. *Foreign Languages Research* (2): 8–13.
Zhang K (2021) Chouxiang kongjian fangwei guanxi goushi de jiegou tezheng yu renzhi liju [Structural characteristics and cognitive motivation of abstract spatial orientation relation constructions]. *Modern Language* 44(3): 360–371.
Zhongguo shehui kexue yuanyu yanjiusuo cidian bianjishi (2016) *Xiandai Hanyu Cidian (Di 7 Ban)* [*Modern Chinese Dictionary (the 7th Version)*]. Beijing: The Commercial Press, pp.981–982.
Zhongguo shehui kexue yuanyu yanjiusuo cidian bianjishi (2011) *Xinhua Cidian (Di 11 Ban)* [*Xinhua Dictionary (the 11th Version)*]. Beijing: The Commercial Press, pp.376.

ORIGINAL ARTICLE

The L3 acquisition of English tense-aspect system by Uygur speakers with L2 Mandarin Chinese

Qianping Gu*, Muradil Adil

School of Foreign Languages, Southeast University, Nanjing, Jiangsu, China

Abstract: This paper examines the role of Lexical Aspect Hypothesis (LAH) and linguistic typological similarity in the L3 acquisition of English tense and aspect among Uygur speakers with L2 Mandarin Chinese (Chinese hereafter). LAH asserts that the emerging verbal inflections at the early stage of language acquisition primarily function as markers of the lexical aspect and thus predicts universality for acquisition of tense and aspect. However, with an assumption of language transfer, the typological closer relationship of Uygur with English in terms of the tense and aspect system was expected to trigger L1 transfer in L3 acquisition. The study analyzed the English tense and aspect forms used by the participants (N = 25) for verbs of four distinct lexical aspects (50 target items) in contexts of past. The result shows that the lexical aspect influences the appropriate use of past tense—past tense marker aligned with telic predicates (achievements and accomplishments), *-ing* with activities (for inappropriate uses), and nonpast with states (for inappropriate uses), and the influence is observed at each proficiency level. The results show little evidence for language transfer in the acquisition of the English past tense, either from L1 Uygur or L2 Chinese; instead, the data suggest that L3 acquisition of tense and aspect is more subject to acquisitional universality (LAH).

Keywords: tense-aspect; Lexical Aspect Hypothesis; L3 acquisition; language

***Corresponding author:** Qianping Gu, School of Foreign Languages, Southeast University, Nanjing, Jiangsu, China; qianping.gu@outlook.com

Received: April 27, 2022; **Accepted:** June 16, 2022; **Available online:** August 23, 2022

Citation: Gu Q and Adil M (2022) The L3 acquisition of English tense-aspect system by Uygur speakers with L2 Mandarin Chinese. *Forum for Linguistic Studies* 4(1): 116–131. DOI: 10.18063/fls.v4i1.1459

1. Introduction

Previous studies on second and third language acquisition of the grammatical categories of tense and aspect have found several factors that influence the use of tense-aspect markers of the target language (e.g., Bardovi-Harlig and Reynolds, 1995; Robison, 1995; Shirai and Kurono, 1998; Eibensteiner, 2019; Vallerossa, 2021). One of those factors, which is a general factor that constantly influences the process of L2 and L3 acquisition, is the (learners' perceived) typological distance between the target language and the previously acquired languages (both L1 and L2), which could easily induce a language transfer in learning the tense and aspect system of a new language. Another

important and unique factor that has been noticed to play a role in language acquisition of tense and aspect is a semantic interaction of the lexical aspect with particular grammatical aspects or viewpoint aspects in Smith's (1997) term. This factor is known as the Lexical Aspect Hypothesis (LAH), which says that the tense-aspect markers used by learners, especially at the beginning of language acquisition, largely reflect the inherent lexical aspects of verbal predicates rather than their real tense meanings (Shirai and Andersen, 1995). As the learners' language proficiency increases, the influences of the two factors on the acquisition of tense and aspect begin to decline. But the influence of the lexical aspect could persist to the high level even when the learner is a native-like speaker (e.g., Salaberry, 2005).

In this study, we investigate the Uygur learners' L3 acquisition of English tense and aspect with a focus on the influence of the lexical aspect and typological similarity on the use of past tense in the L3 interlanguage. In this case of L3 acquisition, the involved languages—Uygur, Mandarin Chinese (Chinese hereafter), and English—are three typologically distinct languages. In terms of the tense system, both Uygur and English have grammaticalized tenses while Chinese is usually considered tenseless (see the discussion below). However, the results of the investigation provide no strong evidence for L1 transfer. Instead, it seems that L3 acquisition of tense and aspect is more subject to acquisitional universal (the LAH).

In the rest of the paper, Section 2 presents the background of the studies of the lexical aspect and the LAH proposed for language acquisition of tense and aspect. Section 3 reviews some studies of language transfer in L3 acquisition. Section 4 gives a brief description of the tense and aspect systems of Uygur, Chinese and English. The details of the present research are presented in Section 5, and the results are presented in Section 6 and a discussion in Section 7. Section 8 concludes the study.

2. Lexical aspect and Lexical Aspect Hypothesis

In the research area of lexical aspect, it has commonly been assumed that verbal predicates, verbs or phrases, bear a certain inherent temporal trait which establishes a temporal structure imposed on the events they describe (e.g., Vendler, 1957). Four classes of verbal predicates have been identified, namely, *achievement*, *accomplishment*, *activity*, and *state*. Each of the four classes is associated with a distinct temporal structure which is usually characterized in terms of three lexical aspectual properties, namely, dynamicity, durativity, and telicity, that is, whether the event is dynamic or stative, whether the event is instantaneous or has certain duration, and whether the verbal predicate imposes any endpoint or boundary for the event. The three binary contrasts provide the necessary features for characterizing the four verb classes as shown in **Table 1**.

Table 1. Lexical aspectual features of four verb classes

Features	**Verb classes/lexical aspectual classes**			
	Achievement	**Accomplishment**	**Activity**	**State**
Dynamicity	+	+	+	−
Durativity	−	+	+	+
Telicity	+	+	−	−

Achievement describes a dynamic, instantaneous event that has a natural endpoint, as illustrated in (1a). Accomplishment describes a dynamic event that has a process (i.e., durative) and progresses towards a specified endpoint, as illustrated in (1b). Activity describes a dynamic, durative event that has no inherent endpoint, as illustrated in (1c). State is not eventive and thus non-dynamic and atelic, but it is durative, as illustrated in (1d).

(1) a. Achievements: *notice a painting*, *arrive at the hotel.*

b. Accomplishments: *build a house*, *create a model.*

c. Activities: *walk around*, *ride a bike*.

d. States: *know the story*, *seem worried*.

The lexical aspect has been found to affect the use of tense-aspect markers in language acquisition—in second and third language acquisition as well as in first language acquisition. In the second and third language acquisition, it has been found that the emerging verbal inflections for tense and aspect function in a distinct way from their use in the target language. Language learners use a particular inflection form for a verb in accordance to its inherent lexical aspectual property while the original semantic meaning of the inflection is largely ignored. This redundant marking of the inherent aspectual property leads to a hypothesis called The Aspect Hypothesis (Andersen, 1991), Primacy of Aspect Hypothesis (Robison, 1990), or LAH (Shirai and Andersen, 1995), which says that the verb inflections function as indicators of the inherent lexical aspects of the predicates, which are different from the original functions.

One of the common errors of language learners in the use of tense and aspect markers is that telic predicates are more likely to be used with the past tense form, especially when talking about a "punctual" and "completive" situation, as in (2a); by contrast, the present participle form (*-ing*), a marker of imperfectivity (progressive), is often used when talking about a situation with no definite duration, as in (2b).

(2) a. And here, do you want something, you *buyed* it.

[In the United States, if you want something, you buy it.] (Robison, 1990: 328)

b. The one guy tell me, "I want to you ***making***" one pant.

[One guy might tell me, I want you to make me a pair of pants.] (Robison, 1990: 326)

Many studies of second or third language acquisition have reported the association of verbal inflection with the lexical aspect, which is in effect even at a high level of language proficiency. For example, Salaberry (2005) studied the influence of L2 Spanish on L1 English speakers' acquisition of L3 Portuguese. The results show that even for the learners with a high level of proficiency in L3 Portuguese, who show a sophisticated, native-like morphological system of L3 tense, there are still significant differences between those L3 learners and Portuguese native speakers in their morphological marking of static verbs. In Bardovi-Harlig and Reynolds's (1995) investigation of L2 acquisition of English simple past tense of adult learners of various language proficiency in English, they found that the acquisition of the simple past tense undergoes a gradual process, at the beginning of which the past tense is significantly undergeneralized, and the lexical aspectual effect is attested at

each level of language proficiency.

The effect of the lexical aspect on verbal morphology is not only observed in second and third language acquisition but it has also been found in child (first) language acquisition development. Prior studies have found such effect in the early stage of child language acquisition of various languages, including English, Italian, French, and Greek (see Bardovi-Harlig and Reynolds (1995) for the relevant references). In these cases, children treat the tense markers as indicating for aspectual meanings rather than their tense meanings. In the current study, we found that the influence of the lexical aspect on the use of the past tense marker and the progressive also exists in the case of Uygur learners' L3 acquisition of English. Thus, it seems that the constraint captured by LAH is ubiquitous in learning a language.

3. Language transfer in third language acquisition

Studies of third language acquisition found that the prior linguistic knowledge influences the acquisition of L3, and such influence diminishes as the L3 proficiency increases. Previous studies argued for different sources of the transferred linguistic knowledge at the initial stage of L3 acquisition, including L1, L2 or both. For example, Jin (2009) investigated the language transfer problem of L1 Chinese–L2 English bilinguals in learning L3 Norwegian. Of the three languages, only Chinese allows null objects in its grammatical system; neither English nor Norwegian allows such linguistic expression. The study found that in making grammatical judgments, the participants accepted up to 72% of Norwegian sentences with null objects, while the acceptance of English sentences with null objects decreased to 57%. This study supports The L1 Factor Hypothesis which says that in learning L3 the learners transfer their L1 linguistic knowledge to their L3 system.

But it is notable that such studies cannot prove that L2 does not become a source of language transfer in the process of learning L3. In fact, many studies have found that L2 is the main source of language transfer in some aspects of third language acquisition, which prompted the researchers to put forward the hypothesis of L2 Status Factor. The main evidence for the role of L2 in L3 acquisition comes from Bardel and Falk's (2007) study of language transfer in the process of learning L3 Swedish or Dutch at an early stage. This study found that the learners' L3 expressions showed syntactic features of their L2. The two researchers explained this in terms of the distinction of procedural memory and declarative memory from the psycho-cognitive research field: L2 and L3, at least in the early stages, are stored in the declarative memory while L1 is stored in the procedural memory, thus L2 and L3 are cognitively closer, explaining why L2 has a greater influence on L3. Another explanation for the influence of L2 on L3 in the literature assumes that L1 is unconsciously suppressed by learners in the process of learning L3 and thus L2 becomes the main source of language transfer.

Moreover, there are two influential theoretical models of L3 acquisition that have received a lot of discussions in the literature. One is the Cumulative Enhancement Model (Flynn et al., 2004). According to this theoretical model, the acquired L1 and L2 provide cumulative linguistic knowledge which can be used by the learners in learning L3, thus both languages can be the source of language transfer for L3 acquisition. The evidence for this theoretical model mainly comes from the comparative study conducted by Flynn et al. (2004) on L3 and L2 learners who studied English as their target language. The results showed that trilingual learners with L1 Kazakh, L2 Russian and L3 English produced the same English restrictive attributive clauses as bilingual learners with L1 Span-

ish and L2 English, but different from bilingual learners with L1 Japanese and L2 English. Note that in terms of language types, both Russian and Spanish are head-initial languages and the modifying attributive clause is placed after the head noun. Kazakh and Japanese are the opposite. They are head-final languages and the attributive clause is placed before the head noun. The results could be explained in terms of the L2 Status Factor hypothesis; however, the researchers believed that there was a more reasonable explanation: once learners have acquired a language with an X grammatical feature, no matter L1 or L2, when learning a new language, if the new language has X feature, the learners will transfer it to the new language. In this experiment, the trilingual learners transferred the head-initial feature from L2 language (Russian), and the bilingual learners with L1 Spanish and L2 English transferred the feature from L1 language (Spanish); however, for the bilinguals with L1 Japanese and L2 English, since L1 Japanese is not a head-initial language, there is no source of transfer for them, which explains why they produced more errors in the language output.

Another theoretical model is called "The Typological Primacy Model" (Rothman, 2011, 2015). This theoretical model also assumes language transfer occurs at the early stage of L3 acquisition; however, it makes a broader hypothesis for the source of language transfer—instead of assuming the transfer is due to the similarity over a specific linguistic feature, it hypothesizes that the major motivation of language transfer is the typological similarity between the three languages. It emphasizes that it is the typological closeness relationship between languages that is perceived by learners at the psychological level that determines the source of transfer, although this may not be the case objectively. The close relationship may exist in all aspects of language: phonology, lexicology, morphology, and syntax. The research evidence for this theoretical model comes from Rothman and Cabrelli Amaro's (2010) study of trilingual learners with L1 English, L2 Spanish and L3 French producing "null subject" expressions in French. Spanish is a language that allows null subjects, while English and French do not allow such linguistic form. The study found that the learners' grammatical knowledge of L3 French allowed null subjects. The researchers believed that this was because the learners perceived the two Romance languages, French and Spanish, to be typologically closer and thus transferred the Spanish grammatical feature to French.

For the L3 acquisition of tense and aspect, researchers who took a language transfer perspective on the learners' L3 tense-aspect representation also argued for different linguistic sources of the transfer. For example, Eibensteiner (2019) found that in the acquisition of perfective and imperfective aspect in L3 Spanish, the L1 German speakers transferred their L2 English aspectual knowledge—note that L1 German and L2 English are equally typologically different from L3 Spanish as both of them are Germanic languages; he argued that this transfer was due to the structural similarities between English and Spanish. Foote (2009) found that participants who were able to recur to a previously acquired Romance language, either L1 or L2, performed better in a judgment task which tested the (im)perfective knowledge of another Romance language (L3), which suggested that language transfer may occur in the tense-aspect domain.

4. Tense and aspect in Uygur, Chinese and English

Tense and aspect are two common grammatical categories that deal with two different types of temporal relations. Tense is the grammaticalized expression for indicating the deictic relation of a particular event, action or state to the speech time, which could be a precedence relation (past tense),

overlapping relation (present tense), and subsequence relation (future tense). According to Comrie (1976), aspect is about how the speaker views the internal temporal structure of a situation, which could be a totality of the situation (perfective aspect) or a partial view of the situation (imperfective aspect). The imperfective aspect is usually divided into progressive, habitual and continuous aspect. It has been found that almost all languages on the globe have grammaticalized formatives for aspect whereas most languages do not have grammaticalized tense.

