

Article

A Longitudinal Study of the Acquisition of the Polysemous Verb 打 *dǎ* in Mandarin Chinese

Jidong Chen *  and Xinchun Wang

Department of Linguistics, California State University, Fresno, CA 93720, USA; xinw@mail.fresnostate.edu

* Correspondence: jchen@csufresno.edu

Received: 16 February 2020; Accepted: 13 May 2020; Published: 18 May 2020



Abstract: Lexical ambiguity abounds in languages and multiple one-to-many form-function mappings create challenges for language learners. This study extends the theoretical approaches to the acquisition of polysemy to the Mandarin verb 打 *dǎ*, which is highly polysemous and among the earliest verbs in child speech. It analyzes longitudinal naturalistic data of nine children (1;05–3;10) from two Mandarin child speech corpora to explore the developmental trajectory of different senses of 打 *dǎ* and the role of input. The results support a continuous derivational and restricted monosemy approach: children initially extract a core feature of 打 *dǎ*, but only apply it in a restricted way, reflected in a small number of senses in a limited set of semantic domains and syntactic frames, revealing an early preference for initial unambiguous form-meaning mappings. Mandarin-speaking children’s production mirrors the semantic and syntactic distribution of the input, supporting the usage-based approach to the acquisition of polysemy that meaning is derived from the confluence of lexical and syntactic cues in the usage patterns in the input. Our research is the first longitudinal study of the emergence and development of polysemous verbs in Mandarin and has pedagogical implications for teaching Mandarin as a second language.

Keywords: Mandarin; child language acquisition; polysemy; verb semantics

1. Introduction

Lexical ambiguity abounds in languages and multiple one-to-many form-function mappings create challenges for child language learners (e.g., Clark 1993). Lexical ambiguity can arise from polysemy or homonymy (e.g., *male* and *mail*). Polysemy is characterized as a single word associated with multiple related sense in contrast to homonymy, a single form associated with multiple unrelated meanings (e.g., Vicente and Falkum 2017), as illustrated by examples from child language (1) (cited from Tomasello 1992) and (2) (extracted from the Tong corpus, Deng and Yip 2018) below.¹

¹ CHI = child. The age of the children is conventionally notated as years;months. Utterances are transcribed in Chinese characters and Pinyin, the official Romanized transcription of Chinese.

McKercher (2001) studied the acquisition of the multiple senses of the preposition *with* (e.g., ‘attribute’, ‘nominal’, ‘instrument’, ‘accompaniment’, ‘manner’, etc.) and proposed two different developmental approaches, the monosemy approach and the multiple-meanings (polysemy) approach. The monosemy approach predicts that children would start with an underspecified (or core) representation and use a range of different senses simultaneously from the beginning, and the multiple-meanings approach predicts that children would acquire each sense of *with* item-by-item (e.g., *with*_{instrument}, *with*_{accompaniment}, etc.) and different senses would emerge at a different time. McKercher (2001) analyzed longitudinal data of six English-learning children with varied length of data between the ages of 1;3 and 4;10 from the Child Language Data Exchange System (CHILDES) (MacWhinney 2000). He found that the children produced a range of semantic roles in their early speech and showed a more general meaning of *with* (i.e., ‘having’) than each of the specific senses, which supported the monosemy approach. Kidd and Cameron-Faulkner (2008) analyzed a dense (one hour five days per week from 2;0.12 to 3;1.30) longitudinal corpus of one English-learning child, and argued that children initially extracted a core feature of *with* (i.e., spatial proximity or co-location) but only use it in a restricted way (i.e., using the core meaning for some time before gradually extending to other senses), which suggested children’s preference for an initial one-to-one form-meaning. Kidd and Cameron-Faulkner (2008) also examined in detail the input of the target child and showed that the relative frequency of senses was similar in the child speech and the input and argued that the input offered reliable cues for the uses of different senses, including the semantics of the verbs and the construction in which *with* occurred most often. A similar argument was made in the study of children’s acquisition of various forms of the verb *go* longitudinally, where a good predictor of children’s usage was the input frequency of *go* in different structures and the specific meanings with particular forms of *go* (Theakston et al. 2002).

1.2. Semantics of the Mandarin Verb 打 *dǎ* and Its Acquisition

Our study aims to extend the theoretical approaches to the acquisition of polysemy in Mandarin Chinese (henceforth Mandarin), focusing on the Mandarin verb 打 *dǎ*. The Mandarin verb 打 *dǎ* is ‘hit/beat’ is one of the most frequently used verbs in Mandarin (e.g., Gao 2001) and is highly polysemous. *The Contemporary Dictionary of Chinese* (2016, 7th edition) lists 24 senses of 打 *dǎ*, e.g., 敲门 *dǎmén* ‘knock on door’, 打架 *dǎjià* ‘fight’, 打家具 *dǎjiājù* ‘make furniture’. The basic meaning of 打 *dǎ* refers to a physical action of the hand and can extend from its basic prototypical meaning to a wide range of actions or events involving a hand or instrument as well as to events that are metaphorically hand-involving (e.g., 打折 *dǎzhé* ‘discount’, 打听 *dǎtīng* ‘inquire’). The Chinese Wordnet (Huang et al. 2010) lists 121 senses of 打 *dǎ* based on a detailed lexical semantic analysis ranging from concrete actions involving hand manipulation (e.g., 打桌子 *dǎzhuōzi* ‘knock on table’) to metaphorically extended senses such as 打天下 *dǎtiānxià* ‘establish power’, 订票 *dǎpiào* ‘buy tickets’, 打光 *dǎguāng* ‘polish’. These dictionary-based depictions provide a descriptive picture of the senses of 打 *dǎ*, but do not offer a systematic categorization and characterization of the many senses and the derivational relations between them.

Gao’s (2001) study of the physical action verbs in Chinese (including 打 *dǎ*) filled this gap. It provided a comprehensive analysis of the lexical semantics of 打 *dǎ* in a cognitive linguistics framework that grounded verbal semantics and argument structures in the nature of human bodies and their interactions with the physical, social, and cultural environment (cf. Lakoff and John 1999). Chinese physical actions verbs could be characterized in terms of typical bodily experiences including, e.g., the body parts involved (e.g., hand, foot, head, mouth, etc.), physical contact, motion, and intention. Gao (2001) analyzed the distribution of different senses of 打 *dǎ* in two large corpora of Mandarin (the Academia Sinica Balanced Corpus and the Beida Institute of Computational Linguistics corpus). She found 152 distinct senses of 打 *dǎ* in 27 semantic domains (e.g., game playing, physical punishment, open, fastening, construction, covering, etc.) in five broad semantic representations (categories) that set up “a linkage of all the sub-fields from the descriptions of the most prototypical actions if 打 *dǎ*

with hand contact as the focus down to the metaphorical uses with very vague or even non indications of hand actions of any type” (Gao 2001, p. 166). Table 1 presents the five semantic representations and the distinct senses in each semantic category, adapted from Gao (2001, pp. 163–65). Gao also found that the two semantic categories defeat and physical punishment are the most frequent in both corpora, despite the high degree of polysemy.

Table 1. Semantic representations and senses of the target verb 打 dǎ (Gao 2001).

