

ARTICLE

Speech representation used by Mandarin Chinese-speaking children aged three to six years

Wei GUAN¹  and Haitao LIU^{2,3,*}

¹School of Foreign Languages, Zhejiang University of Finance and Economics, China

²Zhejiang University, China

³Guangdong University of Foreign Studies, China

*Corresponding author: Department of Linguistics, Zhejiang University, Hangzhou, China; and Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China. Email: htliu@163.com

(Received 05 August 2020; revised 16 November 2021; accepted 03 December 2021)

Abstract

This study investigates how Mandarin Chinese-speaking children use Mandarin Chinese, a language lacking tense markers, to represent characters' speech in their story narratives. Eighty participants, from three to six years of age, completed an elicited narrative task based on a wordless picture book. The representing forms and signals that they employed in representing activities were assessed. The results showed a significant age-related increase in the overall use of speech representation by the cohort of children. Regarding representing forms, direct representation exhibited its expected dominance among all age groups, and its presence in children's narratives grew significantly with age. Concerning representing signals, *shuō* 'say' was the most prevalent across all ages and susceptible to significant age effects, and, with advancing age, children's representing signals expanded from single verbs to more complex syntactic constructions. In addition, no significant gender effects were observed regarding the representing forms or signals that Mandarin-speaking children used.

Keywords: speech representation; Mandarin Chinese-speaking children; forms of representation; representing signals

Introduction

Speech representation (SR), or reported speech, normally refers to the recurrent tendency of speakers to incorporate the utterances that they have ascribed to other speakers in either the real world or the figurative world into their communication. In the following example, the narrator – a Mandarin Chinese-speaking child in our study – represented how the boy in the storybook asked the waiter to return his pet frog.

- (1) Ránhòu tā shuō zhè shì wǒ de qīngwā.
then he say this be I Nominalizer frog
‘Then he said, “This is my frog.”’

SR usually requires the introducer of the representation to precede the represented discourse. The introducer, commonly presented within the frame of 'represented speaker + speech verb' (in Example [1], *tā shuō* 'he said'), marks an introduction to the represented discourse (in Example [1], *zhè shì wǒ de qīngwā* 'This is my frog').

Children's use of SR

SR is a demanding language skill for children since it involves a complex integration of various domains, such as mental, pragmatic, and syntactic abilities. In narratives, for example, representing characters' speech requires a young narrator to encode information about the mental state of the characters (e.g., their behavior and intentions) through an ability known as theory of mind (Astington & Jenkins, 1995; Premack & Woodruff, 1978), and to shift between the perspectives of the narrator and the characters. Through representing speech, the narrator attributes a particular intentional state to the characters (Bamberg & Damrad-Frye, 1991). Therefore, representing character speech has been widely adopted as an evaluative device for analyzing the narrative abilities of typically developing (TD) children, generally demonstrating an age-related increase (Chang, 2000, 2003; Chang & Huang, 2016; Drijbooms, Groen & Verhoeven, 2017; Ukrainetz, Justice, Kaderavek, Eisenberg, Gillam & Harm, 2005; Zhang, McCabe, Ye, Wang & Li, 2018). It is also an effective index of the narrative competence of children with language disorders who are less likely to incorporate character speech than their TD peers (Hemphill, Picardi & Tager-Flusberg, 1991; Hemphill, Uccelli, Winner, Chang & Bellinger, 2002; Manhardt & Rescorla, 2002).

SR also involves children's pragmatic knowledge (Hemphill *et al.*, 1991) in comprehending their listeners' needs and understanding their responsibility as speakers to communicate unambiguously, for instance, by explicitly differentiating various perspectives when they shift. Representing introducers are perspective-shift markers that inform listeners of "the occurrence of an adjacent representation" (Güldemann, 2012, p. 118) from the characters and consequently facilitate the listener's correct interpretation of a represented discourse. Previous studies have found that children's tendency to represent speech without introducers (i.e., to use unframed direct representation) decreases with age (Hickmann, 1993; Özyürek, 1996), thus indicating children's increasing awareness of their listeners' need for understanding.

Another form of knowledge that is essential to representation is children's competence in making "syntactic adjustments" in SR (e.g., indexical, tense, and word order shifts) (Goodell & Sachs, 1992, p. 397). The representing forms that heavily rely on these syntactic abilities have been the focus of much research on children's use of SR (e.g., Ely & McCabe, 1993; Goodell & Sachs, 1992; Hickmann, 1993; Köder, 2016; Nordqvist, 2001a; Özyürek, 1996). Many studies have shown that direct representation, or direct speech (either with or without introducers), is the preferred form of representation for children of different ages and languages. A similar trend was found in Ely and McCabe's (1993) research on English-speaking children (aged four to nine), in Hickmann's (1993) study of English-speaking children (aged four to ten), in Özyürek's (1996) research on Turkish-speaking children (aged five to 13), in Köder's (2016) study of German and Dutch subjects (aged two to four-and-a-half years), and in Zhang *et al.*'s (2018) research on Mandarin-speaking children (aged four to six).

As opposed to the dominance of direct speech, the SR that English subjects used in Goodell and Sachs's study (1992) depicted a U-shape among the three age groups.

Specifically, those aged four and eight most frequently used indirect representation, or indirect speech, and only six-year-olds preferred direct speech. Nordqvist (2001a) found that only 15-year-old Swedish children preferred the indirect form when compared to the other six age groups. Such inconsistent findings regarding the dominant form across previous studies might be attributed to two factors: the ways in which narratives were obtained and the children's ages.