Uygur is an agglutinative language. It has grammaticalized tenses which are expressed by morphology agglutinated to the verb stem with an obligation to indicate for person. There are two types of past tense in Uygur. One is simple past and the other one is termed distant past in the traditional grammar. The simple past tense is the common form to talk about past while the distant past tense indicates a sense of remoteness. An example is presented in (3).

(3) a. Simple past:

bardim

bar-dim

go-PST.1SG

"I went"

b. Distant past:

*bar**gan**men*

*bar-**gan-**men*

go-DIST.PST-1SG

"I went"

The imperfective aspect (progressive) formally distinguishes between present and past tenses. The verb morphology of the imperfective form inflects for tense and also shows agreement in person of the subject. For example, in (4), the imperfective form o'qi-watidu indicates present tense, progressive, and third person of the subject.

(4) U hazir kitab(ni) o'qi-watidu.

U	*hazir*	*kitab(-ni)*	*o'qi-watidu*
S/he	*now*	*book-ACC*	*read-PRS.PROG.*3SG

"S/he is reading a book now."

For a past tense example, consider (5). The imperfective form o'qi-ywtkan-di indicates past tense, progressive, and third person of the subject.

(5) Ahsham kechqurun u kitob(ni) o'qi-ywtkan-di.

Ahsham	kechqurun	u	kitob(-ni)	o'qi-ywtkan-di

yesterday evening s/he book-ACC read-PROG-PAST.3SG

"Yesterday evening s/he was reading a book."

The Chinese language has received more controversy over the issue of whether it has grammaticalized tense. The common view taken in the literature is that Chinese has no tense but only aspect (Li and Thompson, 1981). The most common tense-aspect marker is the verbal suffix -le, which is analyzed as a perfective aspect marker and may also be used as a past tense marker (Comrie, 1976; Ross, 1995; Sybesma, 2019; Gu, 2022). What is notable about -le is that it has certain requirement of the verbal predicates it combines with. It requires the verbal predicate to be dynamic, that is, it is compatible with achievements, accomplishments and activities but incompatible with states, which are illustrated in (6)–(9) respectively.

(6) *wǒ rènchū*-le *tā*.
I recognize-PFV him
"I recognized him."

(7) *wǒ shǔjià kàn-le sān-bù diànyǐng*.
I summer vacation watch-PFV three-CLF film
"I watched three films during the summer vacation."

(8) tā yóu-le yǒng.
he swim-PFV swim
"He swam."

(9) *tā céngjīng zài Nánjīng jūzhù*(*-le).
he once LOC Nanjing live-PFV
"He once lived in Nanjing."

It has also been pointed out that sentences with some activity verbs are less natural when they are suffixed with *-le* unless there are other modifications available in the sentences or the sentences are used in certain context which provides additional information (see the references in Sybesma, 2019). Normally such activity verbs are those VPs that have a bare noun as the object. For example, the sentence in (10) contains an activity *qí zìxíngchē* "ride a bike" suffixed with *-le*; however, the sentence sounds less natural unless *qù xuéxiào* "go to school" is used to indicate the goal of the action—in such case *-le* actually is combined with an accomplishment *qí zìxíngchē qù xuéxiào* "ride a bike to school" rather than an activity predicate *qí zìxíngchē* "ride a bike", or the sentence is used in a context in which it is used to answer a question about the transportation.

(10) *tā qí-le zìxíngchē ?(qù xuéxiào).*
he ride-PFV bike go school
"He rode a boke to school."

In other words, it is less common that activity verbs are suffixed with *-le*, though such expressions are not entirely ungrammatical.

There are two imperfective aspect markers in Chinese. The progressive aspect is expressed by a preverbal particle *zài* with the verb form remaining unchanged. An example is presented in (11).

(11) *tā* *(dāngshí/zhèhuì)* *zài* *chī* *yī-ge* *píngguǒ*.
he at.that.time/now PROG eat one-CLF apple
"He was/is eating an apple (at that time/now)."

The other imperfective aspect is durative aspect, which is expressed by the verbal suffix *-zhe*. It is used to emphasize the existence, continuity or durativity of a state or activity. For example, in (12), the use of *-zhe* diminishes the dynamic meaning of the verb *fàng* "put" and gives rise to a stative interpretation.

(12) *zhuō-shàng* *fàng-zhe* *yī-tái* *diànnǎo*.
table-on put-DUR one-CLF computer
"There is a computer on the table."

English has distinct grammaticalized tenses. The past tense is expressed by the verbal suffix *-ed* with exception of some irregular forms. In the aspect system, it has a fully grammaticalized progressive aspect which is expressed in a periphrastic construction, that is, "be + V-ing", which can be used in all tenses. The aspectual opposition is expressed by the non-progressive form (perfective aspect) and the progressive form (imperfective aspect). The former entails culmination of the denoted event if the predicate is telic (e.g., *eat an apple*) whereas the latter is a viewpoint into the internal temporal structure of the event and thus does not entail culmination when interacting with telic predicates.

5. The present study[1]

5.1. Research questions

The present study investigates the following questions:

(a) In the process of the Uygur learners' L3 acquisition of English tense and aspect, does the lexical aspect influence the use of past tense? In other words, does the LAH also apply to L3 acquisition of English with L1 Uygur and L2 Chinese?

(b) Is there any occurrence of language transfer in this L3 acquisition, either from L1 Uygur or L2 Chinese, or both?

5.2. Participants

The research subjects consisted of 25 participants who were undergraduate students at Southeast

[1.] The present study largely follows the study of Bardovi-Harlig and Reynolds (1995) which investigated the role of lexical aspect in the L2 acquisition of tense and aspect.

University, 13 of which were male and 12 were female. All of them acquired Uygur as their first language and Mandarin Chinese as their second language. However, they were exposed to the three languages differently before entering the university, especially Chinese and English. For Group One (N = 11), the participants received education from elementary school to high school in Xinjiang. They studied all subjects in Uygur and learned Chinese at school as their second language. They did not study English until they went to university. By the time when they participated in the study, they had studied English for two or three years. The participants of Group Two (N = 14) received education from elementary school to high school in Chinese, and they began to learn English at primary school. Overall, their Chinese and English are better than the participants of Group One.

All the participants had taken College English Test (CET) band 4 or 6. They were classified into three levels according to their scores of the test. Level 1 (the lowest level) included those who had taken only CET-4 and the scores were below 490; Level 2 (the medium level) included those who had only taken CET-4 with a score above 520 and those who had taken CET-6 with a score below 460; Level 3 (the high level) included those who had taken CET-6 with a score above 490. The distribution of participants by their English level and education background is presented in **Table 2**.

5.3. Materials

The materials used in the current study for collecting data from the participants included a questionnaire for basic information and a short test which tested the participants' knowledge of English tense and aspect. In addition, a short interview was conducted before the participant took the test, which was for the purpose of getting a general idea about the participants' language proficiency of Chinese and English.

(a) The questionnaire included two types of information of the participants. One type was the demographic information which included gender, date of birth, current grade, and their education background prior to college. The other type of information concerned their language proficiency in Chinese and English, including their test results of MHK (test of Mandarin Chinese proficiency for ethnic minority), the years of learning English, and English level (CET-4 or 6).

(b) The test included 34 short passages which contained 74 test items with 50 target items and 24 distractors that were verb forms not under investigation here. The passages varied in length from one sentence to three sentences. For each passage, the time reference for each eventuality was established through the use of time adverbials or verb tense and the two investigators double checked that there was no ambiguity for any test item. In this task the base form of each verb was given, and the participants were asked to fill in the blank with an appropriate tense form. The verbs were tested in various person such as the first, second and third person singular and plural. The 50 items includ-

Table 2. Distribution of participants by English level and education background

		Education background	
Level	**Number**	**Group 1**	**Group 2**
1	14	7	7
2	8	4	4
3	3	0	3
Total	25	11	14

ed 12 achievement verbs, 12 accomplishment verbs, 14 activity verbs (with one verb being tested twice), and 12 state verbs. Two sample tests are presented as follows:

(13) Last week, John was on his way to a singing lesson in Chicago and ____ (*stop*) at Starbucks for an iced coffee. He _____ (*notice*) a young woman looking at him. John _____ (*smile*) at her, out of politeness.

(14) There is a bank in a nearby town. You ____ (*need*) to take the bus to get there.

(c) More information about the Chinese and English language proficiency of the participants was obtained by a short interviews prior to the test. During the interviews, the participants were asked about their learning experience of English and Chinese as well as their metalinguistic knowledge of the tense and aspect systems in Chinese and English. This provides more information of the participants for the current study.

5.4. Data collection

The experiment was conducted in the spring academic semester and the data of the 25 participants were collected in three weeks. First, we invited the participants to a proper environment in which there are no other interruption so as to ensure the quality of the answers. The participants first took the short interview, and then completed the questionnaire for the demographic information and information of their language proficiency levels of Chinese and English. Afterwards, they took the cloze test. No time restrictions were imposed. For each participant, the entire experiment took about 15 to 25 minutes to complete.

6. Results

The participants' answers were classified into several categories according to the morphology of the verb form they provided: (a) past, which included the simple past tense forms and regularized past tense forms for irregular verbs such as *writed*; (b) non-past, which included simple present such as *likes*, and uninflected base forms such as (he) *like*; (c) progressive, which included cases of an inflected verb (V-*ing*) with no auxiliary, present progressive, and past progressive; (d) perfect, which included perfect forms of all tenses; and (e) other forms, which included all the other remaining supplied forms. The incorrect spelling of the response was overlooked; the response was classified only based on the ending morphology of the supplied verb form. For example, the incorrect spelling *syaed*, which was intended to be the past tense form of the verb *stay*, was classified as the past category. The ungrammatical uninflected verb for perfect aspect was also ignored such as *has snow* which should be *has snowed*. The perfect progressive form for activity verbs such as *had been swimming* was classified as progressive.

6.1. The use of simple past

The participants produced 1,850 responses in total to the cloze test, and 1,250 of them were responses for the target items. Among the responses to the target items, 920 responses were in the simple past tense, which suggests overall the participants had a quite high level of acquisition of past tense. The details of the simple past responses by lexical aspectual class and English language proficiency levels are presented in **Table 3**.

Table 3. The use of simple past by lexical aspectual class and English level

Level	Number of participants	States (12 items)	Activities (14 items)	Accomplishments (12 items)	Achievements (12 items)
1	14	90 (53.6%)	129 (65.8%)	144 (85.7%)	135 (80.4%)
2	8	68 (70.8%)	78 (69.6%)	80 (83.3%)	81 (84.4%)
3	3	25 (69.4%)	30 (71.4%)	30 (83.3%)	30 (83.3%)

Table 3 shows that the participants of all proficiency levels had acquired the English tense knowledge to certain degree. Even for the lowest level (Level 1), the participants were able to use the past tense for more than half of the target items of all aspectual classes. The table also shows that both the lexical aspectual property of verb predicate and the language proficiency level influence the use of simple past. First, the three levels of English proficiency exhibit a similar pattern: accomplishments and achievements have similar level of appropriate use of simple past (between 80.4% and 85.7%) and both are higher than the level of appropriate use of simple past for states and activities (between 53.5% and 71.4%). This suggests a preference to use past tense for telic predicates. Note that such preference was found for all language proficiency levels. The low level of English proficiency shows a high level of appropriate use of simple past for accomplishments (85.7%) and achievements (80.4%), which are apparently higher than states (53.6%) and activities (65.8%). Similarly, for the two higher levels, the appropriate use of simple past for accomplishments (L2: 83.3%; L3: 83.3%) and achievements (L2: 84.4%; L3: 83.3%) is apparently higher than states (L2: 70.8%; L3: 69.4%) and activities (L2: 69.6%; L3: 71.4%). This suggests that the LAH persists into the higher level of language proficiency. Second, the three levels of English proficiency show apparent differences from each other in their uses of past tense for state verbs. The low level (Level 1) produced appropriate use of past tense for only 53.6% of the states in the test whereas the higher level (Levels 2 and 3) did it for 70.8% and 69.4% of the states respectively. The medium level (Level 2) does not show apparent difference from the high level (Level 3) with respect to the use of past tense for the four lexical aspects. These suggest that the low level of language proficiency is more subject to the constraint of the lexical aspect for using past tense. In other words, the influence of the lexical aspect diminishes when the language proficiency increases.

6.2. Activity verbs

It is also worth taking a look at the participants' alternative responses to the simple past form as they also reveal that the lexical aspect plays a role in the acquisition of tense and aspect. In the case of activity verbs which were tested in contexts such as (15), the major alternative to the past tense form is progressive, as shown in **Table 4**. Note that in Bardovi-Harlig and Reynolds' (1995) study, the native speakers predominantly used the simple past tense form for the stative verb in such case, which showed the contrast between native speakers and non-native speakers in tense-aspect knowledge.

(15) Last week, James swam (swim) every day. Now he's getting bored of it.

Table 4 shows that there are three alternatives to the past tense for activity verbs that the participants of the current study might use, namely, nonpast, perfect form, or progressive with the progressive as the more common choice. This pattern holds for all of the three levels of language proficiency. As indicated by the percentages of progressive, the participants of the three levels of English

Table 4. The distribution of tense-aspect markers in activity verbs

Level	Past	Nonpast	Progressive	Perfect
1	129 (65.8%)	18 (9.2%)	37 (18.9%)	9 (4.6%)
2	78 (69.6%)	10 (8.9%)	16 (14.3%)	7 (6.3%)
3	30 (71.4%)	2 (4.8%)	8 (19.0%)	2 (4.8%)

Table 5. The distribution of various progressive forms in activity verbs

Level	θ prog.	Simple past prog.	Present prog.	Perfect prog.	
				Past	Present
1	23 (11.7%)	2 (1.0%)	0	1 (0.5%)	3 (1.5%)
2	11 (9.8%)	3 (2.7%)	0	0	2 (1.8%)
3	1 (2.4%)	3 (7.1%)	0	1 (2.4%)	0

proficiency almost have the same chances of using progressive (between 14.3% and 19.0%), though the number of L3 participants is too small to make any statistically significant conclusion.

The progressive responses by the participants show an improvement of using the correct grammatical form for progressive as the language proficiency increases. The responses consist of several forms, which show contrast between levels of language proficiency as shown in **Table 5**. In particular, the data show that as language proficiency increases, a predominant use of θ progressive forms such as θ *dancing* and θ *swimming* is gradually replaced by the use of the simple past progressive such as *was swimming*. For example, Level 1 participants produced θ progressive forms for 11.7% of their responses and 1.0% for the simple past progressive, whereas Level 3 participants produced 2.4% for the θ progressive forms and 7.1% for the simple past progressive.