Semantic Representations	Specific Senses
Type 1: Physical action focusing on hand contact	Defeat (play games, battle, fight, kill, attack, break, smash) Physical punishment (beat, hit, punch, spank, whip, slap, strike) Open (turn on, take out, unpack) Fastening (pack up, knot, tie)
Type 2: Hand action (mostly with instrument)	Construction (make, build, dig, drill, burrow) Launching (shoot, fire, send, report, pump, post, set up, signal) Cover (dress up, pretend, spray, wax, powder, plaster, clout, drench, label, paint, polish, wrap) Insert (inject, nail, hammer, knock) Possess (fetch, catch, have, hunt, ladle, hold) Mark (type, work out, engrave, press, label, print, stamp) Sound source (drum, knock, cap, pound, crow, flap, ring, tap, whistle) Motion (stir, return, move, mix) Reflection (flash, reflect) Collection (gather, reap, sweep, get in, get) Removal (prune, peel, knock out, get rid of, rob, thresh) Engagement (work)
Type 3: Physical action with physical contact unspecified	Physiological reaction (shiver, yawn, doze off, snore, sneeze, cheer up, hiccup, nod, stupefy) Gymnastic feat (roll, tumble, loop) Posture (bare, remain, zazen)
Type 4: Metaphorical uses (hand action traceable)	Social interaction (call, contact with, gesticulate) Business deal (discount, buy, invest) Authoritative conduct (issue, score) Legal activity (go to court)
Type 5: Metaphorical uses (hand action untraceable)	Mental activity (plan, seek, make, think of, calculate, disturb, consider, decide, draw analogy, concern, estimate) Verbalization (ask about, greet, bet, draw, interrupt, talk, chat, cry out, discuss, question, speak) Opposition (defend, protest) Visual contact (meet, bump into, look at)

The verb 打 dǎ has been found among the earliest verbs in child speech and input frequency and children’s physical development, growth environment, and cognitive understanding of the actions correlate directly with multiple senses (Gao 2015). Little research has been conducted to examine monolingual Mandarin-learning children’s semantic acquisition of polysemous verbs. Two recent studies were most relevant to our inquiry of the acquisition of polysemous verbs in Mandarin. Sak and Gao (2016) examined the semantic knowledge of preschool bilingual Mandarin-English children in Singapore using elicited descriptions of pictures or videos depicting 打 dǎ actions. They analyzed 31 senses in the data and found that the semantic domains of “social interactions” (e.g., 打电话 dǎdiànhuà ‘call’) and “physical punishment” (e.g., 打耳光 dǎěrguāng ‘slap in the face’) are the most frequent, whereas semantic domains such as “fastening” and “possession” are least used. English was found to negatively affect the bilingual children’s use of 打 dǎ. It is argued that children’s exposure to action events and the competition of near synonyms of 打 dǎ accounted for the observed usage.

Zhang et al. (2010) examined the acquisition of eight polysemous words, among which three were polysemous verbs (i.e., 看 kàn ‘look’, 走 zǒu ‘walk’, and 给 gěi ‘give’), in a longitudinal corpus (weekly one-hour recordings) of a young monolingual Mandarin-learning child from age 1;6 to 3;0. They proposed three possible approaches for learning polysemous words: (1) a continuous approach, where multiple senses could be derived via metaphor or metonymy; (2) an independent approach, where each sense was acquired independently without derivational relations; and (3) a mixed strategy using both (1) and (2). Their analyses of the longitudinal emergence of the different senses of the three verbs supported the continuous derivational approach, and the derivational routes could proceed in three

possible ways, i.e., radiationally (i.e., multiple senses derived from one single sense simultaneously), serially (i.e., new senses derived from the previously acquired senses), or both radiationally and serially. The results also suggest that derivational routes could vary from verb to verb, subject to usage factors such as input frequency and functional needs: the verb 看 *kàn* 'look' showed a serial route where the basic sense emerged first and different derived senses appeared longitudinally, whereas the other two verbs, 走 *zǒu* 'walk' and 给 *gěi* 'give', showed an independent route, where multiple senses appeared simultaneously around a similar time, and the derived sense could appear early (e.g., the sense 'leaving' for 走 *zǒu* appeared before the prototypical sense 'walk'). Since 打 *dǎ* was not examined, it remains an empirical question if and how its acquisition fits in the proposed developmental routes in Zhang et al. (2010).

1.3. Research Questions

The current study aims to examine monolingual Mandarin-learning children's semantic acquisition of the polysemous verb 打 *dǎ*. Based on prior research, the acquisition of the multiple senses of 打 *dǎ* could potentially proceed in three different routes, following the monosemy approach (McKercher 2001), the restricted monosemy approach (Kidd and Cameron-Faulkner 2008), or the polysemy approach (similar to the independent approach proposed in Zhang et al. 2010). The continuous strategy in Zhang et al. (2010) shares with the monosemy approach and the restricted monosemy approach the emphasis on the derivational relations between all senses, but they differ in the claim about the starting point of the sense derivation in acquisition: the basic core or functionally most salient and frequent sense (Zhang et al. 2010), the basic core sense with derived senses fully accessible simultaneously (McKercher 2001), or partially accessible with restricted uses and slow extension to other senses (Kidd and Cameron-Faulkner 2008). It is thus interesting to investigate empirically how multiple senses of 打 *dǎ* develop longitudinally to evaluate the different approaches. This study aims to answer the following research questions:

1. What is the developmental trajectory (i.e., emergent order) of different senses of 打 *dǎ* in Mandarin?
2. How do Mandarin-learning children proceed in acquiring different senses of 打 *dǎ* in a multiple-meanings (polysemy) approach, a monosemy approach, or the restricted multiple-meanings (polysemy) approach?
3. How does input, including syntactic, semantic, and contextual cues, contribute to the acquisition of different senses of 打 *dǎ*?

2. Materials and Methods

We analyzed longitudinal naturalistic corpus data of 9 children (age range 1;05–3;10) in two Mandarin child corpora, one child from the Tong corpus (Deng and Yip 2018; MacWhinney 2000) and eight children from the Taiwan Corpus of Mandarin Chinese (TCMC) (MacWhinney 2000). The Tong corpus contains hour-long monthly recordings of interactions between the target Mandarin-learning child Tong and his caregivers (mostly mother, father and occasionally grandparents) from age 1;07 to 3;4, including a total of 22 transcripts (see Table 2), the largest dataset among the 9 children. The TCMC contains monthly naturalistic data from 10 children, age range from 1;05–4;03 with varied ages of the start and the end of data collection, and length of data recordings (see Table 1). The data from eight of the children were included in the analysis due to their early age (between 1;05 and 2;07) at the start of the data collection to explore the early emergence of the use of 打 *dǎ* and 2 children who were at 3;01 and 3;6 at the start of the data collection were excluded. A total of 375 tokens of 打 *dǎ* were produced in the child speech and a total of 809 tokens of 打 *dǎ* in the adult speech.

All utterances containing the target verb 打 *dǎ* in the children's speech were extracted with the Computerized Language Analysis (CLAN) program (MacWhinney 2000), and sample early utterances that contain 打 *dǎ* are shown in Table 3. At least 3 utterances above and below the target utterance were also extracted with the CLAN program to provide contextual information to determine the specific

senses. The verb 打 dǎ in the target utterances were coded for (1) the type of semantic representations and (2) the specific sense based on Gao’s (2001) lexical semantic analysis of 打 dǎ (cf. Table 1).