Given that the production of direct speech is easier than that of indirect speech (which will be further explored in the Discussion section), it can be expected that children manifest general preferences for direct representation among various narrative types, including personal narratives (Ely & McCabe, 1993), picture book narratives (Nordqvist, 2001a), caregiver-child spontaneous speech (Köder, 2016), and story retelling (Hickmann, 1993; Özyürek, 1996). However, Goodell and Sachs (1992) adopted a different eliciting process in which the subjects in one of the story-retelling tasks were exposed to a story that contained indirect speech only. Most groups' apparent preference for the indirect form was corroborated by the findings of Serratrice, Hesketh, and Ashworth (2015), in that children could be successfully primed to use more indirect speech than those without training. Another explanation might be that Nordqvist's (2001a) indirect speech users were older than most of the ones using direct speech. Nordqvist explained that these (pre)adolescent narrators had adopted indirect speech as a strategic choice, particularly when they felt little motivation to complete the task of narrating a children's story.

Investigations into representing forms have led to an interest in the introducer, also called the quotative index (Güldemann, 2012). As a metalinguistic frame (Hickmann, 1993), the quotative index helps a representor or narrator in narratives to reorient listeners from the discourse in the immediate narrative situation to that in the narrated situation, thereby serving "a discourse organizational function" (Hasund, Opsahl & Svennevig, 2012, p. 38). Apparently, children's use of introducers, which help listeners recognize incoming represented speech, indicates that their syntactic ability in organizing discourse and their pragmatic awareness of listeners' needs are not mutually exclusive. As demonstrated in certain studies, older children (aged seven to 13 years) more consistently include quotative indexes than younger children (between four and five years of age), who tend to omit these indexes (Hickmann, 1993; Özyürek, 1996). Some researchers have attributed this tendency to older children's growing sensitivity to their listeners' needs and the related communication requirements (Goodell & Sachs, 1992; Nordqvist, 2001a; Özyürek, 1996).

Apart from organizing discourse, a representor employs various lexical signals, including both traditional speech verbs and novel quotative indexes such as "like" in "I am like, No!", and "go" in "And she went, 'No you're not!'" (both cited from Levey, 2003, pp.311, 315), which mark a speech act to express the representor's interpretation of the represented content (Levey, 2003; Özyürek, 1996; Spronck, 2012). Hence, these signals serve "subjective and interpersonal functions" (Hasund *et al.*, 2012, p. 38). Furthermore, these representing signals – identified as verbs of saying in most studies – provide a good place for researchers to look in order to find evidence of children's developing representing skills. Goodell and Sachs (1992) have revealed that the prevalent usage of the generic verb *say* diminishes as age increases, while the occurrences of other generic verbs (e.g., *tell* and *ask*) and non-generic verbs (e.g., *complain* and *beg*) gradually increase. Özyürek (1996) disclosed a similar pattern in his work on Turkish children. Simultaneously, discrepancies could also be identified among these studies. The five-year-olds in Özyürek's research preferred only generic verbs, whereas both the four- and six-year-olds

investigated by Goodell and Sachs (1992) included both generic and non-generic verbs in their representations. The differences in the retelling materials to which the children had been exposed could partly account for such contrasting performances, with Goodell and Sachs's materials containing various generic and non-generic speech verbs and Özyürek's containing no such verbs. However, it is notable that these studies focused almost exclusively on children's use of single speech verbs. Nordqvist (2001a, p.243) found that, unlike children, adults added adverbial modifiers to verbs, with more detailed descriptions of the represented speech events (e.g., *frågade hunden surt*, 'the dog asked sourly'). So far, little is known about children's use of alternative signals other than single verbs.

The gender effect on children's SR has not been assessed by many researchers, and even fewer significant differences have been identified thus far. One exception is the work of Ely and McCabe (1993), which found that girls (four- to nine-year-olds) in personal narratives significantly outperformed boys, not only in representing forms (i.e., direct speech and narrativized speech) but also in speech verbs (i.e., *go* and non-generic verbs). In a later study on children (two- to five-year-olds) in dinnertime conversations, Ely, Gleason, Narasimhan, and McCabe (1995) only tested the gender variation in the children's total frequency of representation because of the rather few occurrences. They found that girls' standardized representation rate was approximately double that of boys, albeit with no statistical significance. Besides the different age ranges that might partly account for the inconsistent findings across the two studies, another potential reason was that representations elicited by the dinnertime conversations were fewer than those elicited by the prompted personal narratives. Moreover, Levey (2003) compared the representing signals produced by boys and girls (aged 10 and 11) and found that girls demonstrated a less frequent use of *say* and a greater use of the relatively new quotative index *go* than did boys, which indicated that girls exercise a more involved conversational style.

Most developmental studies on SR have mainly been concerned with children who speak English and other European languages (e.g., Swedish, German, Turkish, and Dutch). They have included children of various age ranges, but few have examined SR among young children aged, for example, from three to six. Previous findings have suggested that children's narrative competence develops rapidly in their preliterate years before they enter school (Berman & Slobin, 1994; Chang, 2000; Curenton, 2011; Lai, Lee & Lee, 2010). Zhang *et al.* (2018) have added corroborative evidence, documenting that SR develops significantly in Mandarin-speaking children from the ages of three to six. However, in their study, as well as in others on Mandarin SR in children's speech (Chang, 2000, 2003; Chang & Huang, 2016), SR has been adopted as one of the narrative components assessed and strictly confined to direct and indirect speech forms.