According to the LAH, the progressive expresses such a meaning that the action progresses or continues at the time the speaker is referring to, which is compatible with the inherent lexical aspectual property of activity verbs, and thus learners may use the progressive for activity verbs rather than past tense.

6.3. State verbs

The alternative to the past tense for state verbs favors a different form from the one for activity verbs. An example of the discourse containing a state verb is presented in (16).

(16) John finally decided to let Mary leave because she <u>seemed</u> (seem) unhappy with him.

In this case, the responses of all language proficiency levels show a preference for nonpast (present tense) as an alternative to the past tense while the low uses of progressive and perfect are largely negligible, as shown in **Table 6**.

Table 6 also shows that as the participants' language level advances, the use of past tense for state verbs increases and the use of nonpast decreases, suggesting that the influence of the lexical aspect on the use of past tense weakens when the language proficiency increases.

The use of simple present tense is consistent with the durative feature of state verbs. Conceiva-

Table 6. The distribution of tense-aspect markers in state verbs

Level	Past	Nonpast	Progressive	Perfect
1	90 (53.6%)	69 (41.1%)	5 (3.0%)	4 (2.4%)
2	68 (70.8%)	25 (26.0%)	1 (1.0%)	2 (2.1%)
3	25 (69.4%)	7 (19.4%)	3 (8.3%)	1 (2.8%)

bly, a quality that is true in the past is normally true at the present. For example, if Mary was smart in the past, then it is often the case that she is smart now. Only the present tense has the meaning of "continued existence" (Andersen and Shirai, 1994) and thus we can expect that the present tense is a common alternative to past tense for state verbs.

It is also noticeable that as found in prior study (Bardovi-Harlig and Reynolds, 1995), the participants in the current study also rarely used the progressive for state verbs, although the meaning of progressive is compatible with the durative property of state verbs. This suggests that the learners of all language proficiency levels are aware of the restriction of progressive to dynamic verbs and thus do not overgeneralize it to the state verbs.

7. Discussion

The results presented above confirm that the lexical aspectual property of verbs plays an important role in the acquisition of tense and aspect. The participants of all language proficiency levels treat the telic predicates (achievements and accomplishments) as the best cases of past tense carriers but show a lower use of past tense for activity and state verbs. Thus, the data seem to suggest that L1 transfer is absent in the process of L3 acquisition of tense and aspect since the L1 Uygur is also a tensed language; instead, it is the universal acquisition constraint (the LAH) that exerts more influence on the L3 acquisition, which has also been attested for L1 and L2 acquisition.

The study found that in the process of L3 acquisition, the use of past tense is influenced by the lexical aspect of verbs even though the L1 also has grammatical tense. The performances of the participants of all English proficiency levels show a similar pattern of using past tense for the four verbal classes. They all showed a significantly higher use of past tense for the telic predicates (accomplishments and achievements) than for activity and state verbs, and for the alternatives to the past tense for activity and state verbs, they also showed the same pattern, that is, the progressive is the major tense-aspect marker alternative for activity verbs and the nonpast for state verbs. Such similarities suggest that the inherent lexical aspect is semantically related with the tense-aspect markers, and they are mainly used by learners to indicate the lexical aspectual meaning rather than to perform the tense-aspect function (the LAH). At an earlier stage of learning the tense and aspect system, the use of past tense is undergeneralized. It is not until the learners move to a higher level of language proficiency, when the past tense is used across lexical aspectual classes, the past tense is used for its real function.

The findings of the current study do not support The L1 Factor Hypothesis for L3 acquisition (Jin, 2009) which emphasizes the role of L1 in learning the grammar of L3. For the current study, the L1 Uygur, like English, is a tensed language and the verbal morphology inflects for the past tense. Thus, if L1 transfer occurred in the L3 acquisition of tense and aspect, we would expect that

the use of past tense should be highly and equally accurate for all lexical aspectual classes. But this prediction was not borne out. The use of past tense for activity and state verbs is not significantly high when compared to the use of past tense for accomplishment and achievement verbs. It does not seem to be the case that the learners transfer their Uygur tense knowledge in acquiring the English past tense.

It is hard to conclude that the results of the current study support the L2 Status Factor for L3 acquisition (Bardel and Falk, 2007). As shown in Section 4, the verbal suffix *-le* in Chinese, which is the L2 of the participants, has such a distribution that it is typically used for the accomplishments and achievements but never for the state verbs, and for the activity verbs, it is acceptable to use *-le* with them in more restricted contexts and thus a low frequency in the input. In other words, the distribution of *-le* is subject to the lexical aspect of verbs in a way that the LAH predicts for the use of past tense for languages like English. Moreover, semantically, when *-le* used with accomplishments, achievements, and activities, its default interpretation is past tense. Therefore, it seems that the use of past tense of the participants in the current study could be a result of L2 transfer. However, such an explanation cannot explain why the LAH is commonly attested for L2/L3 acquisition even when the L1/L2 has no counterpart of the Chinese *-le*.

The same argument holds for the Cumulative Enhancement Model (Flynn et al., 2004). According to this model, the use of past tense across lexical classes by the participants is a result of transferring linguistic knowledge of both L1 and L2. For example, the learners transferred the past tense knowledge from L1 Uygur and the lexical aspectual restriction on the use of *-le* from L2 Chinese. This explanation seems implausible as well since the LAH has been widely attested.

The Typological Primacy Model (Rothman, 2011, 2015) cannot explain the results of the current study either. According to this model, we might expect an L1 transfer occurs for the L3 acquisition of tense and aspect by the Uygur learners since both Uygur and English have grammatical tense and thus the learners may perceive them as typologically closer. However, like the L1 Factor Hypothesis, this cannot explain the influence of lexical aspect on the use of past tense. So, it seems that linguistic typology does not play any significant role in the L3 acquisition of tense and aspect.

The ubiquitous influence of the lexical aspectual class seems to suggest that the acquisition of tense and aspect, which undergoes a specific sequence, is something universal rather than a result of language transfer of prior linguistic knowledge. However, there might be some other explanation for it. For example, Bardovi-Harlig and Reynolds (1995) put forward a hypothesis from Andersen (1990) as an alternative explanation. Andersen hypothesized that if a particular form received a distributional bias (i.e., more common) in a certain linguistic environment against another form which occurs in the same environment, the learner might misperceive the meaning of the form. In the case of the use of simple past, the telic predicates (accomplishments and achievements) are more common than activity and state verbs, and have a completive reading, which might cause the learner to misperceive it as a perfective aspect marker. This explains why telic predicates are selected as the best carriers of simple past.

8. Conclusion

The current study investigated L3 acquisition of English tense and aspect by Uygur learners with

L2 Chinese. Specifically, it investigated whether the lexical aspectual class influences the use of simple past for verbs (the LAH), and whether the acquisition of tense and aspect shows an occurrence of language transfer, either from L1 or L2 or both. The results of the study show that the lexical aspect plays an important role in the use of past tense by the Uygur learners of all levels of English proficiency, thus it confirms the LAH. It is argued that there is no clear evidence for language transfer involved in this case of L3 acquisition. However, more evidence is needed for a definitive conclusion regarding the language transfer factor. For example, future studies may consider a comparison of the Uygur learners' English tense-aspect system (L3 Acquisition) with that of the Chinese speakers' (L2 Acquisition), which would provide more (counter)evidence for language transfer from L1 Uygur.

Conflict of interest

The authors declared no conflict of interest.

Author contribution

Both authors contributed equally to the paper and are co-first authors.

References

Andersen RW (1990) Models, processes, principles and strategies: Second language acquisition inside and outside the classroom. In: VanPatten B and Lee JF (eds.) *Second Language Acquisition-Foreign Language Learning*. Clevedon: Multilingual Matters, pp.45–78.

Andersen RW (1991) Developmental sequences: The emergence of aspect marking in second language acquisition. In: Huebner T and Ferguson CA (eds.) *Crosscurrents in Second Language Acquisition and Linguistic Theories*. Amsterdam: John Benjamins, pp.305–324.

Andersen RW and Shirai Y (1994) Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16: 133–156.

Bardel C and Falk Y (2007) The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research* 23: 459–484.

Bardovi-Harlig K and Reynolds DW (1995) The role of lexical aspect in the acquisition of tense and aspect. *TESOL Quarterly* 29(1): 107–131.

Comrie B (1976) *Aspect*. Cambridge: Cambridge University Press.

Eibensteiner L (2019) Transfer in L3 acquisition: How does L2 aspectual knowledge in English influence the acquisition of perfective and imperfective aspect in L3 Spanish among German-speaking learners? *Dutch Journal of Applied Linguistics* 8(1): 67–83.

Flynn S, Foley C, and Vinnitskaya I (2004) The cumulative enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition. *International Journal of Multilingualism* 1: 3–17.

Foote R (2009) Transfer in L3 acquisition: The role of typology. In: Leung YI (ed.) *Third Language Acquisition and Universal Grammar.* Clevedon: Multilingual Matters, pp.89–114.

Gu Q (2022) Telicization in Mandarin Chinese. *Journal of Linguistics* 1–33.

Jin F (2009) Third language acquisition of Norwegian objects: Interlanguage transfer or L1 influence? In: Leung YI (ed.) *Third Language Acquisition and Universal Grammar.* Clevedon: Multilingual Matters, pp.144–161.

Li CN and Thompson SA (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Robison RE (1990) Aspectual marking in English interlanguage. *Studies in Second Language Acquisition* 12: 315–330.

Robison RE (1995) The aspect hypothesis revisited: A cross-sectional study of tense and aspect marking in interlanguage. *Applied Linguistics* 16(3): 344–370.

Ross C (1995) Temporal and aspectual reference in Mandarin Chinese. *Journal of Chinese Linguistics* 23: 87–135.

Rothman J (2011) L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27: 107–127.

Rothman J (2015) Linguistic and cognitive motivation for the typological primacy model of third language (L3) transfer: Considering the role of timing of acquisition and proficiency in the previous languages. *Bilingualism: Language and Cognition* 18(2): 179–190.

Rothman J and Cabrelli Amaro J (2010) What variables condition syntactic transfer: A look at the L3 initial state. *Second Language Research* 26: 189–218.

Shirai Y and Andersen RW (1995) The acquisition of tense/aspect morphology: A prototype account. *Language* 71: 743–762.

Shirai Y and Kurono A (1998) The acquisition of Tense-Aspect marking in Japanese as a second language. *Language Learning* 48(2): 279–244.

Smith CS (1997) *The Parameter of Aspect* (2nd edn.). Dordrecht: Kluwer.

Vallerossa F (2021) The role of linguistic typology, target language proficiency and prototypes in learning aspectual contrasts in Italian as additional language. *Languages* 6: 184.

Vendler Z (1957) Verbs and times. *The Philosophical Review* 66(2): 143–160.

Salaberry R (2005) Evidence for transfer of knowledge of aspect from L2 Spanish to L3 Portuguese. In: Ayoun D and Salaberry R (eds.) *Tense and Aspect in Romance Languages: Theoretical and Applied Perspectives.* Philadelphia: John Benjamins, pp.179–210.

Sybesma R (2019) Finiteness and Chinese main clauses. *International Journal of Chinese Linguistics* 6(2): 325–344.

ORIGINAL ARTICLE

The development of Chinese university students' intercultural competence in a short-term study-abroad program

Hongshan Zuo[1*], Yaning Hu[2]

[1] College of Foreign Languages & Cultures, Sichuan University, Chengdu, Sichuan Province, China
[2] Changsha Dongya Middle School, Changsha, Hunan Province, China

Abstract: With increasing internationalization in higher education, more and more Chinese universities are providing students with exchange programs in cooperation with foreign universities. Hitherto, however, research about Chinese university students' intercultural competence in a study-abroad context has been relatively rare. The present study, adopting a mixed-method research design, examines the development of Chinese students' intercultural competence in a short-term study-abroad program, as well as the factors that might have a role to play in the process. It is found that the students made significant progress in intercultural competence after participating in the program and factors such as attitudes towards intercultural communication, cultural knowledge storage, English language proficiency and engagement in intercultural communication are important contributors to the development of intercultural competence in a short-term study-abroad context.

Keywords: study-abroad program; short-term; intercultural competence; development

***Corresponding author:** Hongshan Zuo, College of Foreign Languages & Cultures, Sichuan University, Chengdu, Sichuan Province, China; zuohongshan@scu.edu.cn

Received: May 3, 2022; **Accepted:** June 20, 2022; **Available online:** August 28, 2022

Citation: Zuo H and Hu Y (2022) The development of Chinese university students' intercultural competence in a short-term study-abroad program. *Forum for Linguistic Studies* 4(1): 132–146. DOI: 10.18063/fls.v4i1.1460

1. Introduction

In the current age of globalization, during which interactions among culturally diverse people are increasingly more common, the development of intercultural communicative competence (ICC) is of great importance in tertiary education. ICC is defined as a set of abilities to interact effectively and properly across cultures in various contexts (see Arasaratnam, 2015, 2016; Deardorff, 2009; Fantini, 2000). While the influences of classroom interventions on the development of ICC have been explored by researchers (e.g., Schuetze, 2008; Wang and Kulich, 2015), the impact of short-term study-abroad (STSA for short) experiences has been less well documented (Czerwionka et al., 2015). With more and more university students participating in STSA programs, research investigating ICC development during short-term programs abroad bears both academical and practical implications.

Drawing on Fantini's (2000) theoretical framework of ICC, which proposes that intercultural knowledge, skills, attitudes and awareness converge together as main components that contribute to ICC, the current research focuses on Chinese English-as-a-Foreign-Language (EFL) learners' development of ICC in a STSA program.

2. Defining and measuring intercultural competence

The concept of ICC has been given numerous definitions in the literature over the past several decades (Byram, 1997; Lustig and Koester, 2003; Deardorff, 2004; Fantini, 2000, 2006). Many researchers focused on the components of ICC while defining this concept. Pusch (1994), for example, proposes that the most important aspects of ICC include thoughtfulness, cognitive flexibility, ambiguity tolerance, behavioural flexibility, and intercultural identity. Byram's (1997) multi-dimension theory of ICC postulates that intercultural knowledge, skills, attitudes and awareness converge together as main components that contribute to ICC. In similar vein, Fantini (2000, 2006) holds that ICC is a complex of abilities with the four main dimensions of knowledge, attitudes, skills and awareness. He argues that awareness is the most important dimension of ICC and builds a graphical model in which attitudes, skills, and knowledge surround the core component of awareness. Deardorff (2004) notes that ICC is composed of three dimensions, including attitudes, skills, and knowledge. Among the three components, she emphasizes that attitudes are the most critical one.