Table 2. Information about the target children in the longitudinal naturalistic corpora.

Child	Age Range	# Files	# Utterances	Types	Tokens	Token Freq. of 打 dǎ in child Speech	Token Freq. of 打 dǎ in Input	
1	Tong	1;07–3;04	22	9110	5240	30,798	175	294
2	Chou	2;01–3;04	16	4991	1482	14,025	33	44
3	JC	2;02–3;05	14	4955	1183	10,710	22	90
4	Pan	1;07–3;09	19	2661	942	7634	40	129
5	Wang	2;05–3;04	12	3149	1092	10,133	19	44
6	Wu	1;07–2;01	12	2785	922	7030	15	92
7	Wuys	2;07–3;10	10	1396	758	5223	23	40
8	Xu	1;06–2;05	11	2700	651	4315	14	7
9	Yang	1;05–2;09	13	1929	733	4787	34	69
Total							375	809

Table 3. Sample utterances containing the target verb 打 dǎ in child speech².

Child Age	Samples	Pinyin	Glossing	Translation
1;07.18	打开啦。	dǎkāla	hit open SFP	‘Opened!’
1;07.18	打球。	dǎqiú	play ball	‘Play with the ball.’
1;08.22	打牌。	dǎpái	play card	‘Play cards.’
1;11.21	打妈妈。	dǎmāma	hit mommy	‘Hit mommy.’
2;00.19	打锣了。	dǎluó	play gong SFP	‘Play the gong.’
2;00.19	妈妈给打开。	māmāgěi dǎkāi	mommy give hit-open	‘Mommy open (for me).’
2;01.17	我们来打牌。	wǒmēnlái dǎpái	we come play card	‘Let’s play cards.’

To explore the relationship between the child’s output and input, we further extracted and coded the semantics of the target verb 打 dǎ in the caregivers’ speech in the Tong corpus, the biggest dataset among the 9 children. We also coded all the target utterances in the speech of both Tong and his caregivers for the syntactic contexts or frames in which 打 dǎ occurs, e.g., VV (verb compound), VNP (verb followed by an object noun phrase), NPVNP (subject noun phrase followed by the verb and an object noun phrase), to see if the emergence of the multiple senses of 打 dǎ are closely tied to certain syntactic frames or constructions in the child and the caregivers’ speech. The rationale for this additional construction-based analysis comes from the findings that children are sensitive to the syntactic frames in which a verb occurs and use them to infer verb meanings (e.g., Gleitman 1990; Lee and Naigles 2005), and surrounding linguistic context plays a role in the sense distribution of verbs (e.g., Theakston et al. 2002).

Two native Chinese-speaking authors coded the semantics of all the target utterances in the child speech and the input independently. Both authors checked each other’s coding and any discrepancies were resolved and agreed upon on a case-by-case basis. Where utterances were ambiguous, linguistic context of the target utterances (i.e., preceding and following utterances) was used to determine the meaning in the corresponding transcripts. The intercoder agreement was high, with 98% agreement.

3. Results

We analyzed the semantic distribution of the total of 375 tokens of 打 dǎ in the child speech and the total of 809 tokens of 打 dǎ in the adult speech. All the children were found to have started to produce 打 dǎ at a young age, around the first or second data session of each individual child, between the ages of 1;06 and 2;07. The verb 打 dǎ is among the top 10 most frequent verbs in the children’s speech. In the sections below, we present the overall distribution of the different senses of 打 dǎ, followed by

² SFP = sentence final particle.

an examination of the earliest emergent uses (i.e., the first 10 tokens) and the longitudinal emergence of different senses across the nine children. Next, we focus on the comparison of the distributional patterns between the child Tong and his input.

3.1. Distribution and Emergence of Different Semantic Categories and Senses of 打

The token frequencies of 打 were calculated by the semantic representation types. Figure 1 presents the overall proportions of the different semantic categories of 打 across all nine children. The predominant type of semantic representation for all the children is Type 1 (physical action focusing on hand contact) with a mean proportion of 93% (ranging from 75% to 100%). Four out of the nine children (JC, Pan, Wu, and Xu) used only the Type 1 meaning (100%). Type 2 (hand action mostly with instrument) and Type 4 (metaphoric uses with hand action traceable) were used with a low mean proportion of 1.78% (Type 2) and 4.6% (Type 4), with the latter type being mostly produced by one child, Yang (24% of his usage), to refer to only one specific event of calling by phone. Types 3 (physical action with physical contact unspecified) and 5 (metaphoric uses with hand action untraceable) meanings were used minimally across the children with a mean proportion of 0.06% (Type 3) and 0.13% (Type 5), respectively. This suggests that the core meaning of 打 as a physical action involving hand contact is acquired early and used prevalently among the multiple senses, and metaphoric uses with hand action traceable seem to emerge before untraceable hand action. Less concrete meanings seem to be produced less and at later ages.

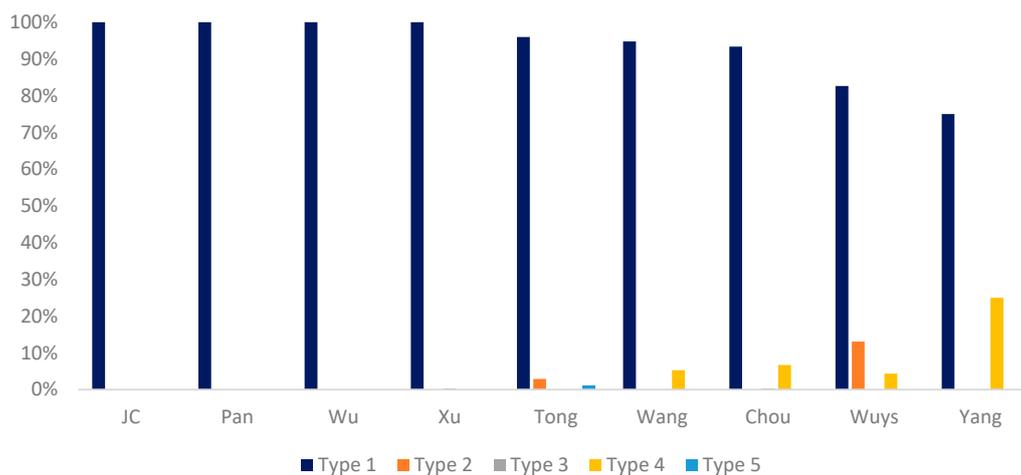


Figure 1. Proportions of different types of semantic categories of 打 by individual child. Type 1: Physical action focusing on hand contact; Type 2: hand action (mostly with instrument); Type 3: physical action with physical contact unspecified; Type 4: metaphorical uses (hand action traceable); Type 5: metaphorical uses (hand action untraceable).