SR in Mandarin Chinese

Many studies have been devoted to Mandarin SR, particularly its forms, across various genres. Drawing on Leech and Short's (1981) influential SR model, which was developed for English novels, some analysts have examined Mandarin Chinese materials, including literary works (Tang, 2005; Zhao, 1987), news reports (Gao, 2013), and courtroom trials (Luo, 2013). They have observed a range of representing forms in Mandarin Chinese such as direct speech, indirect speech, free direct speech (i.e., unframed direct form), free indirect speech (i.e., unframed indirect speech), and narrative report of speech acts (NRSA). Classified as 'voice representation' in our analysis, an NRSA typically contains

one clause with a speech verb followed by a noun phrase. It gives a minimal account of a speech act, without specifying the details of, for example, what words were uttered in the speech event (e.g., *Tā jièshào le jingjì fāzhǎn qíngkuàng* ‘He introduced the economic development’).

While Mandarin SR has similar forms to English classifications, its unique characteristics are also noticeable. In an analysis of Chinese literary texts, Shen (1991) has noted that Mandarin Chinese does not have tense markers or complementizers (e.g., *that* in English), and that personal pronouns are usually the only markers of the direct-indirect distinction. She has further proposed, in the cases where these pronouns are missing, the ambiguous form arises. Consider example (2) from our data, in which the six-year-old narrator dropped the subject in the represented content.

- (2) xiǎo nánhái shuō nǎ gè dōu xiǎng yào.
 little boy say which Classifier all want have
 ‘The little boy said, (I/he) want(ed) everything.’

Unlike unambiguously direct or indirect speech in Mandarin, example (2) misses the diagnostic signal (i.e., the dropped personal pronoun in this case) for the direct-indirect distinction. Missing pronoun, or “zero pronoun”, is a pronounced feature of Mandarin, which means that pronoun that is understood in context tends to be left unspecified (Li & Thompson, 1981, p. 658). Due to the zero pronoun, example (2) becomes ambiguous; it could be interpreted as either direct speech (with the addition of the first-person pronoun) or indirect speech (with the addition of the third-person pronoun).

Similarly, the limited syntactic constraints on direct and indirect speech contribute to another unique grammatical form in Mandarin SR: namely, the mixed form. Having been found in classical written Chinese, modern spoken and written Mandarin, this form combines the features of both direct and indirect speech, which, in most cases, are indexical expressions such as pronouns (Dong, 2008). See example (3) from our data, in which the six-year-old narrator demonstrated how the character asked the waiter to return his pet frog.

- (3) ránhòu tā shuō, zhè shì wǒ de
 then he say this be I Nominalizer
 qīngwā, néng bù néng huán gěi tā.
 frog can not can return give he
 ‘Then he said: “This is my frog”, and that if it was possible to give (it) back to him.’

Rather than direct and indirect forms that reflect the respective perspectives of the character and the narrator, example (3) mixes the two points of view. The first represented clause resembles a direct form, as the demonstrative *zhè* ‘this’ and the pronoun *wǒ* (‘my’) mark the character’s perspective; the second clause represents indirect speech, as *tā* (‘him’) indicates the narrator’s viewpoint.

In addition to the classification of forms, there has been focus in Mandarin SR form studies on the direct-indirect distinction. Personal pronouns (e.g., *wǒ* ‘I’, *tā* ‘he’) are almost unanimously regarded as the main, or even the only, diagnostic criterion (Hagenaar, 1996; Huang, 2009; Li, 1986; Luo, 2013; Shen, 1991; Yue, 2011). Other indexical features constitute another domain for distinction, including demonstratives (e.g., *zhè* ‘this’, *nà* ‘that’), spatial or temporal deictics (e.g., *zhèlǐ* ‘here’, *nàlǐ* ‘there’, *jīntiān*

'today'), and deictic verbs (e.g., *lái* 'come', *qù* 'go') (Dong, 2008; Huang, 2009; Wu, 2014; Yue, 2011). Furthermore, as Yue (2011) has suggested, expressive elements (e.g., interrogatives, imperatives, exclamations, and vocatives) (Mayes, 1990) only occur in direct speech and may serve as markers for the Mandarin direct-indirect distinction, although Yue did not provide any examples to support his suggestion. In addition, Huang (2009) has observed that word order, or "pivotal construction" (Li & Thompson, 1981, p. 607), may indicate the indirect speech of Mandarin imperatives where a noun phrase simultaneously connects two verbs. Consider example (4) from our data.

- (4) ránhòu tāmen jiào tā zǒu
 Then they ask he leave
 'Then they asked him to leave.'

The five-year-old narrator indirectly represented the imperative by using the pivotal construction, where *tā* 'he' is the object of the first verb (the speech verb), *jiào* 'ask', and the subject of the second verb, *zǒu* 'leave'.

Compared with forms, the quotative index in Mandarin SR has received less attention. As demonstrated by Gao (2013), who explored speech verbs in news reports, and Xu (1996), who examined verbs introducing direct speech in the news, *shuō* 'say' (as in *Tā shuō* "wǒ de róngyù guīyú wǒ zǔguó" 'She said, "My honor belongs to my motherland"; Xu, 1996, p. 53) was the most frequent representing signal in their data. These findings should be unsurprising, given that *shuō* 'say' is the most commonly used speech verb, as well as the 12th most frequently used word in Mandarin Chinese (Xiao, Rayson & McEnery, 2009). Although most studies of introducers have focused their interest on speech verbs, several researchers have noticed that the construction of representing signals is more complex than that of single speech verbs. For instance, a verb of speech can be modified by an adverb that describes, for example, the manner of the representing act (e.g., *dòngqīng de shuō* 'say affectionately') (Gao, 2013; Yue, 2011). In addition, *shuō* 'say' and an immediately preceding non-genetic verb can work collaboratively to be an introducer (e.g., *dīngzhǔ shuō* 'advise and say') (Fang, 2006; Gao, 2013).