Researchers generally agree that knowledge, attitudes, skills, and awareness are the main components of ICC (see, e.g., Lustig and Koester, 2003; Byram, 1997; Fantini, 2000, 2006; Deardorff, 2004, 2006; Spitzberg and Changnon, 2009; Peng et al., 2015). Based on this idea, scholars have designed various assessment tools to measure the extent to which an individual has the skills related to various dimensions of ICC, for example, "Assessment of Intercultural Competence" (AIC) (Fantini, 2000, 2006) and "Intercultural Development Inventory" (IDI) (Hammer et al., 2003).

In China, researchers have also developed a number of assessment tools to measure Chinese EFL learners' ICC (e.g., Wang, 1990; Liu, 2004; Wu, 2013; Peng et al., 2015). An early study by Wang (1990) designed a set of socio-cultural questions on "common culture" (e.g., social customs) and "formal culture" (e.g., geography, history, literature, and other subjects). A similar study by Liu (2004) attempted to incorporate culture test in language test. On the basis of large-scale empirical research, Liu concluded that the culture test in his study was a theoretically sound and practically feasible measurement system for assessing students' intercultural competence. In a more recent study, Peng et al. (2015) adapted Fantini's (2000, 2006) "Assessment of Intercultural Competence" and designed "Intercultural Competence Evaluation Scale for Chinese College Students". The scale enjoyed high validity and reliability (as measured by Cronbach's α, see Peng et al., 2015). In the present study, the scale was employed to measure the changes in the students' ICC before and after taking part in the STSA program.

3. Research on the development of ICC in a study-abroad context

Increasing globalisation has led more and more people to set out to study in foreign countries. Studying abroad is an experiential learning process which immerses the students in another culture (Kolb, 1984) and many researchers believe that studying abroad is an ideal way to improve one's

intercultural competence (Deardorff, 2006; Watson and Wolfel, 2015). In recent years, quite a few scholars have used empirical research approaches to explore the development of students' ICC in study-abroad contexts (e.g., Behrnd and Porzelt, 2012; Gregersen-Herman, 2015; Czerwionka et al., 2015). For example, Behrnd and Porzelt (2012) compared the development of intercultural competence of college students with and without study-abroad experiences. The results showed that while there is no significant difference in intercultural competence between the two groups at the beginning of the study, the improvement of intercultural competence of overseas students is more significant over time. Gregersen-Herman (2015) investigated the development of college students' cross-cultural ability and found that merely placing students in a multicultural environment without intervention or guidance cannot improve their cross-cultural ability. Czerwionka et al. (2015) measured the change of students' cross-cultural knowledge in a STSA program. They found that the students' attention to cross-cultural knowledge changed after studying abroad. Besides, the students' intercultural knowledge increased during their exchange abroad, with the most significant increase in cultural and historical knowledge.

Over the past decade, STSA programs of one to eight weeks have been gaining increasing popularity among Chinese university students. Compared with long-term study-abroad programs, STSA programs are less time-consuming and more economical, and are therefore favoured by many Chinese students. However, at present there is still controversy concerning the role of STSA programs in the development of students' ICC. While a number of researchers believe that STSA programs can broaden the students' international perspectives, cultivate global mindset, enhance cultural sensitivity, and empower the students to cope with cultural conflicts, some scholars question the influence of STSA programs on the development of the students' intercultural competence (Dwyer, 2004; Medina-Lopez-Portillo, 2004). Medina-Lopez-Portillo (2004), for example, studied the relationship between the length of study-abroad programs and the development of students' intercultural sensitivity. The study focused on 28 students who went abroad for exchange, 18 of whom participated in long-term exchange programs, and ten participated in short-term exchange programs. Results showed that the length of the project has a significant impact on the improvement of the students' cross-cultural sensitivity. However, Dwyer's (2004) research discovered that the advantages of long-term exchange programs over short-term exchange programs are not absolute. Dwyer noted that while having an entire academic year abroad has a more significant impact on students, "in some cases, students in short-term exchange programs may be more likely to derive more lasting benefits from their study abroad experience than students who participate in long-term study abroad programs" (Dwyer, 2004: 161).

In recent years, some Chinese scholars probed into the development of ICC of Chinese students in a study-abroad context (Lu and Li, 2012; Cui, 2013). Lu and Li (2012) used "Intercultural Sensitivity Inventory" to explore the differences in cross-cultural sensitivity of Chinese postgraduate students at home and abroad. The results showed that Chinese graduate students who studied abroad had a higher level of cross-cultural sensitivity than their counterparts in China, and they were more likely to accept and identify with different cultures and to enjoy diverse cultural interactions. Cui (2013) explored the influence of STSA experience on Chinese EFL students' cultural sensitivity and language proficiency by observing their cultural sensitivity, motivation, and relationship with the host family. The results showed that the overall level of language proficiency of the study-abroad group improved. While these studies in general showed a positive effect of studying abroad on the

development of Chinese ELF learners' intercultural sensitivity, more research is needed to examine the overall development of ICC in a study-abroad context.

4. Research methods

4.1. Research questions

This study set out to investigate the development of Chinese EFL students' intercultural competence in a STSA program. Specifically, two research questions have been formulated:

Research Question 1: What is the impact of a STSA program on the development of ICC of Chinese EFL students?

Research Question 2: What factors have a role to play in the students' development of ICC in the STSA program?

4.2. Research setting

The study took the ** program between a public research university in **southwest** China and a public research university in the **northeast** of the United States as a specific case to explore the impact of short-term exchange programs on ICC development of Chinese students. The program spanned two weeks, in which Chinese students studied in the American university and experienced American university teaching, academic lectures, campus life, as well as other cultural activities, around themes of international diplomacy, international business and global leadership. The American university appointed a highly international teaching team for the program, including an African-American program director and an Anglo-American supervisor, a Turkish-American lecturer, and three teaching assistants from the U.S., Japan, and China respectively. This, in addition to the study-abroad context, offered the students ample opportunities to interact with people from diverse cultural backgrounds. One of the authors of this paper also participated in the program as a student, which facilitated the data collection process. To ensure the reliability of the study, the data of this participant was not included in data analysis.

4.3. Participants of the study

Thirty-five members of the program participated in the study[1], the demographic information of the participants can be seen in **Table 1**. All of them were students of ** University, and their majors ranged from English, Literature, International Politics, to Law, Finance, Microelectronics, and Civil Engineering.

All the thirty-five participants took part in the online questionnaire survey. Twelve participants were invited to participate in in-depth semi-structured interviews concerning the strategies that they employed to develop ICC in the study-abroad context.

4.4. Research instruments

In order to fathom thoroughly the development of the students' ICC in STSA programs, this

[1] Originally there were 36 participants in the study-abroad program as one of the authors was also a member of the program as a student at the time of investigation. To ensure the reliability of research, the data of the author was taken out, leaving a total of 35 participants (see Table 1).

Table 1. Demographic information of the participants

	Categories	No. of participants
Gender	Female	27
	Male	8
Age	19–23 yrs	19
	24–29 yrs	16
Education background	Undergraduate	16
	Graduate	19

Table 2. Reliability test results of the self-assessment ICC questionnaire

Dimension	Number of items	Cronbach's alpha	Sample items
Knowledge of domestic culture	3	0.896	Knowledge of the domestic country's history, geography, and social politics
Knowledge of foreign culture	7	0.862	Knowledge of foreign history, geography and social politics
Attitude	4	0.877	Willing to communicate and study with foreigners from different cultures
Cross-cultural communication skills	9	0.911	Skills to negotiate and explain one's own culture to the satisfaction of the interlocutor when cross-cultural misunderstandings occur
Cross-cultural cognitive skills	3	0.848	Skills to acquire knowledge related to cross-cultural communication through direct contact with foreigners
Cross-cultural communication awareness	3	0.929	Awareness of cultural similarities and differences when communicating with foreigners
Total	29	0.954	

study employed a variety of ways to collect data, including questionnaire surveys, semi-structured interviews and field observations.

The questionnaire used in this study consisted of three parts. The first part was the participant's personal information, including gender, age, major, grade, previous study abroad experience, and English proficiency. The second part was the intercultural competence self-assessment scale, adapted from Peng et al.'s (2015) "Intercultural Competence Evaluation Scale for Chinese College Students", including six dimensions: knowledge of domestic culture, knowledge of foreign culture, attitude, cross-cultural communication skills, cross-cultural communication skills, and cross-cultural communication awareness, with a total of 29 items.

The researchers piloted the questionnaire with students who have taken similar exchange programs in ** University and calculated the reliability of the questionnaire. As can be seen (**Table 2**), the Cronbach's α coefficient of each dimension was between 0.848 and 0.929, and the Cronbach's α coefficient of the overall questionnaire was 0.954, indicating that the questionnaire enjoyed high reliability.

Following Peng et al. (2015: 149), the scale data were weighted in the following manner:

The weight of the first-layer index u_i, which is related to target layer U is:

$$A = (a_1, a_2, a_3, a_4, a_5, a6) = (0.05, 0.30, 0.19, 0.25, 0.06, 0.15)$$

The weight of the second-layer index u_{ij} under the ith first-layer index u_i is:

$$A_1 = (a_{11}, a_{12}, a_{13}) = (0.33, 0.27, 0.4),$$
$$A_2 = (a_{21}, a_{22}, a_{23}, a_{24}, a_{25}, a_{26}, a_{27}) = (0.18, 0.18, 0.20, 0.14, 0.12, 0.06, 0.12),$$
$$A_3 = (a_{31}, a_{32}, a_{33}) = (0.5, 0.17, 0.33),$$
$$A_4 = (a_{41}, a_{42}, a_{43}, a_{44}, a_{45}, a_{46}, a_{47}, a_{48}, a_{49}) = (0.09, 0.18, 0.15, 0.13, 0.09, 0.09, 0.09, 0.09, 0.09),$$
$$A_5 = (a_{51}, a_{52}, a_{53}) = (0.34, 0.25, 0.41),$$
$$A_6 = (a_{61}, a_{62}, a_{63}) = (0.40, 0.20, 0.40).$$

The third part is open-ended questions to understand the participants' perceptions of intercultural competence, self-assessment of their own intercultural competence, and their strategies for improving intercultural competence.

In addition to the questionnaire, we also conducted semi-structured interviews with 12 students randomly selected from the 35 participants. The semi-structured interviews focused on eight questions, such as, "How do you understand intercultural competence?" "How do you rate your intercultural competence?" "Has the STSA program helped you develop your intercultural skills?" "What methods have you adopted to improve your intercultural competence?". Interviews were conducted in Chinese and lasted approximately 40 minutes for each student.

4.5. Data collection and analysis

The students completed the "Intercultural Competence Evaluation Scale for Chinese College Students" before they left China and did it again upon their return to China. We performed statistical analysis of the questionnaire data using SPSS 24.0.

For the interview recordings, we first transcribed the interviews verbatim, then labelled and coded the transcriptions with the help of the qualitative analysis software MAXQDA. To increase the reliability of the qualitative analysis, the two authors of this paper jointly analysed the transcribed utterances.

5. Results and discussion

5.1. The development of the participants' ICC during the study-abroad program

To understand the changes of students' cross-cultural competence before and after participating in the STSA program, we calculated the scores of each dimension as well as the total score of the questionnaire for each participant according to the data weighting scheme mentioned above (Peng et al., 2015; Wu, 2013). Paired sample *t*-test was conducted to compare the scores before and after studying abroad. Data analysis showed that participants' intercultural competence changed significantly before and after their STSA experience, with remarkable improvements in all six dimensions of intercultural competence, as well as the overall intercultural competence assessment (**Table 3**).

Figure 1 shows the changes in each dimension of intercultural competence of university students before and after they participated in the STSA program.

Table 3. Participants' self-assessment of ICC before and after the STSA program

		Mean	SD	Mean difference	*t*	Sig. (2-tailed)
KD	Before	0.595	0.142	−0.080	−5.088	0.000
	After	0.676	0.107			
KF	Before	0.426	0.106	−0.201	−8.973	0.000
	After	0.627	0.100			
AT	Before	0.700	0.154	−0.069	−3.649	0.001
	After	0.769	0.105			
ComS	Before	0.600	0.131	−0.123	−8.693	0.000
	After	0.723	0.109			
CogS	Before	0.542	0.132	−0.127	−6.881	0.000
	After	0.700	0.139			
AW	Before	0.577	0.176	−0.170	−5.766	0.000
	After	0.747	0.119			

Note: KD = Knowledge of domestic culture; KF = Knowledge of foreign culture; AT = Attitude towards cross-cultural communication; ComS = Cross-cultural communication skills; CogS = Cross-cultural cognitive skills; AW = Cross-cultural communication awareness.

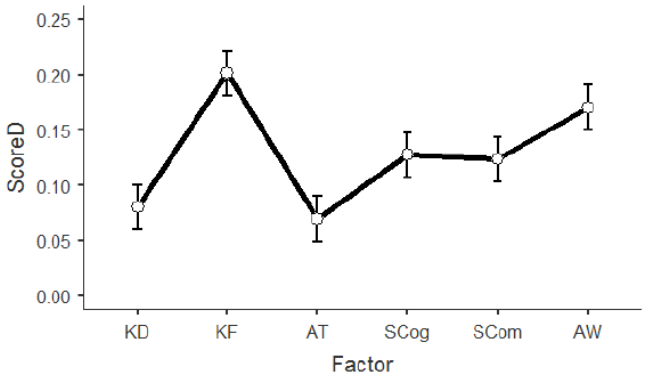


Figure 1. Numerical gains in the participants' ICC scores along the six dimensions.

As shown in Table 3, the students' scores in attitudes towards intercultural competence (AT for short) are at a relatively high level both before and after the STSA program, showing that students had positive attitudes towards cross-cultural communication. In other words, they were willing to communicate and learn with foreigners from different cultures, to tolerate foreigners' different values, living habits and taboos, and are eager to learn foreign languages and understand foreigners. On the contrary, although the "Knowledge of foreign culture" (KF for short) was dramatically improved after studying abroad, it still ranked the lowest among the six dimensions. A possible reason might be that although STSA program helped to enhance the students' knowledge of foreign history, geography, society and politics, the students were unable to acquire systematic and sufficient expertise in such a short time. More extended programs and more elaborate learning are needed to enrich students' knowledge of foreign cultures.