We also tallied the token frequencies of 打 by different senses and calculated the overall proportions of the senses that were used more than once (i.e., a minimum of 2% usage in each child's speech). Figure 2 shows the proportions of different senses by individual child. For the ease of interpreting the distribution, color coding is used to indicate the semantic categories in which the senses belong to: various shades of blue indicate Type 1 senses; different shades of green indicate Type 2 senses, and light yellow indicates senses in Type 4. As shown in Figure 2, the majority of the different senses (75–100%) involves Type 1 physical action with hand contact ('hit', 'spank', 'fight', 'break', and 'open'), among which the sense open dominates (mean proportion 41%). Type 2 senses, hand action with instrument ('play', 'shoot', and 'hold/possess'), account for a small portion of usage (mean proportion 2.7%) and Type 4 sense, call (metaphoric uses with traceable hand action), is used by four children (mean proportion 4.5%), with the most frequent uses from one child, Yang (24%).

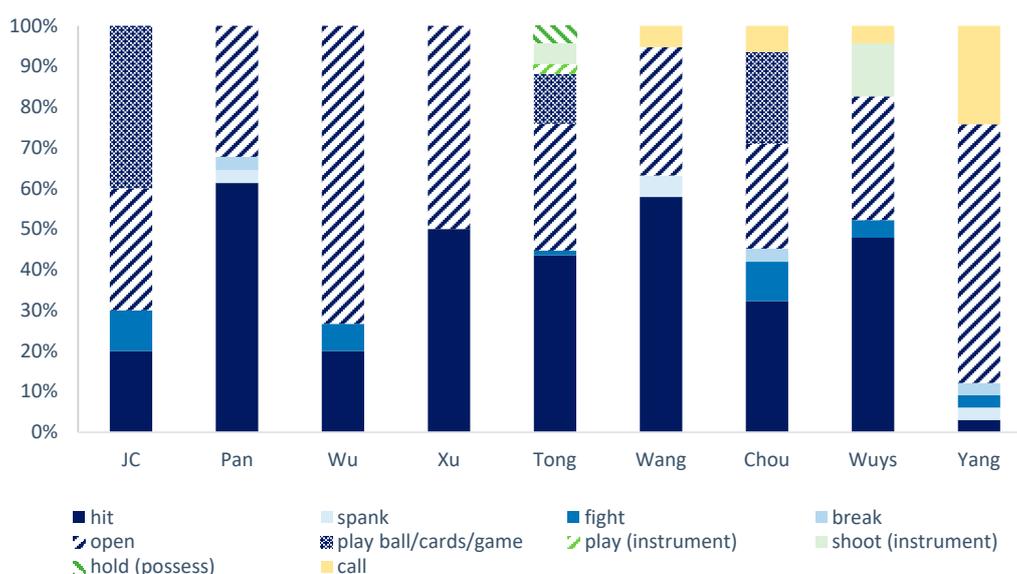


Figure 2. Proportions of different senses of 打 dǎ by individual child³.

To examine the earliest uses of 打 dǎ (emergent categories and senses), the first 10 tokens from each child were analyzed. The distribution is shown in Figure 3. Similar to the overall distribution, the most frequent senses in the earliest tokens of 打 dǎ involve physical action of hand contact, i.e., ‘open’ (40%) and ‘hit’ (31%), followed by metaphorical use of traceable hand action (i.e., ‘call’, 7.8%) and three senses (‘fight’, ‘play games’, and ‘spank’) that also involve physical action of hand contact. A further examination of the two most frequent senses ‘open’ and ‘hit’ shows that both senses occur in specific syntactic and semantic contexts: 100% of all the utterances of the ‘open’ sense occur in a resultative verb compound 打开 dǎkāi ‘hand.action-open’ with an animate agent and an inanimate patient referents, and 100% of the utterances of the ‘hit’ sense occur in a transitive sentence frame with animate agent and patient referents. (e.g., 我打妈妈 wǒ dǎ māma ‘I hit mom’). A similar pattern is observed in the sense ‘call’, which occurs 100% with an object NP 电话 diànhuà ‘phone’ and an animate agent (caller). The sense ‘fight’ occurs only in a compound verb 打架 dǎjià; and the sense ‘play games’ is also used only with inanimate objects (e.g., cards, ball) and an animate agent.

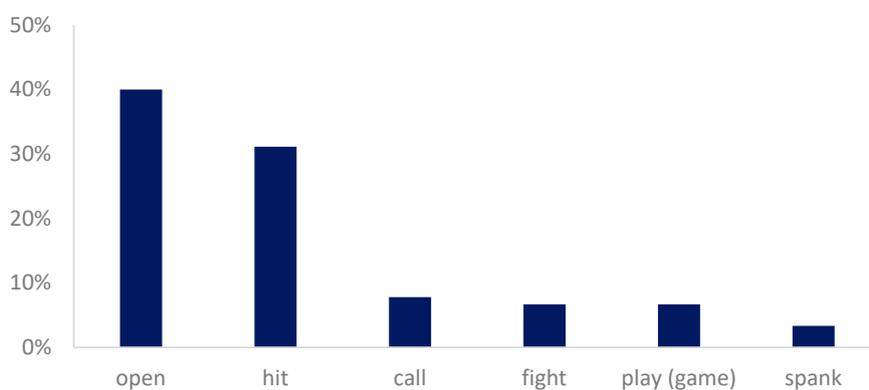


Figure 3. Proportions of the senses of 打 dǎ in the first 10 tokens by all the children.

³ Examples illustrating the different senses in the child speech: hit: 打人 dǎrén ‘hit a person’; spank: 打屁股 dǎpígǔ; open: 打开 dǎkāi; fight: 打架 dǎjià; play (ball): 打球 dǎqiú; play (game): 打牌 dǎpái; play (instrument): 打鼓 dǎgǔ ‘play the drum’; shoot: 开枪 dǎqiāng; hold: 打伞 dǎsǎn ‘hold an umbrella’; call: 打电话 dǎdiànhuà; mark: 打勾 dǎgōu ‘mark with a tick’; knot: 打结 dǎjié ‘make a knot’; thunder strike: 打雷 dǎléi; type: 打字 dǎzì ‘type’; doze: 打盹 dǎdūn ‘doze off’; sneeze: 打喷嚏 dǎpēnti ‘sneeze’; launch: 打到 dǎdào ‘launch (to)’.

3.2. Longitudinal Development of Different Senses of 打 打

We further examined the longitudinal development of different senses of 打 打. We first present the results from the speech of the child Tong, whose dataset is the largest among the nine children (cf Table 2). Tong produced 打 打 in 15 out of the 22 transcript files. Table 4 summarizes the token frequencies of the different senses by age. A total of 10 different senses show up in Tong’s speech and most involve the physical action of hand contact. The top five senses, ‘hit’, ‘open’, ‘play (games)’, ‘shoot’, and ‘launch’, account for 94% of the uses of 打 打. The earliest production of 打 打 at 1;07 includes two senses involving physical action of hand contact, ‘open’ and ‘play (games)’, suggesting that multiple senses can emerge at the same time. Three new senses, ‘hit’, ‘launch’, and ‘shoot’, emerged simultaneously at 1;11, among which ‘hit’ was the most frequent sense (37%) in Tong’s speech from 1;07 to 3;04. The sense ‘open’ is the second most frequent produced sense (30%) overall. The production of the top two senses is consistent with the general pattern observed across the other children that ‘open’ and ‘hit’ are the most dominant senses in early Mandarin-learning children’s speech. Newly emerged senses are also used with previously used senses, e.g., ‘shoot’, ‘hit’, and ‘launch’ emerged with the previously used sense *open* at 1;11. Overall, Tong used at least two or more different senses in 13 out of the 15 transcripts, indicating that the majority of his uses of 打 打 are polysemous (80%).