The current study

The lack of representing constraints, such as tense markers and complementizers, distinguishes SR in Mandarin Chinese from that in European languages. However, compared with the extensive research on children's use of SR in European languages, little is known about how young speakers of Mandarin implement this language skill. This study aimed to examine the developmental features of SR among Mandarin-speaking children (aged three to six years) in the context of engaging in picture book narration.

Our first interest was to probe how Mandarin Chinese-speaking children apply the varied forms of SR when representing characters' speech. It is clear from previous evidence that the forms of Mandarin SR are complex (e.g., Gao, 2013; Luo, 2013; Shen, 1991; Tang, 2005). Some of them, such as direct and indirect speech, have similar counterparts in European languages and have been investigated in research on children aged three to six (Zhang *et al.*, 2018). Nevertheless, a relatively systematic view of the SR forms used by Mandarin-speaking children may require the inclusion of those forms that are unique to Mandarin Chinese (e.g., ambiguous form and mixed form). For example, it might be interesting to discover whether the ambiguous form, emerging partly due to the

tendency to omit subjects in Mandarin, manifests age- or gender-related changes. As we found in several studies reviewed in the Introduction, direct speech was the most frequent representing form for children, especially for those in a similar age range as our subjects (e.g., Ely & McCabe, 1993; Köder, 2016; Zhang *et al.*, 2018). This led us to hypothesize that the direct form would be the most preferred choice by all age groups in our study. Additionally, despite the gender differences identified by Ely and McCabe (1993), we predicted that there would be no significant gender effects in children's use of SR forms, as most of the preceding literature did not document such effects.

Another interest of the current study was to uncover how Mandarin-speaking children adopt representing signals to introduce story characters' speech. Given that *shuō* 'say' is the most frequent speech verb both in Mandarin Chinese (Xiao *et al.*, 2009) and in the speech of children (aged three to six) (Zhang, 2010), and that *say*, its English equivalent, is the dominant representing signal for English-speaking children aged four and six (Goodell & Sachs, 1992), we expected that it would be the preferential signal across the age groups of our study. Furthermore, we anticipated that older Mandarin-speaking children would demonstrate a wider scope of SR signals than younger children. For example, old children might use speech verbs with adverbial modifiers or juxtapose two representing verbs to introduce characters' speech.

Methods

Participants

The analysis involved 80 Mandarin Chinese-speaking, monolingual children in four age bands: three years of age (aged 3; 3–3;10, mean age = 43.1 months, $SD = 2.69$), four years of age (aged 4; 2–4; 10, mean age = 54.45 months, $SD = 3.17$), five years of age (aged 5; 2–5; 10, mean age = 65.06 months, $SD = 2.98$), and six years of age (aged 6; 2–6; 10, mean age = 76.55 months, $SD = 2.87$). Each age group included 20 children, with gender balancing. All the participants were randomly recruited from a daycare center and a kindergarten in the urban communities of a provincial capital city in southeast China. They were all native speakers of Mandarin Chinese, talking in Mandarin with their parents at home and with their teachers in kindergarten or at the daycare center. None of the participants had speech, hearing, or cognitive deficits, as per the reports of their parents and teachers. They all came from families in which at least one parent had earned a college degree. Parental consent was obtained prior to commencing this experiment.

Elicitation material

A series of wordless picture books by Mercer Mayer, particularly the book *Frog, where are you*, has been widely used to evaluate the narrative skills of children of various ages and languages and of children with language deficiencies (Berman & Slobin, 1994; Norbury & Bishop, 2003; Nordqvist, 2001a). It is of interest in this study to investigate how SR is used by Mandarin-speaking children of different ages in story narration. Consequently, 21 linguistics students were invited to choose the one that would trigger the most production of SR from amongst four selected picture books by Mercer Mayer (i.e., *A Boy, a Dog and a Frog*; *Frog, Where Are You*; *Frog Goes to Dinner*; *One Frog Too Many*). Upon observing the details of the pictures in the books, such as characters' body language, facial expressions, eye contact, and mouth shape, the students selected the pages that they

thought explicitly displayed speech acts being performed by the characters. Then, the pages chosen by two thirds of the students were counted. With the highest number of selected pages (11 out of 22 pages), *Frog Goes to Dinner* was chosen as the best elicitation material for the present study.

Procedure

Each child was individually tested and audio-taped by researchers in a quiet room at their schools. They first were instructed to read the book and to ask the researchers questions if they found anything difficult to understand. Researchers would explain to them such details as the identities of the characters or the types of musical instruments, without any implication of speech acts that would possibly elicit SR. When the participants felt they were ready, they were asked to tell the story. During the participants' narration, the researchers minimized their own participation except proposing some minimal questions or offering the acknowledgement tokens (e.g., by saying *ng* 'right', and *ránhòu ne* 'and then').

Transcription and coding system

All the audio-recordings were transcribed by two professional transcribers with transcription training and subsequently checked by two independently-working postgraduate students.