It is noteworthy that in STSA programs, students were immersed in a foreign language and cultural environment, which undoubtedly contributed to their acquisition of foreign cultural knowledge. Surprisingly, as can be seen from **Table 4**, the students' knowledge of domestic culture (KD for short) also experienced a noticeable increase after the STSA program. Interviews showed that many

Table 4. Participants' self-assessment of overall ICC before and after the STSA program (*t*-test)

		Mean	SD	Mean difference	*t*	Sig. (2-tailed)
ICC	Before	0.559	0.108	−0.142	−9.437	0.000
	After	0.701	0.086			

students believed that knowledge of their own culture was a prerequisite for successful cross-cultural communication, and only with this premise could they communicate with people from different cultural backgrounds calmly and confidently. As student F1[2] stated in her interview:

> "Intercultural competence, first of all, includes the understanding of the politics and history of one's own country. Then there is one's view of life and values, some local customs and manners of one's own country, and such a knowledge of one's own culture. It is through the understanding of one's own culture and the mastery of some local knowledge and skills, that we build up cultural confidence. And then when it comes to intercultural communication, we can be neither cringing nor arrogant, and communicate with the other side confidently, even promote our culture." (F1)

Through cross-cultural communication, students gradually realised the importance of understanding their own culture, and which fuelled them to expand their relevant knowledge and deepened their knowledge of Chinese culture. As Jiang (2012) proposed, the more profound the mastery of one's native culture, the easier the understanding of western culture, because culture is both national and global, and there are commonalities and universal laws in different cultures. Therefore, understanding foreign cultures also deepens our understanding of the cultural differences between China and foreign countries, which in turn deepens our understanding of our own culture.

5.2. Factors affecting the development of ICC during STSA programs

Through analysis of the interview data, we identified five main factors that influenced Chinese students' development of intercultural competence in STSA programs.

5.2.1. Attitudes towards intercultural communication

Attitude is an essential component of intercultural competence (Allport, 1954; Orlandi, 1992; Byram, 1997). According to Allport (1954), attitudes are prerequisites for successful intercultural interaction, and people need to be curious, open-minded, and ready to learn about and explore the meanings, beliefs, and behaviours of others. In similar vein, Orlandi (1992) also proposed that among the five components of intercultural competence, i.e., skills, understanding, appreciation, willingness, and ability, willingness was the most important one. Similarly, Byram (1997) stated that part of the success of such interaction will depend on the establishment and maintenance of human relationships, something which depends on attitudinal factors.

Our analysis of the interview data borne out the above views. In response to the question, "What qualities did you observe in the students who were interculturally competent in this exchange program?" All participants believed that students with strong cross-cultural abilities have a higher willingness to contact people from different cultures and actively seek opportunities for cross-cultural communication. As the following two students said in the interview:

[2] This is a pseudonym given to the participant. It is the same with the other participants appearing in the rest of the paper.

> "The students who behaved well (in intercultural communication) were often curious about new cultures that they were not very familiar with. I think it's more… let's say, they want to know more about another culture, another place, another city, another country, 'what would it be like?' They have this kind of curiosity." (M3)

> "They are more willing to step out of their comfort zone and communicate and make friends with people from other countries, compared with those who prefer to communicate with their Chinese classmates." (M2)

According to the interview data, the language difference between Chinese and English did not seem to be a problem for these students who actively participated in cross-cultural communication, as they did not perceive it as a barrier to their participation in cross-cultural communication. They were able to listen actively to gain new knowledge and express themselves confidently in English during intercultural communication.

> "I think it depends on the mindset of the person you are communicating with interculturally. If you are shy about your language, and you feel that it is one of your weaknesses, then this can be a hindrance. But if I accept the fact that my English is not good and I can still communicate bravely, then language has no significant impact on cross-cultural communication." (M2)

In cross-cultural communication, the students not only got a sense of accomplishment in learning new knowledge and improving their abilities but also encountered frequent setbacks and difficulties, resulting in a sense of frustration and loss. Therefore, confusion, conflict and frustration were a constant part of intercultural communication, and students need to adjust their mindsets and deal with these challenges actively.

> "I think there's a lot of joy in intercultural communication, but the frustration it can bring is also deep... It is inevitable to encounter such things (frustration in the process of intercultural communication). I think the ability to resist setbacks in cross-cultural communication must be high." (F2)

As Deardorff (2006) stated, the key to the development of intercultural competence is personal attitudes, such as respect and curiosity about different cultures and values. These attitudes stimulate self-consciousness of culture, a deep understanding of other cultures, and the ability and willingness to behave accordingly. Some scholars pointed out that many Chinese students learned English from an early age, and have been exposed to Western culture both inside and outside the classroom since then, and they are curious and open-minded to Western language and culture (Peng et al., 2015). This study also finds that although the language proficiency and intercultural competence of the participants varied, all of them showed positive attitudes towards cross-cultural communication, which was beneficial to the improvement of intercultural competence.

5.2.2. Cultural knowledge

According to Rosen et al. (2000), the key elements of cross-cultural communication are: valuing one's own culture, understanding the culture of others, using cultural knowledge to strengthen one's own culture, and using culture to create advantages for oneself. A good intercultural interlocutor should have competent cultural knowledge, such as knowledge of domestic culture, knowledge of foreign culture, knowledge of the culture of the destination to bring about successful intercultural communication. Wiseman (2001) addressed the knowledge component a critical indicator of one's

intercultural competence. He highlighted that successful intercultural communication requires information about the people, the context, and communication rules governing the interaction with the members of another culture. Wiseman's research shows that competent intercultural communicators are willing to spend time listening to and learning from different cultures. They know about cultural matters and are good at dealing with intercultural issues. Our analysis of the interview data confirms these views. For example, F1 stated in the interview that:

> "I think the people around me with good cross-cultural communication skills have a common trait, that is, they can integrate Chinese and western cultures. They can talk about the similarities and differences between Chinese and western cultures based on understanding, or even their origin to explain why such similarities and differences would appear, and then treat these differences rationally. Then I think the deeper reason is that they can integrate Chinese and Western cultures. Because they have accumulated a large amount of reading, they have formed an extensive knowledge. Then these students can build up a complete knowledge framework in their mind so that they can integrate these different cultures. And then I think they also have the common trait of always being tolerant and open-minded to different cultures, and will take the initiative to understand each other's culture in the process, probably starting from a minimal level." (F1)

Intercultural knowledge increases intercultural communicators' understanding of others and self to facilitate making accurate predictions and attributions in intercultural communication (Wiseman et al., 1989). The interview data of this study revealed that students considered intercultural knowledge to be an essential element of intercultural competence and were willing to spend time learning about domestic and foreign cultures. For example, student M1 gave the following example in the interview.

> "I had an undergraduate classmate who went to Canada when he was in high school. During his communication with Westerners, he gradually realised that he needed to have a deeper understanding of his own culture. Then he went to learn how to make tea. Right. Then when he returns home every summer vacation, he goes to the mountains to learn how to pick tea, how to make tea and how to taste tea, that is, to study this series of tea culture knowledge. Then, when he was an undergraduate, he had some local Canadian students, and he would take the initiative to exchange tea culture with these students and other international students, and then promote Chinese tea knowledge. Then, in this way, many foreigners would know some of our cultures. In this way, on the one hand, he has been affirmed by others; on the other hand, he also has a considerable sense of trust or confidence in himself. In this way, when he is communicating, he will benefit from this knowledge a lot." (M1)

To conclude, a distinctive feature of higher-intercultural-competent students was a more affluent knowledge base of both domestic and foreign cultures, which enabled them to promote their own culture confidently, to view cultural differences rationally, to digest and integrate different cultures.

5.2.3. English proficiency

Successful communication undoubtedly requires communicative skills, which includes linguistic, sociolinguistic and pragmatic abilities. However, researchers are not unanimous as to whether foreign language proficiency is a necessity for successful intercultural communication. Studies such as Carrell (1984), Koike (1996), Olson and Kroeger (2001), and Hismanoglu (2011) argue that stu-

dents with higher language proficiency are more competent in intercultural communication than students with lower language proficiency. However, the results of Watson and Wolfel (2015) do not align with this view. They conducted a two-year language training program for over two hundred students and examined their intercultural competence before and after participating in a STSA program. Surprisingly, the results of the study failed to obtain significant correlation between language proficiency and the development of intercultural competence.

The interview data in this study partially supports the findings of Watson and Wolfel (2015). When asked how much impact English proficiency had on intercultural communication, most students believed that English proficiency did not hinder intercultural communication if the conversation did not involve expertise. As M1 stated:

> "In similar contexts, or if both sides have related majors and are interested in similar subjects, or if the topic is frequently discussed, then there is no problem in communicating with each other. And if these conditions are not met, then a high level of language proficiency alone will not help intercultural communication." (M1)

While the students generally agreed that language proficiency had little impact on everyday cross-cultural communication, the role of language came to the fore if cross-cultural communication involved more specialised issues. Some students suggested that:

> "In a way, if we are not engaged in profound cross-cultural communication, which may involve a particular area of expertise, our English proficiency is not very important. But suppose we are trying to explore cross-cultural communication deeply, with topics including politics, economics, international relations, and so on, the accumulation of professional vocabulary and the way we express ourselves will make our English proficiency very important... and very important indeed." (F6)

Although the students had different views on the impact of foreign language proficiency on cross-cultural communication, they all agreed that low language proficiency did have a negative impact on their confidence and attitude to participate in cross-cultural communication. The higher the level of language, the more confident the students were in cross-cultural communication and the higher their participation in cross-cultural interaction. On the contrary, lower language proficiency would reduce their enthusiasm to participate in cross-cultural communication and cause them to miss some opportunities for cross-cultural communication, which would consequently result in limited gains in cross-cultural communication competence. In the interview, a student reported that:

> "I can't understand what the professor said for most of the time, but I ask the person next to me the meaning of some words that confused me, it will interrupt the learning process for both of us. There is tremendous nervousness and begin to doubt myself, as I didn't feel that my English is so poor when I was in China, and how could it be so different when I was in America? Then I am afraid to communicate with the teacher, and consequently, there is a lot of information I will miss." (F2)

Although the students held that language did not pose a threat to cross-cultural communication in general, in some cases, inadequate language proficiency did cause problems in cross-cultural communication, as shown in the following example described by the student:

> "When the first time I entered this kind of English working environment, I had many difficulties in adapting to it. For example, our teaching secretary sent us an email with a request to reply, but he used a word we don't often use in everyday English. So, none of us noticed this request, and nobody sent him a message back. The next day when he met us, he was furious about it and doubted our working attitude. Since then, I realised that English does have a vital role to play in intercultural communication. You couldn't tell which side should be blamed for the failure of cross-cultural communication, but this kind of thing often happens. That's where poor language proficiency can lead to." (M4)

As can be seen, the English proficiency of these university students was sufficient for everyday intercultural communication activities. Therefore, they believed that language level had little effect on cross-cultural communication. However, if intercultural communication involved in-depth discussion of professional or academic issues, a good mastery of the English language turned out to be an essential prerequisite. Only with a high level of English proficiency can one navigate through highly specialised intercultural communication and achieve satisfactory communicative effects.

5.2.4. Engagement in intercultural communication

The study found that the students' differential degrees of engagement in intercultural communication resulted in uneven gains in the development of intercultural development. Fredrick et al. (2004) propose that students' engagement in academic settings includes behavioral engagement, emotional engagement and cognitive engagement. This section mainly discusses the students' behavioral engagement since we found this element especially relevant in intercultural communication contexts.

Behavioral engagement refers to students' involvement/participation in academic and social or extracurricular activities and is crucial for achieving positive academic outcomes (Fredrick et al., 2004). The students' active and purposeful participation in intercultural communication means better opportunities to learn and use English because they felt like they were valued and contributive, which was conducive to the development of their intercultural competence. In our study, we found that learners' personality factors had a significant impact on the development of their intercultural competence. Although they joined in the same program, introverted students always had far fewer opportunities for cross-cultural communication than extroverted students because this personality hindered their attempts to engage in conversations with people who had different values, beliefs, and behaviours, as M3 stated:

> "I think I am a little bit introverted, especially when I meet some strangers for the first time, I may not dare to talk to them or whatever, or I will act a little bit reserved." (M3)

Previous research has shown that personality differences can lead to different levels of progress in intercultural competence. Some personality traits, such as optimism, openness, and extroversion, may be associated with higher levels of intercultural competence (Caligiuri, 2000). The present study found that, unlike extroverted students, introverted students were not proactive in intercultural communication, which resulted in a loss of opportunities to interact with locals, thus affected their development of intercultural competence to some extent.

6. Conclusion

This study examined the impact of a STSA program on the development of intercultural communication competence of Chinese students and found significant improvement in students' intercultural competence before and after the STSA experience. Among the six aspects of ICC the most significant change appeared on the improvement of students' knowledge of the foreign culture. In contrast, students' attitudes to intercultural competence did not show noticeable improvement as expected. In the qualitative examination, the study tried to explore the possible reasons underpinning the development.

Examining how STSA programs affect students' intercultural competence bears significant implications for both students and program designers. For students, the results of the study may help them better understand the benefits of short-term exchange programs in developing intercultural competence and the challenges they might encounter when interacting with people from different countries and cultural backgrounds in the study-abroad context. Besides, the study also examined the typical characteristics of the participants who performed well in intercultural communication. This can provide role models for students to follow when they set out to study in foreign countries. For study-abroad program designers, they could offer students courses about intercultural training before the students embark on the study-abroad journey, aiming at enriching their cultural knowledge, building positive attitudes towards intercultural communication, as well as mastering some essential intercultural communication skills.

Although this study used both questionnaire and interview data to examine the development in Chinese students' intercultural competence in a STSA program, it has a number of limitations. First, further consideration needs to refine the assessment on students' intercultural competence because there are potential defects in the form of self-report in indirect evaluation. Doubt about the ability of individuals to demonstrate valid and accurate self-evaluation is the reason why some researchers question whether the data results of indirect evaluation tools can be trusted (Arasaratnam and Doerfel, 2005). Besides, this research only examined the ICC development of 35 students in one study-abroad program. Larger samples from more programs are necessary to validate the findings of this study.

Conflict of interest

No potential conflict of interest was reported by the authors.

Acknowledgements

This research is supported by the Project on the Integration of Production, Education and Research, and sponsored by the Ministry of Education, P.R. China (Project number: 202101034009).

References

Allport GW (1954) *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Arasaratnam LA and Doerfel ML (2005) Intercultural communication competence: Identifying key components from multicultural perspectives. *International Journal of Intercultural Relations* 29: 137–163.

Arasaratnam LA (2015) Research in intercultural communication: Reviewing the past decade. *Journal of International & Intercultural Communication* 8(4): 290–310.

Arasaratnam LA (2016) Intercultural competence. In: Nussbaum J (ed.) *Oxford Research Encyclopedia of Communication*. Oxford: Oxford University Press, pp.1–23.