Table 4. Token frequencies of different senses of 打 打 in the child Tong’s speech by age.

Senses\Age	1;07	1;08	1;11	2;00	2;01	2;02	2;03	2;05	2;06	2;07	2;09	2;10	3;01	3;03	3;04	Total	prop.
hit (person/object)	0	0	7	2	4	0	6	5	3	4	1	31	0	0	1	64	37%
open	1	0	4	1	1	1	0	0	0	0	0	22	13	6	3	52	30%
play (games)	2	2	0	0	11	0	8	0	0	0	0	0	0	0	2	25	14%
shoot	0	0	4	0	0	0	0	1	1	0	0	0	0	0	4	10	6%
launch	0	0	2	0	0	1	4	0	2	0	0	0	0	0	0	9	5%
hold	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	7	4%
fight	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2	1%
play (instrument)	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2	1%
knot	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	1%
thunder strike	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1%
mark	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1%

Do other children show similar developmental trajectories? Table 5 presents longitudinally the different senses produced by Tong and the eight children in the TCMC corpus. Compared with Tong, fewer senses were produced, which is probably due to the lower frequency of data collection and fewer tokens of 打 打 in each child’s speech in the TCMC corpus (see Table 2). The early production (1;06–2;01) of 打 打 is limited to a single sense for the children in the TCMC corpus, and two senses, ‘open’ and ‘hit’, are the most frequent. The sense ‘sneeze’ appeared once for the child Xu at age of 1;06, and the sense call appeared three times in the speech of one child Yang at 1;10. Multiple senses were used simultaneously from age 2;02 and a limited number of new senses emerged gradually, including ‘play games’ (2;02), ‘break’ (2;02), ‘spank’ (2;02), ‘type’ (2;02), ‘fight’ (2;03), ‘doze’ (2;04), ‘turn on’ (2;09), ‘hold’ (2;10), ‘shoot’ (3;0), and ‘sweep’ (3;03). The most frequent senses are ‘open’ and ‘hit’ through ages 2;0–3;10. The developmental pattern in the eight children from the TCMC corpus is thus similar to that in the speech of Tong. Individual difference is also observed in the specific senses produced, which could tie to the particular contexts of the speech produced (e.g., ‘sneeze’, ‘call’, or ‘type’).

To summarize, the earliest production of 打 *dǎ* emerges around age 1;06. A small set of different senses of 打 *dǎ* are produced by the children between 1;06 and 3;10, centered around the core meaning of concrete physical action with hand contact. Metaphorical senses are overall very infrequent in terms of both types and tokens. Multiple senses emerge and are often used simultaneously with the senses produced earlier. The small set of senses of 打 *dǎ* typically involves specific contexts that are frequent in a child's daily interactions and bodily experience (e.g., calling, playing games, fighting, sneezing), which usually occur in specific syntactic and semantic contexts that are inherent to the meaning of the specific senses (e.g., an open event typically involves an animate agent and an inanimate patient).

3.3. Comparison between Tong and His Input

To explore the role of input on the semantic development, we compared the production of 打 *dǎ* in the speech of the child Tong and his input. The Tong corpus is selected because of the higher tokens of 打 *dǎ* in both the child and the adults' speech due to a relatively longer period of data collection (cf. Table 2). As shown in Figure 4, Tong is similar to his caregivers in the overall distributions of the semantic representations of 打 *dǎ*. He used 打 *dǎ* as meaning physical action with hand contact (Type 1) most frequently and 95% of his 打 *dǎ* belong to this semantic category, suggesting that the core meaning of hand action is a prototypical usage. In addition to the dominant Type 1, Tong's caregivers show a wider range of semantic categories, including metaphorical uses of traceable hand action (Type 4, 10%) and hand action with instrument (Type 2, 7%). Physical action with unspecified physical contact (Type 3, 2.8%) and metaphorical uses with untraceable hand action (Type 5, 2%) also show minimal uses.

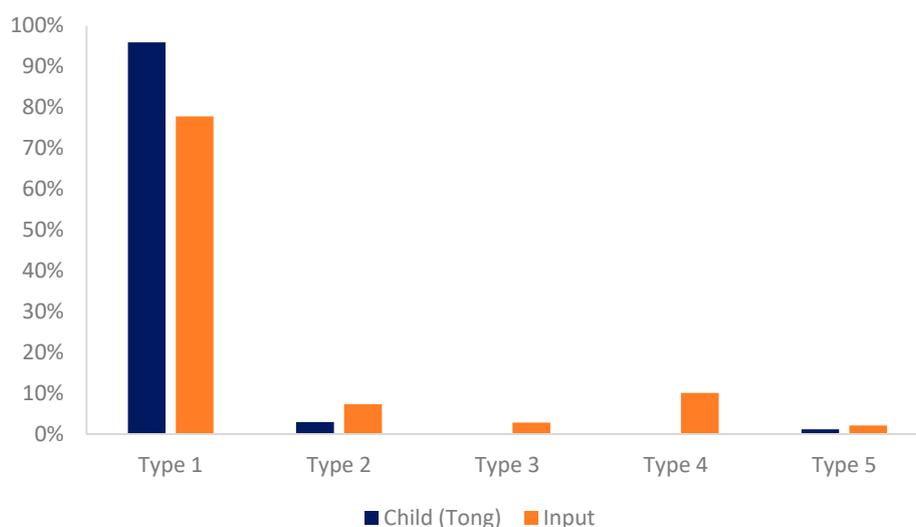


Figure 4. Proportions of the semantic representations of 打 *dǎ* in Tong and his input. Type 1: Physical action focusing on hand contact; Type 2: hand action (mostly with instrument); Type 3: physical action with physical contact unspecified; Type 4: metaphorical uses (hand action traceable); Type 5: metaphorical uses (hand action untraceable).

We further compared the distributions of different senses of 打 *dǎ*. As shown in Figure 5, both Tong and his caregivers produced a variety of different senses. The most frequent sense in Tong's speech is 'hit' (37%), followed by 'open' (30%), 'play (games)' (14%), and 'shoot' (6%), whereas in the caregivers' speech, those senses are also frequent but more evenly distributed, 'open' (17%), 'hit' (17%), 'shoot' (15%), and 'play (games)' (14%). The senses 'launch' and 'hold/possess' show similar proportions of usage in both Tong and his caregivers' speech (4–5%). Overall Tong's use of the different senses of 打 *dǎ* reflects the distributional pattern in his caregivers' speech.