In order to better describe and analyze Mandarin-speaking children's representing activities performed in narratives, we proposed the following preliminary classification schemes for representing forms and representing signals in Mandarin Chinese contexts. The coding unit used here was the clause, which is a unit that contains a predicate expressing a single situation (i.e., activity, event, or state) (Berman & Slobin, 1994) and has been widely adopted for analyzing narratives (including the use of SR) produced by children speaking European languages (Drijbooms *et al.*, 2017; Nordqvist, 2001a) as well as Mandarin Chinese (Chang, 2000, 2003; Zhang *et al.*, 2018).

Forms of representation

As noted in the Introduction, previous studies have revealed that Mandarin SR is more complex than the simple direct-indirect dichotomy. This complexity renders the well-recognized continuum view of SR (which is based on European languages) (e.g., Evans, 2012; Köder, 2016; Köder & Maier, 2016; Leech & Short, 1981) applicable to the interpretation of Mandarin SR. However, undeniably, the direct-indirect distinction is still essential to the formation of the SR continuum, wherein direct and indirect speech, with their defining features, serve as diagnostic points to classify and locate various representing forms on SR's continuous scale (Köder, 2016). In addition to the continuum perspective, we also took the functional view according to which differentiating direct and indirect speech primarily lies in the deictic center or perspective of the speaker: direct speech adopts the perspective of the represented speaker or source speaker, and indirect speech manifests the representor's own point of view (Coulmas, 1986; Köder, 2016; Li, 1986; Nordqvist, 2001a; Tannen, 2007).

In this study, four types of indicators were classified for perspective differentiation in Mandarin. They were mainly adapted from previous categorizations of English (Goodell

& Sachs, 1992) and Mandarin Chinese (Dong, 2008; Huang, 2009; Yue, 2011) and a more comprehensive view concerning different languages (Coulmas, 1986). First, deictic expressions that concern the context-dependent expressions of person, time, and place are an important criterion. They include personal pronouns (e.g., *wǒ* 'I', *tā* 'she'), demonstratives (e.g., *zhè* 'this', *nà* 'that'), spatial deictics (e.g., *zhèlǐ* 'here', *nàlǐ* 'there'), temporal deictics (e.g., *jīntiān* 'today', *nàtiān* 'that day'), and deictic verbs (e.g., *lái* 'come', *qù* 'go'). For example, the represented content in direct speech is normally anchored on an *I-here-now* perspective, with indicators such as *wǒ* 'I', *zhèlǐ* 'here', *jīntiān* 'today' and *lái* 'come'. Second, the modifications frequently used in conversations can also serve as diagnostic tools of the source speaker's viewpoint, such as vocative expressions (e.g., *bàba* 'dad'), exclamations (e.g., *a* 'ah'), expressive constructions (e.g., *nǐhǎo* 'hello'), and Chinese sentence-final particles (e.g., *ne*, *ba*, *a*, *ya*), all of which typically reflect conversation (Li & Thompson, 1981). The third perspective marker, which is usually bound to the source speaker's view and manifests in direct speech, is voice features, such as changes in speech rate, volume, tone, intonation, pitch, etc. (Klewitz & Couper-Kuhlen, 1999; Li, 1986; Oliveira & Cunha, 2004). Fourth, word order, or pivotal construction, can indicate indirect forms of Mandarin imperatives. It shows the representor's adjustment associated with his/her perspective. However, we should be aware that when indirectly representing Mandarin statements and questions, word order tends to remain consistent and is thus an invalid indicator.

Based on the perspective indicators mentioned above and observations of the present data, we proposed a tentative framework for analyzing the representing forms that Mandarin-speaking children use. The coding system of classification integrates the semantic, syntactic, and prosodic features for identifying a speaker's perspective as manifest in the representing occurrences. The framework includes the following seven types (Mandarin samples, along with the translations, are included in Appendix A):

- (1) *Direct speech*. This is marked by a quotative index; its represented message, either a clause or a single word (e.g., an interjection or expression of politeness like *zài jiàn* 'goodbye'), features the perspective of the source speaker.
- (2) *Unframed direct speech*. Its content reflects the source speaker's point of view but is free from a representing introducer. It is also called zero quotative (Mathis & Yule, 1994).
- (3) *Indirect speech*. Its represented content shows the current speaker's vantage point, which is marked by a preceding frame of introducer.
- (4) *Voice representation*. This prototypically has only one clause that contains a verb of speech with or without a noun phrase following it. It gives a minimal reference to the occurrence of a verbal activity without the specification of its details.
- (5) *Mixed form*. This indicates a mingling of features of both the source speaker's and the representor's perspectives and, in most cases, involves mixes of deictic expressions.
- (6) *Ambiguous form*. Containing a quotative index, this form can be either direct speech or indirect speech due to the absence of features that specify the speaker's viewpoint in the represented message.
- (7) *Unclear utterance*. With neither a framing introducer nor any features of the source speaker's involvement in conversational exchanges, it is unclear whether the utterance is represented speech or thought, or whether it is a non-representing statement of the current speaker.