Behrnd V and Porzelt S (2012) Intercultural competence and training outcomes of students with experiences abroad. *International Journal of Intercultural Relations* 36(2): 213–223.

Byram M (1997) *Teaching and Assessing Intercultural Communicative Competence*. Clevedon, UK: Multilingual.

Caligiuri PM (2000) The big five personality characteristics as predictors of expatriate's desire to terminate the assignment and supervisor-rated performance. *Personnel Psychology* 53(1): 67–88.

Carrell PL (1984) Inferencing in ESL: Presuppositions and implications of active and implicative predicates. *Language Learning* 34: 1–21.

Cui D (2013) A study on the changes of English proficiencies acquired by Chinese students' short-term learning experiences in Canada. PhD Thesis, Shanghai International Studies University, China.

Czerwionka L, Artamonova T, and Barbosa M (2015) Intercultural knowledge development: Evidence from student interviews during short-term study abroad. *International Journal of Intercultural Relations* 49: 80–99.

Deardorff DK (2004) *The Identification and Assessment of Intercultural Competence as a Student Outcome of International Education at Institutions of Higher Education in the United States*. Raleigh: North Carolina State University.

Deardorff DK (2006) Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education* 10(3): 241–266.

Deardorff D (2009) Implementing intercultural competence assessment. In: Deardorff E (ed.) *The SAGE Handbook of Intercultural Competence*. Thousand Oaks: Sage, pp.477–491.

Dwyer MM (2004) More is better: The impact of study abroad program duration. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10: 151–164.

Fantini AE (2000) A central concern: Developing intercultural competence. *SIT Occasional Papers Series*, 1: 25–42. Brattleboro: School for International Training.

Fantini AE (2006) *Assessment Tools of Intercultural Competence* [online]. Brattleboro, VT: School for International Training. Available at: http://www.sit.edu/publications/docs/feil_appendix.pdf (Accessed: 10 June 2019).

Fredricks JA, Blumenfeld PC, and Paris AH (2004) School engagement: Potential of the concept, state of the evidence. *Review of Educational Research* 74(1): 59–109.

Gregersen-Herman J (2015) The impact of exposure to diversity in the international university environment and the development of intercultural competence in students. In: *The European Higher Education Area*. Cham: Springer, pp.73–92.

Hammer MR, Bennett MJ, and Wiseman R (2003) Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations* 27(4): 421–443.

Hismanoglu M (2011) An investigation of ELT students' intercultural communicative competence in relation to linguistic proficiency, overseas experience and formal instruction. *International Journal of Intercultural Relations* 35(6): 805–817.

Jiang M (2012) *Xichao yu Xinchao* [*Tides from the West & the Renaissance*]. Beijing: People's Publishing House.

Kolb DA (1984) *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs: Prentice-Hall.

Koike DA (1996) Transfer of pragmatic competence and suggestions in Spanish foreign language learning. In: Gass SM and Neu J (eds.) *Speech Acts across Cultures: Challenges to Communication in a Second Language*. New York: Mouton de Gruyter, pp.257–281.

Liu B (2004) An interface study of intercultural communicative competence and language testing. Unpublished Thesis, Shanghai International Studies University, China.

Lu X and Li X (2012) Comparison of the intercultural sensitivity of Chinese postgraduates at home and aboard. *Academic Exploration* 9: 105–107.

Lustig MW and Koester J (2003) *Intercultural Competence: Interpersonal Communication across Cultures* (4th edn.). Boston: Allyn and Bacon.

Medina-Lopez-Portillo A (2004) Intercultural learning assessment: The link between program duration and the development of intercultural sensitivity. *Frontiers: The Interdisciplinary Journal of Study Abroad* 10: 179–200.

Olson CL and Kroeger KR (2001) Global competency and intercultural sensitivity. *Journal of Studies in International Education* 5: 116–137.

Orlandi M (ed.) (1992) *Cultural Competence for Evaluators Working with Ethnic Minority Communities: A Guide for Alcohol and Other Drug Abuse Prevention Practitioners*. Rockville: Office for Substance Abuse Prevention, Cultural Competence Series 1.

Peng R, Wu W, and Fan W (2015) A comprehensive evaluation of Chinese college students' intercultural competence. *International Journal of Intercultural Relations* 47: 143–157.

Pusch M (1994) The chameleon capacity. In: Lambert RD (ed.) *Educational Exchange and Global Competence.* New York: Council on International Educational Exchange, pp.205–210.

Rosen R, Digh P, Singer M, et al. (2000) *Global Literacies: Lessons on Business Leadership and National Cultures*. New York: Simon and Schuster.

Schuetze U (2008) Exchanging second language messages online: Developing an intercultural communicative competence? *Foreign Language Annals* 41(4): 660–673.

Spitzberg BH and Changnon G (2009) Conceptualizing intercultural competence. In: Deardorff D (ed.) *The Sage Handbook of Intercultural Competence*. Thousand Oaks: Sage, pp.2–52.

Wang YA and Kulich SJ (2015) Does context count? Developing and assessing intercultural competence through an interview-and model-based domestic course design in China. *International Journal of Intercultural Relations* 48: 38–57.

Wang Z (1990) An analysis of data from a socio-cultural test in China. *Foreign Language Teaching and Research* 4: 32–36.

Watson JR and Wolfel R (2015) The intersection of language and culture in study abroad: Assessment and analysis of study abroad outcomes. *Frontiers: The Interdisciplinary Journal of Study Abroad* 25: 57–72.

Wiseman RL (2001) *Intercultural Communication Competence* [online]. Available at: http://commfaculty.fullerton.edu/rwiseman/ICCCpaper.htm.

Wiseman RL, Hammer MR, and Nishida H (1989) Predictors of intercultural communication competence. *International Journal of Intercultural Relations* 13: 349–370.

Wu W (2013) A comprehensive evaluation on Chinese college students' intercultural competence. PhD Thesis, Huazhong University of Science & Technology, China.

ORIGINAL ARTICLE

The effect of content retelling on incidental vocabulary acquisition for Chinese EFL learners

Jingyi Jin[1], Shuguang Li[2*]

[1] School of Teacher Education, Nanjing Normal University, Nanjing 210023, Jiangsu Province, China
[2] School of Foreign Languages and Cultures, Nanjing Normal University, Nanjing 210023, Jiangsu Province, China

Abstract: As the foundation of English learning, the acquisition of vocabulary has always been a hot topic in the fields of second language acquisition and foreign language teaching. Contrasted with intentional vocabulary acquisition, incidental vocabulary acquisition (IVA) relates to lexical gains as a by-product of main cognitive activities. In the field of teaching English as a foreign language (TEFL) in Chinese high schools, the incidental acquisition of English vocabulary has increasingly attracted the academic attention in recent years. However, few empirical studies have focused on the incidental acquisition of English vocabulary engendered by doing content retelling tasks. In light of the inadequacy, this study adopted the quality audio-visual material as input and content retelling as output (forming an input-output-input circle), aiming at exploring the effect of retelling on Chinese high school EFL learners' IVA. Results indicated that learners who retold the content of the audio-visual material between two viewings can pick up more words. In this process, the attempted use of new words in oral reproduction plays a positive role in strengthening the immediate acquisition and long-term retention of the target words.

Keywords: incidental vocabulary acquisition; audio-visual material; content retelling; pick up and retention

***Corresponding author:** Shuguang Li, School of Foreign Languages and Cultures, Nanjing Normal University, Nanjing 210023, Jiangsu Province, China; lishuguang@njnu.edu.cn

Received: May 6, 2022; **Accepted:** June 22, 2022; **Available online:** August 28, 2022

Citation: Jin J and Li S (2022) The effect of content retelling on incidental vocabulary acquisition for Chinese EFL learners. *Forum for Linguistic Studies* 4(1): 147–162. DOI: 10.18063/fls.v4i1.1461

1. Introduction

As a fundamental component in communication and a basic element of language learning, vocabulary is considered as a cornerstone of language proficiency. In order for a foreign or second language learner to reach the intermediate or even advanced levels of proficiency, he/she need to acquire "many thousands of words" (Hulstijn and Laufer, 2001: 540). However, the cruel fact is, in class, the time for teaching vocabulary is limited. Only a small portion of class time can be allocated specifically for deliberate and direct vocabulary teaching and learning. Fortunately, learner's mental

dictionary can also be enlarged incidentally. As noted by Huckin and Coady (1999: 190), "Incidental acquisition is the primary means by which second language learners develop their vocabulary beyond the first few thousand most-common words."

With the rapid development of science and technology, education informatization has entered a new historical stage, offering EFL learners better opportunities to further develop their academic language skills. Owing to technological developments, vocabulary learning can be more interactive inside and outside of the classroom. Audio-visual input materials, which appear to have a huge advantage in information selection and arrangement, have recently gained more attention in the field of second language acquisition. In digital era, they can act as up-to-date forms to expand the ways students learn so as to meet their diverse needs. Holding the view that input alone is not sufficient in vocabulary learning, this study adopted content retelling task as output, considering its benefit of offering learners the chance to pick up words in the context and retrieve in the context created by themselves. Different from previous research, which usually followed a simple input-output sequence, this study forms an input-output-input circle. Under such circumstances, learners own the opportunity to watch the audio-visual material again and may compare their expression of the target words with the original use.

2. Literature review

2.1. Incidental vocabulary acquisition and retention

According to Nagy et al. (1985), who first proposed the term "incidental vocabulary acquisition" when studying L1 language acquisition, children acquire several thousands of words per year of their mother tone through learning in context. Instructions dealing with words, nevertheless, are impossible to "cover that much ground" (Nagy et al., 1985: 251). Since the proposal of IVA, numerous researchers have paid attention to it and conducted empirical studies to examine its role. Joe (1998: 357) noted that incidental vocabulary learning, as an effective way to learn word meanings in the context, happens without the specific intent on vocabulary. Richards and Schmidt (2010: 276) defined incidental learning as the process of learning something without specific intention. Contrasted with intentional learning, it also means learning one thing while intending to learn another. To sum up, incidental vocabulary acquisition relates to lexical gains as a by-product of the main cognitive activities where learners are not explicitly attached to vocabulary learning.

Since vocabulary acquisition (or vocabulary uptake) occurs in a comparatively short period, vocabulary retention refers to the process in which knowledge is preserved in the long-term memory, and can be located, identified and retrieved accurately in the future (Sousa, 2006). Craik and Lockhart (1972: 675) believed that retention is a function of depth. And the depth to which it is processed is determined by various factors, such as the amount of attention on stimulus, its compatibility with the analyzing structures, and the length of available time to process. Compared with vocabulary uptake, vocabulary retention includes more complex progress such as memorizing, recalling, and recognizing.

2.2. Vocabulary gains from audio-visual input

Researchers around the world have studied the relationship between textual input and learners' incidental vocabulary acquisition from different perspectives. Most studies examined learners' IVA

during reading (Douglas, 2016; Horst et al., 1998; Hwang and Nation, 1989; Min, 2008; Suk, 2017). But Laufer (2003: 275) argued that in instructed foreign language context, the content-focused reading alone is unlikely to be the best source of vocabulary acquisition. Reading alone has some unavoidable deficiencies, and the most obvious one is its slow proficiency. To deal with it, some researchers turn to listening input, but most of them admit that vocabulary gains are quite small under such circumstances (van Zeeland and Schmitt, 2013; Vidal, 2003, 2011). The data in Brown et al.'s (2008) study even demonstrate that none of the target items' meaning is retained by learners in the listening-only mode three months after encountering the input.

To compensate for the disadvantages of reading and listening input, the use of audio-visual materials has been suggested and won more and more attention. Neuman and Koskinen (1992), in a review of research on television as comprehensible input, confirmed the general benefits of multimedia on language learning. Multimedia expose viewers to both picture and sound messages, in which case, learners can form the word-meaning connection of new words more easily. Owing to its' entertaining nature, television is more interesting and attractive, creating a meaningful and appealing context. Two important studies conducted by Secules et al. (1992) and Weyers (1999) have also confirmed the positive role of audio-visual materials, which turned out to facilitate students' listening comprehension and help them be more confident and competent in communication. Both of the two studies support the pedagogical value of authentic video in bringing about language development.

In a nutshell, audio-visual input can more effectively influence the learners on account of the interplay between different representational modes of the target language. In the audio-visual context, language learners get involved in a more natural way. Besides, multimodality input fosters learners' motivation and enthusiasm through stimulating both visual and auditory sensory systems.

In an era of information explosion, students can have access to different types of resources, and are able to learn English in dynamic environments. The input materials of vocabulary acquisition are no longer limited to reading or listening alone. A tremendous amount of authentic audio-visual materials is available on the Internet, which can cater to different learning needs.

2.3. The role of output tasks

As indicated by Swain (1985, 1995), there is no automatic conversion from explicit to implicit knowledge, and learners' understanding ability and output ability are not absolutely balanced. In order to be fluent and accurate in foreign language, language learners need not only comprehensible input, but also comprehensible output. And the latter is of utter importance.

The important role of output activities has been emphasized by many scholars. Chinses scholar Wen (2008) put forward the "output-driven hypothesis" for the teaching of English-skill courses to English majors in China. According to her, output-driven tasks shorten the distance between input and output, and accelerate the process of transformation from declarative knowledge to procedural knowledge. For learners with intermediate level proficiency or above, this hypothesis conjectures that output is more powerful than input. Output-driven hypothesis as prototype, the production-oriented approach (POA) has been elaborated as a system through nearly ten years' research (Wen, 2015). Output conquers a predominant status in this system, because POA starts teaching with language production and ends with production. As for input, it functions as an enabler to help learners accomplish productive activities.

Defining output both as product and process, Anthony (2008) believed that output creates more opportunities for learners to deal with the new words, and "push" them in using them. Since exposure to input alone is not enough for non-native learners, a new concept called "intentionally targeting language output" is mentioned, which requires teachers to "scaffold students to produce precise, coherent, and appropriate language" (Anthony, 2008: 481).

To encourage learner's further engagement with target words, various output tasks have been recommended and examined. Some typical and most frequently used ones include: 1) matching the target words with their definitions (Hu and Nassaji, 2016; Laufer and Rozovski-Roitblat, 2011; Min, 2008; Paribakht and Wesche,1997; Peters, 2012); 2) providing the L1 translation of the target words, or translate L1 words into the English vocabulary from the reading text (Laufer and Rozovski-Roitblat, 2011; Min, 2008; Peters, 2012); 3) selecting words to fill gaps in the text, sentences, or text summary (Hulstijn and Laufer, 2001; Lu, 2013; Yang et al., 2017; Zou, 2017); 4) composition writing incorporating stipulated words (Eckerth and Tavakoli, 2012; Kim, 2008; Lu, 2013; Zou, 2017).