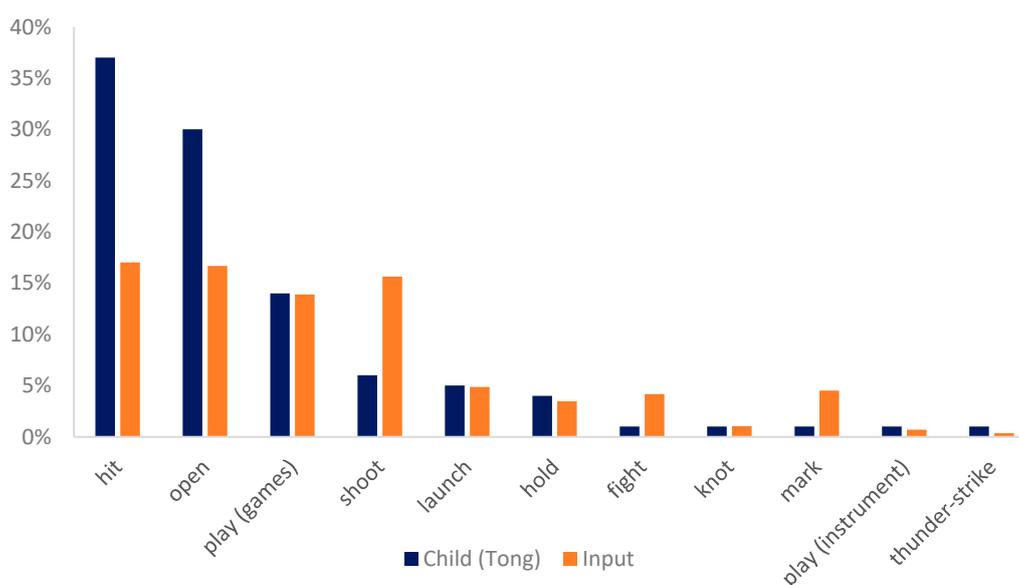


Figure 5. Proportions of different senses of 打 *dǎ* in Tong and his input.

To investigate if the development of the multiple senses of 打 *dǎ* is closely tied to certain syntactic contexts, we analyzed the syntactic frames or constructions that 打 *dǎ* occurs in the speech of Tong and his caregivers. Figure 6 presents the proportions of different syntactic frames. A total of 30 different syntactic frames were identified in Tong's use of 打 *dǎ*. As shown in Figure 6, the most frequent type of syntactic frame is 打 *dǎ* used with another verb in the form of a verb compound as the predicate of a sentence (50%), followed by transitive frames with an overt object NP (23.84%) and with overt subject and object NPs (9.88%), and locative construction (VPP, 5.23%). Within the sentences containing a verb compound with 打 *dǎ*, a variety of syntactic constructions were produced. The most frequent type is the basic transitive frame (38%), followed by negative constructions (30%), modal verb constructions (12%), *ba* constructions (7%), and bare verb compound (7%).⁴ Despite the varied syntactic frames, the verb compounds with 打 *dǎ* show very limited types—only four different verb compounds were produced, among which 打开 *dǎ-kāi* 'hand.action-open' (open) and 打死 *dǎ-sǐ* 'hit-be.dead' (kill) account for 93% of the compound predicates. The adults used a wider range of syntactic frames with 打 *dǎ* than the child, i.e., a total of 50 different syntactic frames. As shown in Figure 6, the most frequent syntactic frame is 打 *dǎ* in the form of a verb compound as the predicate of a sentence (28%), followed by transitive frames with an overt object NP (25.17%) and with overt subject and object NPs (15.52%), and negative construction (13.1%). The adults and the child Tong are thus similar in the overall distribution of the syntactic frames, and verb compounds with 打 *dǎ* and the transitive frames account for the majority of the sentence forms.

⁴ Examples of the syntactic constructions are shown below. Basic transitive frame: e.g., 打死了七个蚊子 *dǎsǐ le qī ge wénzi* 'kill LE seven mosquitos' (I killed seven mosquitos); negative construction: e.g., 没打开 *méi dǎ kāi* 'not hand.action-open' (it did not open); modal verb constructions: e.g., 你需要打开 *nǐ xūyào dǎkāi* 'you need open' (you need to open this); *ba* constructions: 把那打开 *bǎ nà dǎ-kāi* 'ba that hand.action-open' (open that); bare verb compound: 打死 *dǎ-sǐ* 'hit-be.dead' (kill).

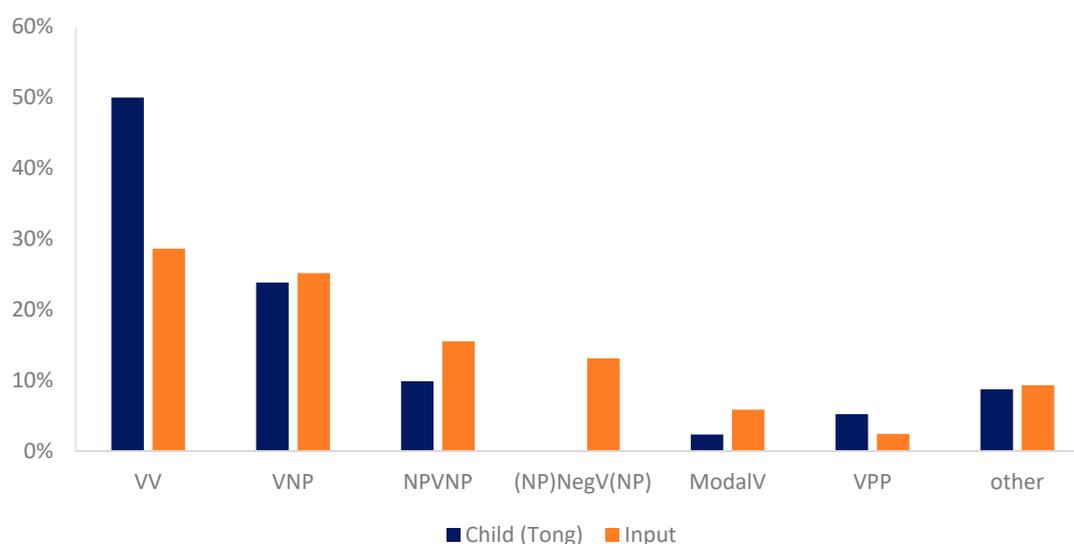


Figure 6. Proportions of syntactic frames of 打 *dǎ* in Tong's and his caregiver's speech. (VV = verb compound; VNP = verb followed by an object noun phrase; NPVNP = subject noun phrase followed by the verb and an object noun phrase; VPP = verb followed by a prepositional phrase)⁵.

4. Discussion

This study investigates a fundamental issue in child language acquisition, i.e., how do children figure out the mapping between the forms and the meanings in the ambient language, in the case of polysemy. When multiple meanings are available for the same form, how is the form-meaning mapping correctly established? We focus on the semantic development of the polysemous verb 打 *dǎ* based on the analyses of longitudinal corpus data from nine Mandarin-learning children from the age of 1;05 to 3;10. The verb 打 *dǎ* is highly polysemous in Mandarin and it is also among the first verbs used by all nine young Mandarin children.

We examined the developmental trajectory (i.e., emergent order) of different senses of 打 *dǎ*. The results show that young Mandarin-learning children produce a small set of different senses of 打 *dǎ* between 1;05 and 3;10, centered around the core meaning of concrete physical action with hand contact (e.g., 'hit', 'open', and 'play games'), and metaphorical senses are overall very infrequent in terms of both types and tokens. Physical actions through an instrument or with unspecified physical contact or metaphorical uses with untraceable hand action are infrequent. This emergent order of concrete before abstract or metaphorical meanings is congruent with findings in prior research that the most prototypical meaning tends to be acquired before the more metaphorically and metonymically motivated meanings (e.g., Nerlich et al. 2003). Booth and Hall (1995) investigated children's (3-, 6-, 9-, and 12-year-olds) understanding of the polysemous cognitive verb *know* and found that children's development of the verb meanings showed an effect of the abstractness and conceptual difficulty hierarchy—low levels of meaning (e.g., perception, recognition, recall, understanding) developed earlier and faster than high levels of meaning (e.g., metacognition, evaluation). Concrete and conceptually less difficulty senses therefore emerge earlier than more abstract and conceptually different senses. In the case of the acquisition of the Mandarin verb 打 *dǎ*, we could also see the effect of the general abstractness and conceptual difficulty hierarchy.