Our framework does not include unframed or free indirect speech because it is far less clearly marked in Mandarin Chinese (Hagenaar, 1996). As a stylistic device thoroughly examined in the literary field (e.g., Leech & Short, 1981), this form in Indo-European languages contains a mingling of features of tense and person selection associated with indirect speech and unembedded sentences typical of direct speech. However, since verb tense is absent and the subject of a sentence is often left implicit in Chinese contexts, making a distinction between free indirect speech and free direct speech nearly impossible in many cases. When presented without a frame, indirect representation does not refer to speech events, but rather to facts (Hickmann, 1993). Consequently, if there are no original utterances that can be referred back to, it becomes extremely difficult to distinguish free indirect speech from statements of fact. However, this was the exact case in the current study on wordless storybook narration.

Representing signals

Together with a reference to the source speaker, speech verbs in introducers often signal the representation of verbal activities occurring. As mentioned in the Introduction, previous research has mostly focused on single speech verbs as representing signals, which are semantically divided into generic and non-generic speech act verbs (Goodell & Sachs, 1992; Özyürek, 1996). The adverbial modifiers that adult representors use (Nordqvist, 2001a), and the mixed forms (such as *say like*) that teenagers use (Levey, 2003), have revealed that representors may have other signals available that are more syntactically complex than single-word verbs.

Our exploration into Mandarin-speaking children's use of representing signals attempted to combine traditional semantic criteria with syntactic differentiation. Based on the reformulation of the semantic categories above, Li and Thompson's (1981) analyses of Mandarin verbs, and the observation of our data, we classified children's representing signals into six categories (Mandarin samples, along with the translations, are included in Appendix B):

- (1) *Shuō* 'say'. This is a typical generic verb, and its equivalents in other languages are the most common lexical signals of SR.
- (2) *Neutral verb*. A generic verb other than *shuō* 'say' that is also semantically unmarked, such as *jiang* 'speak' and *gāosù* 'tell'.
- (3) *Non-generic verb*. A verb that characterizes the represented content by indicating either illocutionary force or the manner of speaking, such as *mà* 'scold' and *hǎn* 'yell'.
- (4) *Adverb + de + representing verb*. This contains a speech verb and, generally, the accompanying construction of a "manner adverb" (Li & Thompson, 1981, p.322) with the particle *de*, which offers a detailed description of the verb – for example, *shēngqì de shuō* 'angrily said'.
- (5) *Verb₁ + Verb₂*. This is a "serial verb construction" (Li & Thompson, 1981, p.594). The representing verb (verb₂ in the form), which is normally *shuō* 'say', has a preceding verb (verb₁ in the form) either presenting illocutionary force, such as *zéguài tā shuō* 'rebuked him and said', or referring to a kinetic behavior immediately prior to speaking (Semino & Short, 2004), such as *zhǐzhe tā shuō* 'pointed at him and said'.

- (6) *Representing verb + object*. This is a verb phrase consisting of a representing verb and its object that is normally the production of speech such as words and sentences: for example, *shuō yī jù huà* 'said one sentence'.

Reliability

All of the narratives produced by the subjects were independently coded by two annotators. The kappa coefficients for the coding of the representing forms and signals were .89 and .93, respectively, indicating substantial agreement between the annotators. All the disagreements were resolved through discussions between the coders and the first author. Eventually, the first author checked all the coded examples before including them in the analysis.

Results

This study found that SR was a quite common narrative method that almost all children in the four age groups adopted in their story narratives. Except for a boy and a girl aged three and a boy aged four, 96.25% of the subjects (77 out of 80) used, at least once, a certain form of SR.

When analyzing both representing forms and signals, the differences in the number of clauses that each participant produced had to be considered. To avoid the influence of different story lengths, the proportional measures of each representing form and signal were calculated for every participant. For example, a child's direct speech rate was the frequency of direct speech over the total number of narrative clauses produced by that child. For a clearer comparison, we adopted the standardizing method (Drijbooms *et al.*, 2017; Ely & McCabe, 1993) and multiplied all the measures by 100 to find the rates per 100 clauses.

Overall use of SR

To assess the development in Mandarin-speaking children's total rates of SR use (displayed in Table 1), ANOVA tests were conducted with age and gender as the factors. The results revealed a significant age-related influence on children's use of SR as a whole ($F [3, 72] = 17.97, p < .001, \eta^2 = .42$). LSD post hoc analyses showed that with the exception of five- and six-year-olds with marginally significant differences ($p = .074$), comparisons among all the other age groups resulted in statistically significant differences: between three- and five-year-olds, three- and six-year-olds, and four- and six-year-olds ($p < .001$), between four- and five-year-olds ($p < .01$), and between three- and four-year-olds ($p < .05$). However, neither significant gender effects ($F (1, 72) = 1.07, p = .30, \eta^2 = .015$) nor an interaction between age and gender ($F [3, 72] = .014, p = .998, \eta^2 = .001$) were observed in the children's overall use of SR.

Forms of representation

Table 1 presents the mean rates (per 100 clauses) of representing forms for each age group in decreasing order. To further analyze the developmental trend of children's representing forms, a set of MANOVAs were conducted, with age and gender as the independent factors and the mean rates of seven forms as the dependent variables.