Although the output tasks mentioned above may provide language learners with more chance to use new words and increase the possibility of their acquisition and retention, more class time is inevitably devoted to language-focused learning. As a result, the time allocated for discourse comprehension is occupied to some degree. Some studies (Kim, 2008; Lu, 2013) adopted the composition writing task, but the learners were explicitly instructed to incorporate the target words in the task, and thus lost part of their initiative.

Nevertheless, it doesn't mean the incompatibility of content-focused activity and vocabulary acquisition. Nation (2007) recommended the integration of different strands and pointed out that language-focused learning could occur in the context of meaning-focused work. That is to say, students can try to do output tasks that are content-focused, and meanwhile, have an opportunity to learn vocabulary incidentally.

With the purpose to examine the role of generative processing in IVA, Joe (1998) asked learners to read and retell the content of a text to incorporate and clarify unfamiliar words. The results showed that more use and retrieval of target words in retelling was likely to be an indicator of better IVA. But for lack of the comparison between reading only or vocabulary-focused condition and content retelling condition, the contribution of retelling remained unclear. Sun (2017) offered evidence of the usefulness of content-focused generative activities, finding that the participants who were required to read the picture books and share their interpretation and opinions about the input text in groups outperformed those who read only, or read with vocabulary instruction. Although there was no clear demand that students should pay attention to and use the target words, such generative task was regarded to create opportunities for them to recycle particular words.

Nevertheless, Yang et al. (2017) to some extent doubted the effect of text-based tasks on vocabulary uptake where there was no explicit focus on the target words. Checking the task design, it is easy to find that when students were required to write a short essay on the related topic, there is no need for them to recycle the original text or related vocabulary. Therefore, there was no surprise that no target words were used in the essays. Consequently, in the delayed post-test, the learners' performance in this condition was inferior to that of the word-focused conditions. Differently, Rassaei (2015) set story summary writing tasks after reading input. It turned out that in this case, participants would use target words in their summary writing. The results supported the idea that summary

writing provided chance for learners to use the new words, thus consolidating vocabulary knowledge.

These are evidences which support that summing up or retelling the input material launches incentive to use the target words, and generates vocabulary learning. The results are in accordance with some related theory as well. Firstly, the adoption of content retelling is also in keeping with Swain's (1995) Output Hypothesis. Swain claimed that the activity of producing may bring learners' attention to what they do not know or only know partially. The language production stage reveals the gap between what they are able to say and what they want to say. At the same time, production of language creates the opportunity for learners to test their hypotheses about the use of target language. When they attempt to use certain grammar rule, they may wonder if the rule could bring about successful communication, or may lead to negative feedback.

In addition, retelling included all three basic elements of the Involvement Load Hypothesis (ILH), i.e., "need", "search", and "evaluation" in one task. First, it develops the need for learners to tap lexical resources they reckon useful in their output. Second, when there are lacunae in their oral performance, learners are motivated to search for missing pieces when they watch the video again. Third, the process of revisiting the video is also the process of evaluating. They reflect on how they performed by comparing their use of new words in the self-provided context with the model in the video. In this respect, ILH confirms the rationality of the input-output-input circle.

Finally, this output task can also fit into the framework of Technique Feature Analysis (TFA) which added "motivation" and "retention" based on ILH. When retelling, learners need to use the newly learned words, which is exactly the process of generation. And because the input text is not available when learners do the output task, they have to retrieve what they have heard from memory.

Since audio-visual material is selected as input material in this study, oral content retelling is adopted as the output task, considering the consistency in linguistic medium. In this respect, Nguyen and Boers (2019) studied the usefulness of input-output-input activities on IVA for Vietnamese EFL learners in college. Students needed to watch a TED Talk video twice and sum up the content in English after their first viewing. Their study shows that the summary task is indeed beneficial to IVA, but the vocabulary uptake rate is rather low.

Vocabulary acquisition has attracted considerable attentions. Although researchers have carried out many studies on this topic and got fruitful achievements, few studies have concentrated on the incidental acquisition of English vocabulary engendered by doing content retelling tasks. In addition, most studies have examined vocabulary learning from text-based input, but comparatively little is known about the effectiveness of audio-visual input. Further, most of the researches only focus on college students' IVA, but few on high school English vocabulary learning and teaching.

In light of the inadequacies, the present study adopted authentic audio-visual material as input, and content retelling as output, forming an input-output-input circle. It is intended to investigate the effect of content retelling on EFL learners' vocabulary uptake and retention through empirical research. Hence, the following two questions are addressed:

What is the effect of content retelling between two viewings of an audio-visual material on vocabulary acquisition?

Do students use or attempt to use target words in their retelling? If yes, is the (attempted) use of target words positively associated with the acquisition and retention of the meaning of these words?

3. Method

3.1. Participants

Two classes of Chinese senior high school learners of English (n = 68) participated in this quasi experiment. The participants, at the age of 15 or 16, were all grade 10 students of a secondary school, and were in their 7th year of formal English instruction. There were 34 participants in each group. A t-test for independent samples was carried out to evaluate the learning proficiency of the participants in the two groups, which indicated no significant difference: $F = 0.293$, $t = 0.146$, $p = 0.885 > 0.05$. Besides the participant learners, two English teachers were invited as the assistants in this experiment to help select the input material and target words, and second mark the vocabulary tests.

3.2. Instruments

3.2.1. Input material

For EFL learners, TED Talks have many obvious advantages. They are authentic and unmodified learning materials, which promote the development of language skills and can assist to cultivate students' learning ability. Since most of the TED Talks are relatively short (most 5–18 min), they are thus applicable for classroom English vocabulary teaching.

To select a proper input material, fifteen senior high school learners, who were roughly at the same English level with the participants, were invited to select the ultimate input TED Talk of the experiment. Based on their feedback, a TED Talk whose title is *Got a wicked problem? First, tell me how you make toast* was chosen as the input material.

Since most of the EFL learners in China have better English reading skills than listening skills, they are prone to count on reading skills to circumvent listening input. To advocate an all-round development, this study focuses on the listening component. The one-way academic listening, which resembles attending English lectures in the university, can equip students with the ability to adapt to learn from English lectures. Thus, a TED Talk without captions was used as the input material.

3.2.2. Target words

Considering the predictability of word meanings, participants' vocabulary level and the volunteer students' feedback, twenty words were selected as the potential target items in this study (see Appendix I). The actual target words for evaluating the effect of content retelling by making a comparison between the pretest and the posttest would be decided according to participants' pretest results.

3.2.3. Vocabulary knowledge test

To measure vocabulary knowledge, a meaning recall test was adopted in this study. It is a receptive knowledge test where target words are provided in English as hints, and participants are required to provide the word meanings in their mother tongue (L1). A learner can be acknowledged as knowing the meaning of a word, when he/she is able to recognize the word in listening and recall

the meaning.

Since the input material was a TED Talk video without captions, participants were not shown the target words in written form. To ensure the consistency in linguistic medium, spoken forms of the target words were supplied as prompts in the test. Records of general American English pronunciation of the target words were downloaded from the e-dictionary. The audio recordings were edited together to form a complete record with twenty-second intervals between each word. This interval had been proved by a pilot study to be long enough for participants to provide the L1 translations.

Altogether, the participants did the test three times: two weeks before the experiment, immediately after the experiment, and two weeks after the experiment, as pretest, immediate post-test, and delayed post-test, respectively. The two post-tests were used to examine the learners' vocabulary gain and retention. In order to avoid participants' recognition memory for the word sequence, the order of the target words was randomized each time.

3.2.4. Text comprehension test

Fifteen statements about the content of the discourse were designed to gauge the participants' understanding of the input material. Based on Wagner's (2002) two-factor model of top-down and bottom-up processing to operationalize academic listening comprehension, the true/false test included the three basic question types to test whether learners can: 1) recognize the purpose or main idea of the input; 2) make inferences about implication, logic connections, pragmatic meaning, or the speaker's attitude; 3) identify crucial supporting details and factual information. Examples of the questions of each type are listed as below: "The speaker finds a simple but effective way to solve problems—drawing how to make toast" (main idea); "By rearranging the sticky notes or cards with nodes, people analyze events step by step to make the whole process clear" (implicit); "The speaker mentions a website where people can find various practices which they can learn from" (factual).

3.3. Procedure

The whole process was divided into pretest, experiment, immediate post-test, and delayed post-test (see in **Table 1**).

The pretest was conducted two weeks before the experiment, whose main purpose was to check whether the participants had already known the potential target words, so as to decide the actual tar-

Table 1. Study design

	The retelling condition (*n* = 34)	**The comparison condition (*n* = 34)**
Pretest (2 weeks before)	Vocabulary knowledge test (1)	
Input	The first viewing	
Output	Participants made oral retelling (recorded) within 5 minutes.	Participants were given 5 minutes to recall and review.
Input	The second viewing	
Text comprehension test	A true/false test	
The immediate post-test	Vocabulary knowledge test (2)	
The delayed post-test	Vocabulary knowledge test (3)	

get words.

In the experiment, the participants in the two groups were asked to watch the TED video twice in an effort to answer the comprehension questions. They were encouraged to take out a piece of paper and take some notes. Learners in the retelling group were told that they would participate in the 5-minute content retelling task after the first viewing, and their oral retelling would be recorded. Those in the comparison condition were informed that they would have 5 minutes after the first viewing. During this period, they could review their notes and think over what they have heard.

The immediate post-test was administered instantly after the second viewing to check vocabulary acquisition. After that, all the participants were given the comprehension test, in which they were allowed to refer to their written notes, but the TED Talk video wasn't accessible.

The delayed post-test was carried out two weeks after the experiment to check the participants' retention of the target words. The words tested were the same, but with different sequence. It is worth mentioning that the participants weren't forewarned of the pretest, immediate post-test and delayed post-test.

3.4. Data collection and analysis

The pretest was administered with the aim to remove the words that participants had already known from the potential target word list, and decide the actual target words. From the studies collected in Ellis' book (1999), potential target words should be rejected if they have already been acquired by more than 10% of the participants. The pretest results showed that none of the potential words were known by over 10% of the participants. Therefore, there were 20 actual target words in this study.

The research aims to study the effect of content retelling on vocabulary uptake and retention. In the following parts, the term "uptake" is used to refer to the performance of vocabulary gain immediately after the task, and is measured by the outcome indicated in the immediate post-test. It is determined by subtracting the score of the pretest from that of the immediate post-test. "Retention", which refers to the performance of long-term IVA, is measured by the outcome indicated in the delayed post-test. Thus, it is calculated by subtracting the score of the pretest from that of the delayed post-test.

4. Results

The descriptive statistics for the scores obtained under the two conditions are presented in **Table 2**. Since there are altogether 20 target words in this study, the maximum score is 20.

As is shown above, the retelling group achieved higher mean scores in all the tests. Further anal-

Table 2. Vocabulary meaning gains (SD in parentheses)

Groups	Pretest scores	Immediate post-test scores	Delayed post-test scores	Uptake scores	Retention scores
Comparison	0.32 (0.59)	1.89 (1.57)	1.12 (1.07)	1.56 (1.38)	0.79 (0.95)
Retelling	0.35 (0.74)	3.77 (1.56)	3.18 (2.02)	3.41 (1.16)	2.82 (1.64)

yses were made with the aid of analysis of variance (ANOVA) tests, independent samples *t*-tests, paired samples *t*-tests, and Cohen's effect size *d*.

To assess whether difference lies in the participants' knowledge of the meaning of the target words under the two conditions before the experiment, an independent sample *t*-test was conducted. It revealed no significant difference between the two groups: $t = 0.18$ ($p = 0.53 > 0.05$). Therefore, it's reasonable to ascribe the cause of different word gains to the insertion of the content retelling output task.

One-way ANOVA tests for correlated samples were then conducted to gauge the difference in the learner's scores across the three tests. A significant difference was found: $F(2,66) = 99.34$, $p < 0.05$, Partial Eta Squared = 0.75. It means that at least one test's mean score is significantly different from the others. Then, paired samples *t*-tests were followed to further explore which one(s) is/are significantly different from the others.

Based on the *t*-tests for paired samples to examine which test is different from others, the alpha level of each pairwise comparison test is 0.016. The result showed that, in the content retelling condition, there are significant differences between pretest and immediate post-test ($p < 0.016$), and pretest and delayed post-test ($p < 0.016$). But no significant difference is found between the participants' immediate and delayed post-test scores ($p = 0.048 > 0.016$). It indicates that participants in this group significantly acquire more new words and retain vocabulary knowledge well over time. The same tests were done to the gauge the scores of the comparison group. In this condition, there is also a significant difference in the participants' scores across the three tests. And the differences between pretest and immediate post-test ($p < 0.016$), and pretest and delayed post-test ($p < 0.016$) reach significance as well. Different from what is shown in the retelling group, the difference between the immediate post-test and delayed post-test is also significant ($p = 0.008 < 0.016$).

Repeated independent samples *t*-tests were run to compare the size of vocabulary uptake and retention between the two conditions. The results showed that both the vocabulary uptake and vocabulary retention were significantly larger in the retelling condition than in the comparison condition, with $t = 6.011$ ($p < 0.001$) and $t = 6.245$ ($p < 0.001$), respectively. Cohen's effect size *d* was also computed to examine the difference in uptake and retention between the two conditions. Such a difference was consistently found to be large, with $d = 1.46$ in the case of the uptake scores and $d = 1.51$ in the case of the retention scores.

Since we have found that learners performed better in the retelling condition both in vocabulary uptake and retention, it needs to figure out whether their attempted use of the new words takes credit for the success.

After viewing all the videos which screened the participants' retelling performance, 6 out of the 20 target words were observed to be used in students' content retelling. Among them, "node" occurred most (30 tokens). Another frequently used word was "visualize" (19 tokens). The next one was "toaster" (9 tokens). Other words included "synthesize" (2 tokens), "refine" (2 tokens), and "fluidly" (1 token).

30 of the 34 participants incorporated at least one target words in their retelling. According to the pre-test, the participants didn't know the words before the experiment. After viewing the video, however, they successfully understood the meaning of the words, added them to their arsenal

of vocabulary, and involved them in their oral retelling to express ideas, in both semantically and grammatically appropriate ways. What's more, it deserves to be mentioned that they did not simply copy or make duplicates of the TED speaker's utterances. Instead, new expressions, sentences, and context were created. Here are some examples of how the target words were included in the content retelling.

S11: You can use **node**s and links to make our ideas into a systems model…Some Americans make bread with **toaster**, and some Europeans make with pan.