⁵ Examples of the syntactic frames are shown below. VV: 打死 *dǎ-sǐ* 'hit-be.dead' (kill); VNP: 打怪物 *dǎ guàiwù* '(I) shot the monster'; NPVNP: 我打妈妈 *wǒ dǎ māma* 'I hit mom'; VPP: 打到上面了 *dǎ dào shàng miàn le* 'launch to the top'.

Furthermore, within the concrete physical action with hand contact, three different senses, ‘hit’, ‘open’, and ‘play (games)’, are the most frequent, suggesting the influence of typical hand actions such as hitting, opening, or playing games in young children’s daily events. This pattern corroborates prior findings that giving physical punishment emerges as one of the earliest senses of 打打 *dǎ* in monolingual children (1;9–2;3) (Gao 2015) and bilingual Mandarin children (Sak and Gao 2016). The emergence of the concrete core senses of 打打 *dǎ* is thus in line with the cognitive linguistics approaches that regard meaning as perceptually grounded—the primary experiential scene in which 打打 *dǎ* is embedded in events where somebody uses their hand to act on something, and from this experience derives many senses originated from hand actions such as playing games, defeating someone, opening, fastening, construction, covering, possession, marking, launching, insertion, collection, removal, and working.

How do children acquire the different senses of 打打 *dǎ*? The developmental trajectory supports the continuous derivational approach (Zhang et al. 2010) and the restricted monosemy approach (Kidd and Cameron-Faulkner 2008)—children initially extract a core feature of 打打 *dǎ*, i.e., volitional physical action focusing on hand contact, but only use it in a restricted way. The knowledge of the core feature of 打打 *dǎ* is revealed in children’s simultaneous production of different senses (e.g., ‘open’, ‘hit’, ‘spank’, ‘fight’) in the first uses. The limited productivity is reflected in the small number of related concrete senses of 打打 *dǎ* and the slow increase in the sense types across the sampled ages, 1;06 to 3;10. The dominant senses tend to center around events involving concrete hand motion (‘hit’ and ‘open’). New senses do not emerge in large numbers and are closely tied to the immediate contexts of the children’s social interactions. The limited productivity is further seen in the limited set of syntactic frames that 打打 *dǎ* occurs in, including verb compounds and transitive sentences.

How do children go beyond the limited productivity? Children are likely to be able to recognize the multiple senses (mostly Type 1 senses) as being related in a network model consciously or subconsciously due to the concrete nature of the hand action feature in these related senses. The more abstract metaphorically derived senses (e.g., Type 5 senses such as 打算 *dǎsuàn* ‘plan’), on the other hand, may be acquired in an independent approach (cf. Zhang et al. 2010) due to the less visible derivational relations to the concrete core sense, and may be treated as unrelated to the core meaning at the beginning. Even in adult language, some meanings may be so different from the core that they may be stored and learned separately (cf. Theakston et al. 2002). Theakston et al. (2002) analyzed the acquisition of the highly versatile English verbs “go”, and found that children acquire different constructions in different contexts without evidence that these uses are initially related. It is possible to explain these results in the sense that the various usage patterns of these multifunctional verbs are not linked to one another, but initially represent different syntactic frames. Our current data only contain early speech of Mandarin-learning children (1;05–3;10) and show limited uses of abstract metaphorical senses. Future research should explore whether abstract metaphorical senses are acquired in an independent approach and how they are integrated with the basic core senses to build a complete semantic network with data from older children.

How does input, including syntactic, semantic, and contextual cues, contribute to the acquisition of senses of 打打 *dǎ*? Our results from the comparison of the child Tong and his input show that Tong’s uses of 打打 *dǎ* reflect the semantic and syntactic distributional patterns in the input, as both Tong and his caregivers use multiple senses of 打打 *dǎ* from the beginning and both use the physical action involving hand contact most frequently. The most frequently produced senses are also used similarly, even though Tong tends to use the prototypical senses dominantly and the adults’ usage is more evenly distributed across the frequent senses. The result provides further support for the usage-based learning that the most frequent senses tend to be the earliest senses in child speech and children’s knowledge of the multiple meanings correlates with parental uses (Adricula and Pielke 2019; Booth et al. 1997; Theakston et al. 2002). Computational modeling, utilizing actual child-directed data, show that distributional models are sufficient to reasonably distinguish verb senses, but further information is needed to better predict children’s learning (Parisien and Stevenson 2009).

5. Conclusions

To conclude, we could summarize the acquisition of the polysemy of 打 *dǎ* as the following. Children start with the concrete prototypical meaning of 打 *dǎ* as a physical action verb that bears the semantic features [+hand, +volition, +contact, +force] and [-instrument] and gradually extend to physical actions that may involve an instrument ([+instrument]), or without hand contact ([-contact]), and finally to metaphorical senses that share the physical action and volition features with the prototypical sense of 打 *dǎ*, but lack the contact or instrument features ([-contact, -instrument]). Traceable hand action senses may be learned before the more abstract senses with untraceable hand action. The conceptual difficulty in terms of levels of abstractness of the meanings plays a role in children's emergent order of different senses of 打 *dǎ*. The development trajectory also supports the restricted monosemy approach in the early acquisition of polysemy and future research needs to examine further how children develop the full range of the diverse meanings of 打 *dǎ* beyond age 3;10. Furthermore, the inherent lexical semantics of 打 *dǎ* is not the only factor that determines the acquisition. The learning process is also modulated by the distributional patterns of 打 *dǎ* in the input, i.e., the semantic and the syntactic contexts of 打 *dǎ*, which suggests that statistical learning may play a role in the acquisition of different senses (e.g., Adricula and Pielke 2019; Kidd 2012). Note that the order and the derivational route of acquisition of different senses of polysemous verbs may vary, subject to usage-based factors such as input distributional patterns and functional saliency and needs (cf. Zhang et al. 2010). Future research should examine other highly frequent polysemous verbs in Mandarin child speech to map out similarities and differences in development and provide a comprehensive account for early semantic development.