Table 1. Means and standard deviations of forms of speech representation

	Age 3		Age 4		Age 5		Age 6	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
<i>Direct speech</i>	1.38 (1.83)	1.56 (2.26)	5.82 (5.75)	4.23 (2.69)	10.16 (7.10)	9.72 (5.06)	11.66 (10.94)	12.42 (4.78)
<i>Voice representation</i>	0.67 (1.42)	0.50 (1.08)	1.27 (1.77)	3.23 (1.98)	2.78 (2.91)	4.37 (1.62)	4.34 (2.63)	5.78 (3.24)
<i>Indirect speech</i>	0.26 (0.83)	0.43 (0.91)	1.08 (1.46)	1.33 (2.37)	1.23 (2.37)	0.23 (0.72)	1.35 (2.35)	0.52 (1.19)
<i>Unframed direct speech</i>	0.99 (2.13)	1.52 (2.00)	0.00 (0.00)	0.73 (1.62)	0.00 (0.00)	0.00 (0.00)	0.29 (0.90)	0.54 (0.90)
<i>Unclear utterance</i>	0.40 (1.26)	0.29 (0.93)	0.00 (0.00)	0.50 (1.58)	0.53 (1.53)	1.53 (2.27)	0.29 (0.90)	0.20 (0.6)
<i>Ambiguous form</i>	0.00 (0.00)	0.22 (0.70)	0.24 (0.77)	0.00 (0.00)	0.26 (0.81)	0.69 (1.46)	0.19 (0.60)	0.60 (1.31)
<i>Mixed form</i>	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.59 (1.24)	0.86 (2.03)
Total	3.70 (2.30)	5.01 (1.99)	8.42 (6.58)	10.01 (5.02)	14.96 (8.42)	16.54 (7.45)	18.69 (12.39)	20.92 (8.45)

Note. The numbers are percentages for the targeted forms per 100 clauses. Standard deviations are in parentheses.

The results revealed that there was a significant multivariate effect of age influencing the children's use of representing forms (Wilks' $\lambda = .37$, $F [3, 72] = 4.64$, $p < .001$, $\eta^2 = .28$). This was largely due to the significant univariate effects of age on the use of the two most frequent forms: direct speech ($F [3, 72] = 13.67$, $p < .001$, $\eta^2 = .36$) and voice representation ($F [3, 72] = 15.02$, $p < .001$, $\eta^2 = .36$). Significant age effects were also found in unframed direct speech ($F [3, 72] = 3.53$, $p < .05$, $\eta^2 = .13$) and mixed form ($F [3, 72] = 3.69$, $p < .05$, $\eta^2 = .13$). In other words, four out of the seven SR forms that children used developed quite considerably over the ages of three to six. As for direct speech, the LSD post hoc tests revealed that the significant differences remained between the age groups of three- and five-year-olds, three- and six-year-olds, and four- and six-year-olds ($p < .001$), and between four- and five-year-olds ($p < .01$). Marginally significant differences were found between three- and four-year-olds ($p = .055$). These findings suggest that with the exception of five- and six-year-old age groups ($p = .255$), all the older groups employed significantly more direct speech than did the younger ones. Likewise, the LSD post hoc tests on voice representation showed significant differences between almost all age groups, including three- and five-year-olds, three- and six-year-olds, and four- and six-year-olds ($p < .001$), as well as three- and four-year-olds, and five- and six-year-olds ($p < .05$). Marginally significant differences were found between the adjacent age groups of four and five years ($p = .061$). For unframed direct speech, multiple comparisons showed significant age effects between three- and five-year-olds ($p < .01$), and between three- and four-year-olds, and three- and six-year-olds ($p < .05$). The LSD post hoc tests revealed that

the six-year-olds used the mixed form more than any younger group ($p < .01$) since it was absent from the narratives of children through age five but was present thereafter.

However, the MANOVAs did not yield significant effects of gender (Wilks' $\lambda = .86$, $F [1, 72] = 1.60$, $p = .15$, $\eta^2 = .15$) or age-gender interaction (Wilks' $\lambda = .53$, $F [3, 72] = .60$, $p = .96$, $\eta^2 = .05$) on SR forms.

According to Table 1, direct speech was the most frequent form across age groups. This leads to the question of whether this preference relates to the pronoun usage of Mandarin-speaking children. As reviewed in the Introduction, personal pronouns are often recognized as the sole diagnostic criterion for the Mandarin direct-indirect distinction, due to the lack of tense shifts and complementizers. However, given the zero pronoun in Mandarin, skepticism concerning pronoun usage in direct speech may arise. Therefore, we turned to the one-sample t -test to determine child speakers' tendency to employ personal pronouns in direct speech. If a strong tendency was attested, a compelling argument for interpreting children's dominant use of direct form would be provided in the Discussion. The one-sample t -test was effectively applied by Köder and Maier (2016, 2018) to measure the accuracy of pronoun interpretation in SR.

Using the t -tests, we analyzed the mean use (in contrast to the omission) of personal pronouns in the direct speech of all the age groups, and compared them with the chance level of .5 (pronoun use vs. pronoun omission). As the one-sample t -tests indicated, Mandarin-speaking children's mean use of personal pronouns in direct speech was .89 ($SD = .33$), .74 ($SD = .45$), .73 ($SD = .45$) and .80 ($SD = .41$) for the age groups of three-, four-, five-, and six-year-olds, respectively. These measurements were significantly higher than the chance level of .5: three-year-olds, $t (8) = 3.50$, $p < .05$, $d = 1.17$; four-year-olds, $t (33) = 3.06$, $p < .01$, $d = .53$; five-year-olds, $t (69) = 4.27$, $p < .001$, $d = .51$; six-year-olds, $t (102) = 7.42$, $p < .001$, $d = .73$. This means that among young speakers of Mandarin Chinese (three- to six-year-olds), personal pronouns serve as crucial indicators in their uses of direct speech.

Representing signals

Table 2 shows the mean standardized rates (per 100 clauses) of the six types of representing signals in decreasing order. MANOVA tests were carried out to evaluate the development of children's use of representing signals, with age and gender as the fixed factors and the means of signals as the dependent variables.