S24: They try to analyze the procedure, and they did it more **fluidly**… When we **visualize** our thoughts by drawing or writing notes, we can create our map and make our ideas more logical and clearer… So next time, when difficult problem shows up, we can make our ideas visible and touchable by using **node**s and links, especially in group.

S27: Also, people should make group notes. They will **refine** things again and again.

In the retelling videos, there are also some obvious long pauses (over 2 seconds). The pauses appear in the middle of a sentence, and certain target words are appropriate to fill in the blank. Below are some instances of the attempt.

S3: That is to make things visible, to [**visualize**] them. It can make our thoughts clearer.

S8: It is messy at first, but after the colleagues draw them [**collaboratively**], I mean, together, it is clearer.

S16: Many people stick notes in different order, and they [**refine**] to make complex things to be simple. People work [**collaboratively**], and group works better than individual.

The data showed that if the participants used the target words in their retelling, the uptake rate and retention rate of these words are 85.2%, and 64.8%, respectively. And what is interesting is that, even though learners may fail to utter the exact word in their output, the meaning of nearly half of those words were retrieved in the post-tests.

In the listening comprehension test, the participants in the retelling condition gained a mean score of 12.26 (SD = 1.24), and those in the comparison group gained 12.21 (SD = 1.55) out of the maximum score of 15. According to the t-test for independent samples, there was no significant difference between the two conditions, with $t = 0.17$, $p = 0.86 > 0.05$.

5. Discussion

The objective of this study is to examine the effectiveness of the content retelling output task. Turn first to vocabulary uptake. Our findings suggested that, the participants, when required to do the retelling task, learned 3 (15%) words out of the 20 target words on average. Though the lexical gain is arguably modest, it exceeded the gain in the comparison condition, where there was no insertion of the content retelling oral task between two viewings of a TED Talk video.

When it comes to vocabulary retention, findings were to some extent similar to what was found for uptake—participants in the retelling group outperformed their counterparts in the comparison group. The latter only memorized less than 1 (5%) word.

The scores of the retelling group didn't have significant difference between immediate post-test and delayed post-test. While for comparison group, the scores of the delayed post-test were significantly lower than that of the immediate post-test. It revealed that the content retelling output task was helpful for participants to entrench what they learnt in their long-term memory.

In the light of the outcome of vocabulary uptake, as well as retention, the potential benefits of implementing the content retelling task, a content-focused output task, are quite obvious. The use of content retelling follows two main principles: firstly, it should still focus on the content of the text; secondly, it should create opportunities for students' use of the new words, and thus lead to their vocabulary acquisition.

This research showed that when learners watched an audio-visual material twice with an insertion of retelling task, they acquired more words compared with their counterparts who were not required to do the output task. The announcement of the retelling task provided learners with a clear purpose to deal with the input material in the first viewing, and may thus listen attentively and notice some relevant words. Through retelling, learners reorganized and recreated the content according to what they saw, heard, and understood. They consolidated the word messages they encountered in order to use these words in the new context created by themselves. Especially, when the learners found their retelling incomplete, or noticed that some points were not clearly and accurately stated, they were motivated to seek for related words in the second viewing in order to fill in the information gaps or correct their mistakes. The three steps in the input-output-input circle are greatly interlinked. In this way, although learners were not stipulated to pay special attention to and use the new words, they were closely associated with these words with the guidance of this content-focused task.

In the acquisition and retention of vocabulary, the attempted use of new words played a predominant role. This productive attempt involved mental rehearsal with target words as well as deeper cognitive processing, in which learners not only remembered and developed a basic understanding of the words grammatically and semantically, but also tried to incorporate them in their own reproduction. With practical use and retrieval of the target words, the learning pathway was enhanced. Besides, EFL learners were more likely to be provided with better conditions for word generation owing to this task. When a new word was used in their oral retelling, learners were motivated to assess the appropriateness of this target lexical item in a new context. And when they met the word again in the second viewing, they evaluated their production by referring to the original model in the audio-visual input. Learners not only paid attention to particular meaning of the word, but also tried to find out whether it fits in the specific linguistic context. In this way, form, meaning, and usage were successfully combined in the meaningful context, leading to better opportunities to retain newly learned words in the long-term memory.

Different from what was designed by Yang et al. (2017), the content retelling task in this study entailed a greater need to incorporate relevant words from the original text. When the output task was more closely connected with the input material, learners were unconsciously guided to use the target words more. The general properties of the target word were explicitly connected with the word in the retrieval and generation. In addition, participants were also able to remember over half of the words which they were attempted to use at least two weeks later. According to Hermann Ebbinghaus' forgetting curve, after a day or two, we typically forget around 75% of what we have learned if not review or reinforce our learning. And about 14 days later, learners can only remember

less than 25%. Therefore, compared with the natural rate of forgetting, the outcome of this experiment is definitely gratifying.

Based on the research results, the words with the highest rate of acquisition in this experiment are "toaster", "visualize", and "node", which were exactly the top three words learners used in the retelling. It is thus reasonable to infer that there may be a significant correlation between students' attempted use and their vocabulary acquisition.

The results were in accordance with Swain's (1995) Output Hypothesis. To complete the retelling task, learners needed to grasp and remember some lexical terms. And after finishing retelling, they came to realize language deficiencies in their lexical resources. Thus, in the second viewing, it was more likely for them to pay attention to special words to fill the knowledge gaps. In this way, the input-output-input circle well promoted noticing, providing opportunities for learners to attend to lexical items at every step. Besides, the retelling task invited language learners to try new expressions and use the new words. This action may help preserve the newly encountered words in their memory. Again, during the second viewing, they were natural to check whether they used the new words correctly by comparing how they were presented by the speech speaker with their own use. Even though learners might fail to retrieve the exact pronunciation of certain word, the pauses and hesitations indicated their efforts of retrieval, and probably, they had already known the meaning of that word. Therefore, when revisiting the input video, they paid special attention to that word, leading to better acquisition and retention. In this respect, the second viewing was necessary and meaningful.

The findings also concurred with Laufer and Hulstijn's (2001) Involvement Load Hypothesis, and Nation and Webb's (2011) Technique Feature Analysis. When the learners actively uttered new words in the context, they recalled the target lexical items, and used them generatively in new ways they had not heard before. When combining the word with other words, they assessed whether it fitted in the new context at the same time. In the light of ILH, the involvement index of content retelling was fairly high—with the involvement index of 4 (+ need, + search, ++ evaluation).

In a nutshell, attempted use of the target words enhances both learner's uptake and retention of the meaning of the words.

6. Conclusion

This study focuses on the effect of content retelling on vocabulary uptake for EFL learners. Under the guidance of the existing researches and output theories, a quasi-experiment was designed and successfully conducted. Through systematic analysis of the statistical data, the major findings have been obtained. Firstly, content retelling is an effective output task, where EFL learner's practice in verbal rehearsal of what they have heard facilitates their vocabulary acquisition. Secondly, the attempted use of new words plays a predominant role in the acquisition and retention of vocabulary. In addition, this study suggests that significant vocabulary learning could occur through being exposed to audio-visual input. It should also be noted that this study is liable to some limitations. The first lies in the sample, whose size is not big enough, so the result may not represent the general vocabulary acquisition features of all the high school learners. And in this study, only one type of audio-visual input (TED Talk video) and one type of output task (content retelling) were examined. It needs further investigation to find out whether the same trend happens with other audio-visual

materials, and whether content retelling gets the edge on other meaning-focused productive activities. Finally, it is hard to define to what extent a word is regarded as useful or even essential in the retelling output. In the present study, words like "prominent", "spontaneously", and "intuitively" were not used in the retelling, and as a consequence, they were poorly acquired and recalled in the post-tests.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgements

We would also like to thank all the participant teachers and students in the research. We are also grateful to the reviewers for their valuable comments.

References

Anthony ARB (2008) Output strategies for English-language learners: Theory to practice. *Reading Teacher* 61(6): 472–482.

Brown R, Waring R, and Donkaewbua S (2008) Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language* 20(2): 136–163.

Craik FIM and Lockhart RS (1972) Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11(6): 671–684.

Douglas V (2016) Enhancing English academic vocabulary acquisition and retention in intensive English programs with the involvement load hypothesis. *Theory and Practice in Language Studies* 6(12): 2237–2244.

Eckerth J and Tavakoli P (2012) The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research* 1(2): 227–252.

Ellis R (1999) *Learning a Second Language through Interaction*. Amsterdam: John Benjamins.

Horst M, Cobb T, and Meara P (1998) Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language* 11(2): 207–223.

Hu HM and Nassaji H (2016) Effective vocabulary learning tasks: Involvement Load Hypothesis versus Technique Feature Analysis. *System* 56: 28–39.

Huckin T and Coady J (1999) Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition* 21(2): 181–193.

Hulstijn JH and Laufer B (2001) Some empirical evidence for the Involvement Load Hypothesis in vocabulary acquisition. *Language Learning* 51(3): 539–558.

Hwang K and Nation P (1989) Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language* 6(1): 323–335.

Joe A (1998) What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics* 19(3): 357–377.

Kim Y (2008) The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning* 58(2): 285–325.

Laufer B (2003) Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review* 59(4): 567–587.

Laufer B and Rozovski-Roitblat B (2011) Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research* 15(4): 391–411.

Laufer B and Hulstijn J (2001) Incidental vocabulary acquisition in a second language: The construct of task-in-

duced involvement. *Applied Linguistics* 22(1): 1–26.

Lu M (2013) Effects of four vocabulary exercises on facilitating learning vocabulary meaning, form, and use. *TESOL Quarterly* 47(1): 167–176.

Min HT (2008) EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning* 58(1): 73–115.

Nagy WE, Herman PA, and Anderson RC (1985) Learning words from context. *Reading Research Quarterly* 20(2): 233–253.

Nation ISP (2007) The four strands. *Innovation in Language Learning and Teaching* 1(1): 1–12.

Nation ISP and Webb S (2011) *Researching and Analyzing Vocabulary*. Boston: Heinle.

Neuman B and Koskinen P (1992) Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27(1): 94–106.

Nguyen C and Boers F (2019) The effect of content retelling on vocabulary uptake from a TED Talk. *TESOL Quarterly* 53(1): 5–29.

Paribakht S and Wesche M (1997) Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In: Coady J and Huckin T (eds.) *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*. New York: Cambridge University Press, pp.174–200.

Peters E (2012) The differential effects of two vocabulary instruction methods on EFL word learning: A study into task effectiveness. *International Review of Applied Linguistics* 50(3): 213–238.

Rassaei E (2015) Effects of three forms of reading-based output activity on L2 vocabulary learning. *Language Teaching Research* 21(1): 76–95.

Richards J and Schmidt R (2010) *Longman Dictionary of Language Teaching and Applied Linguistics* (4th edn.). Malaysia: Pearson Education.

Secules T, Herron C, and Tomasello M (1992) The effect of video context on foreign language learning. *The Modern Language Journal* 76(4): 480–490.

Sousa DA (2006) *How the Brain Learns?* (5th edn.). Thousand Oaks: Corwin Press.

Suk N (2017) The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading Research Quarterly* 52(1): 73–89.

Sun CH (2017) The value of picture-book reading-based collaborative output activities for vocabulary retention. *Language Teaching Research* 21(1): 96–117.

Swain M (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In: Gass S and Maddern C (eds.) *Input in Second Language Acquisition*. Rowley: Newbury House, pp.235–253.

Swain M (1995) Three functions of output in second language learning. In: Cook G and Seidlhofer B (eds.) *Principles and Practice in Applied Linguistics*. Oxford: Oxford University Press, pp.125–144.

van Zeeland H and Schmitt N (2013) Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System* 41(3): 609–624.

Vidal K (2003) Academic listening: A source of vocabulary acquisition? *Applied Linguistics* 24(1): 56–86.

Vidal K (2011) A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning* 61(1): 219–258.

Wagner E (2002) Video listening tests: A pilot study. *Working Papers in TESOL and Applied Linguistics* 2(1): 1–39.

Wen Q (2008) [On the output-driven hypothesis and the reform of English-skill courses for English majors]. *Foreign Language World* (2): 2–9.

Wen Q (2015) [Developing a theoretical system of production-oriented approach in language teaching]. *Foreign Language Teaching and Research* 47(4): 547–558, 640.

Weyers JR (1999) The effect of authentic video on communicative competence. *The Modern Language Journal* 83(3): 339–349.

Yang Y, Shintani N, Li S, et al. (2017) The effectiveness of post-reading word-focused activities and their associa-

tions with working memory. *System* 70: 38–49.

Zou D (2017) Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research* 21(1): 54–75.

Appendix I: Target words

Target words and their frequency of occurrence
inspection (2), sliced (1), toaster (4), illustrate (1), visualize (6), node (18), tangible (2), diagram (1), intuitively (1), trivial (4), prominent (1), emerge (3), spontaneously (1), fluidly (1), synthesize (1), collaboratively (1), refine (5), template (1), thorny (1), confront (1)

Whioce welcomes both current and future authors to submit their works. Here you can find the information you need about publishing with us.

Why publish with Whioce?

- High visibility of Whioce publications and discoverability of content
- All Whioce journals available via a robust online platform
- DOI registration for all Whioce articles
- All Whioce journals available as separately downloadable items
- High quality print publication from Whioce
- Clear guidelines for authors and editors (see Journals section)
- Whioce offers a dedicated team of skilled professionals, marketers and editors
- Each journal and their accompanying articles receive our individual attention

Contact Us

Advertising Enquiries

We welcome advertisements. Please drop us an email at advertising@whioce.com

Advertising Specifications: Required format is TIFF and PDF. Ads must be completed and sized at 100% (trim: 210*275, bleed 216*281, type area 175*240). Ads to be saved in high resolution format for print publication (600 dpi for images containing pictures and text). Electronic transfer PDFs that are less than 5 MB can be submitted by email to advertising@whioce.com

Subscription Enquiries

If you are interested to subscribe to any of our journals, please drop us an email at subscription@whioce.com

For institutional subscriptions, Whioce offers a site license policy. Institutions may also purchase a separate print subscription. We offer librarians the choice of selecting the media most appropriate to their institution's budget and content usage patterns. Contact us at subscription@whioce.com

Editorial Enquiries

If you have any queries on the editorial content of our journals or if you are interested to submit your manuscripts to us, please drop us an email at editorial@whioce.com. Alternatively, you can submit your manuscript at the journal's home page.

WHIOCE PUBLISHING PTE. LTD.
PROVIDING
FIRST-CLASS SCIENTIFIC INFORMATION
FOR TOP SCHOLARS

Whioce Publishing Pte. Ltd.

Address: 7030 Ang Mo Kio Avenue 5#04-15
Northstar@AMK Singapore 569880
Website: www.whioce.com

ISSN 2705-0602
01
9 772705 060047