Our research is the first longitudinal study of the emergence and development of the polysemous verb 打 *dǎ* in the speech of monolingual Mandarin-learning children and has pedagogical implications for the teaching and learning of Mandarin as a second/foreign language (L2). Zhang et al. (2011) showed that even intermediate and advanced L2 learners of Mandarin had not acquired the wide range of meanings of the verb 打 *dǎ* in a meaning matching (comprehension) task and in their analysis of interlanguage writing (written production) corpus data. They found that the dominant senses that the learners showed better knowledge of centered around a very limited number of senses, 'hit' and 'play (games)' in both the comprehension and production data. Following the development trajectory of child Mandarin learners, L2 learners of Mandarin should be exposed to a small set of senses derived from the concrete core/prototypical senses of 打 *dǎ* before the more abstract senses that are derived metaphorically. Functionally salient and frequent senses should also be introduced before less frequently used senses. Recent studies of L2 acquisition of Mandarin polysemous verbs also revealed that prototypicality (e.g., prototypical senses) significantly predicts the learning outcomes (e.g., Liang 2014). Furthermore, large number of exemplars of different senses should be provided to facilitate the inference of multiple senses of 打 *dǎ* in communicative contexts (similar to what a child learner experiences in naturalistic language learning situations). Explicit instructions on the semantic features of 打 *dǎ* and the derivational relations may also promote the understanding, generalization, and productive uses of the complicated meanings of 打 *dǎ*. Future research should be conducted to explore the appropriate methods and assessments of these pedagogical applications.

Author Contributions: Conceptualization, J.C. and X.W.; methodology, J.C. and X.W.; data analysis, J.C. and X.W.; writing—original draft preparation, J.C.; writing—review and editing, J.C. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsor had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Adricula, Norielle, and Megan Pielke. 2019. A child's acquisition of polysemy: Of, with, and by in child English. In *Proceedings of the 43rd Boston University Conference on Language Development*. Edited by Megan Brown and Brady Dailey. Somerville: Cascadilla Press, pp. 27–41.
- Allwood, Jens. 2003. Meaning potentials and context: Some consequences for the analysis of variation in meaning. In *Cognitive Approaches to Lexical Semantics*. Edited by Hubert Cuyckens, René Dirven and John R. Taylor. Berlin: Mouton de Gruyter, pp. 29–66.
- Booth, James R., and William S. Hall. 1995. Development of the understanding of the polysemous meanings of the mental-state verb know. *Cognitive Development* 10: 529–49. [[CrossRef](#)]
- Booth, James R., William S. Hall, Gregory C. Robison, and Su Yeong Kim. 1997. Acquisition of the Mental State Verb Know by 2- to 5-Year-Old Children. *Journal of Psycholinguistic Research* 26: 581–603. [[CrossRef](#)] [[PubMed](#)]
- Clark, Eve V. 1993. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark, Eve V. 1996. Early verbs, event-types, and inflections. In *Children's Language*. Edited by C. E. Johnson and J. H. V. Gilbert. Hillsdale: Lawrence Erlbaum, pp. 61–73.
- Deng, Xiangjun, and Virginia Yip. 2018. A multimedia corpus of child Mandarin: The Tong corpus. *Journal of Chinese Linguistics* 46: 69–92.
- Evans, Vyvyan. 2005. The meaning of time: polysemy, the lexicon and conceptual structure. *Journal of Linguistics* 41: 33–75. [[CrossRef](#)]
- Gao, Hong. 2001. The Physical Foundation of the Patterning of Physical Action Verbs. Ph.D. dissertation, Lund University Press, Sweden.
- Gao, Hong. 2015. Children's early production of physical action verbs in Chinese. In *The Oxford Handbook of Chinese Linguistics*. Edited by William Shi-Yuan Wang and Chaofen Sun. Oxford: Oxford University Press, pp. 654–65.
- Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition* 1: 3–55. [[CrossRef](#)]
- Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Huang, Chu-Ren, Shu-Kai Hsieh, Jia-Fei Hong, Yun-Zhu Chen, I-Li Su, Yong-Xiang Chen, and Sheng-Wei Huang. 2010. Chinese wordnet: Design and implementation of a cross-lingual knowledge processing infrastructure. *Journal of Chinese Information Processing* 24: 14–23.
- Kidd, Evan. 2012. Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics* 33: 393–418. [[CrossRef](#)]
- Kidd, Evan, and Thea Cameron-Faulkner. 2008. The acquisition of the multiple senses of with. *Linguistics* 46: 33–61. [[CrossRef](#)]
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Lakoff, George, and Mark John. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lee, Joanne N., and Letitia R. Naigles. 2005. The input to verb learning in Mandarin Chinese: A role for syntactic bootstrapping. *Developmental Psychology* 41: 529–40. [[CrossRef](#)]
- Liang, Haiyan. 2014. Factors accounting for acquisition of polysemous shàng 'to go up'-phrases in Chinese as a second language (CSL). *Chinese as a Second Language Research* 3: 201–25. [[CrossRef](#)]
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah: Lawrence Erlbaum.
- McKercher, David A. 2001. The Polysemy of with in First Language Acquisition. Ph.D. dissertation, Stanford University, Stanford, CA, USA.
- Nerlich, Brigitte, Zazie Todd, and David D. Clarke. 2003. Emerging patterns and evolving polysemies: The acquisition of get between four and ten years. In *Polysemy: Flexible Patterns of Meaning in Mind and Language*. Edited by Brigitte Nerlich, Zazie Todd, Vimala Herman and David D. Clarke. Berlin: Mouton de Gruyter, pp. 333–57.

- Parisien, Christopher, and Suzanne Stevenson. 2009. Modelling the acquisition of verb polysemy in children. In *Proceedings of the CogSci2009 Workshop on Distributional Semantics beyond Concrete Concepts*. Edited by Niels Taatgen and Hedderik van Rijn. Austin: Cognitive Science Society, pp. 17–22.
- Sak, Hui Er, and Helena Hong Gao. 2016. The Polysemy of the Chinese Action Verb “Dǎ” and Its Implications in Child Language Acquisition. In *Chinese Lexical Semantics: CLSW 2016*. Edited by Minghui Dong, Jingxia Lin and Xuri Tang. Cham: Springer, pp. 524–33.
- Sandra, Dominiek, and Sally Rice. 1995. Network analyses of prepositional meaning: Mirroring whose mind—The linguist’s or the language user’s? *Cognitive Linguistics* 6: 89–130. [[CrossRef](#)]
- 中国社会科学院词典编辑室 (Chinese Academy of Social Science Institute of Linguistics), ed. 2016. *The Contemporary Dictionary of Chinese*, 7th ed. (现代汉语词典 (第7版)). Beijing (北京): The Commercial Press (商务印书馆).
- Theakston, Anna L., Elena V. M. Lieven, Julian M. Pine, and Caroline F. Rowland. 2002. Going, going, gone: The acquisition of the verb ‘go’. *Journal of Child Language* 29: 783–811. [[CrossRef](#)] [[PubMed](#)]
- Tomasello, Michael. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.
- Tuggy, David. 1993. Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4: 273–90. [[CrossRef](#)]
- Vicente, Agustín, and Ingrid L. Falkum. 2017. Polysemy. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- Zhang, Yunqiu, Jianshe Zhou, and Jing Fu. 2010. The Acquisition Strategies of Polysemous Verbs in Early Mandarin-Learning Child Speech: A Case Study of a Child Learning the Beijing Mandarin (早期汉语儿童多义词的习得策略——一个北京话儿童的个案研究). *Studies of the Chinese Language* (中国语文) 334: 34–43.
- Zhang, Jiangli, Dehong Meng, and Weihong Liu. 2011. The Depth of Monosyllabic Polyseme Acquisition of Learners Who Speak Chinese as the Second Language: A Case Study of the Verb Dǎ (汉语第二语言学习者单音多义词习得深度研究——动词“打”为例). *Applied Linguistics* (语言文字应用) 2: 112–21.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).