The multivariate effect of age on children's use of SR signals was significant (Wilks' $\lambda = .32$, $F [3, 72] = 5.30$, $p < .001$, $\eta^2 = .32$). The subsequent analyses of variance suggested that age exerted significant univariate effects on all types of signals: *shuō* ($F [3, 72] = 11.71$, $p < .001$, $\eta^2 = .32$), non-generic verb ($F [3, 72] = 11.25$, $p < .001$, $\eta^2 = .32$), 'verb₁ + verb₂' ($F [3, 72] = 5.37$, $p < .01$, $\eta^2 = .17$), neutral verb ($F [3, 72] = 4.12$, $p < .01$, $\eta^2 = .15$), 'adverb + *de* + representing verb' ($F [3, 72] = 4.36$, $p < .01$, $\eta^2 = .15$), and 'representing verb + object' ($F [3, 72] = 5.20$, $p < .01$, $\eta^2 = .17$). In terms of *shuō* 'say', the most favored SR signal, the LSD post hoc tests revealed significant differences between the four pairs of groups – three- and five-year-olds and three- and six-year-olds ($p < .001$), three- and four-year-olds ($p < .01$), and four- and six-year-olds ($p < .05$) – and a marginally significant difference was found between four- and five-year-olds ($p = .06$). Yet, there was no evidence that the two oldest groups (five- and six-year-olds) used *shuō* differently ($p = .86$). Similarly, these two groups did not show robust differences when using non-generic verbs. Significant differences, however, were found between the rest of the group

Table 2. Means and standard deviations of signals of speech representation

	Age 3		Age 4		Age 5		Age 6	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
<i>Shuō</i> 'say'	1.38 (1.83)	2.22 (2.45)	8.38 (6.65)	6.84 (4.38)	12.96 (4.84)	9.23 (4.02)	10.69 (9.73)	12.17 (8.14)
<i>Non-generic verb</i>	0.62 (1.33)	0.50 (1.08)	1.78 (2.22)	2.59 (2.38)	3.58 (3.29)	3.90 (1.79)	4.59 (3.74)	4.67 (2.03)
<i>Verb₁ + verb₂</i>	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	1.58 (1.86)	1.66 (3.24)	1.45 (1.61)
<i>Neutral verb</i>	0.31 (0.99)	0.00 (0.00)	0.00 (0.00)	0.25 (0.79)	0.00 (0.00)	0.00 (0.00)	1.09 (1.84)	1.01 (2.01)
<i>Adverb + de + verb_r</i>	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.51 (1.08)	0.57 (1.20)	0.83 (1.17)
<i>Verb_r + object</i>	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.47 (1.03)	0.94 (1.67)

Note. The numbers are percentages for the targeted signals per 100 clauses. Standard deviations are shown in parentheses. *Verb_r* = Representing verb

pairings: between three- and five-year-olds and three- and six-year-olds ($p < .001$), four- and six-year-olds ($p < .01$), and three- and four-year-olds and four- and five-year-olds ($p < .05$). For 'verb₁ + verb₂', multiple comparisons showed that six-year-olds used this signal significantly more than the two youngest groups (three- and four-year-olds) ($p < .01$), and that marginally significant differences existed between three- and five-year-olds and four- and five-year-olds respectively ($p = .088$). Regarding the 'adverb + *de* + speech verb' signal, significant differences were found between the oldest group of six-year-olds and the two youngest groups (three- and four-year-olds) ($p < .01$), with five- and six-year-olds showing a marginally significant difference ($p = .052$). Concerning the 'representing verb + object' signal, children aged six tended to apply it significantly more often than the three younger groups ($p < .01$).

The MANOVAs did not reveal significant multivariate effects of gender (Wilks' $\lambda = .92$, $F [1, 72] = .93$, $p = .48$, $\eta^2 = .077$) or an age-gender interaction (Wilks' $\lambda = .79$, $F [3, 72] = .91$, $p = .572$, $\eta^2 = .075$) on the children's use of representing signals.

Form-signal preference in SR

In addition to analyzing the development of forms and signals separately, we also sought to find out the interrelationship between these two variables by investigating children's preferred signals across different forms. The top-three representing forms (i.e., direct speech, voice representation, and indirect speech), which accounted for 96.46% of all the SR occurrences using representing signals, were chosen for analysis. Meanwhile, representing forms without introducers (i.e., unframed direct speech and unclear utterance), forms of low frequencies (i.e., mixed form and ambiguous form), and instances with

Table 3. Signal preference for different representing forms

	Age 3	Age 4	Age 5	Age 6
Direct Speech				
<i>Shuō</i> 'say'	88.89% (8)	97.06% (33)	90.00% (63)	77.23% (78)
<i>Non-generic verb</i>	11.11% (1)	2.94% (1)	0%	5.94% (6)
<i>Verb</i> ₁ + <i>verb</i> ₂	0%	0%	7.14% (5)	11.88% (12)
<i>Neutral verb</i>	0%	0%	0%	0%
<i>Adverb</i> + <i>de</i> + <i>verb</i> _r	0%	0%	2.86% (2)	4.95% (5)
<i>Verb</i> _r + <i>object</i>	0%	0%	0%	0%
Voice Representation				
<i>Shuō</i> 'say'	0%	12.50% (2)	7.69% (2)	0%
<i>Non-generic verb</i>	75.00% (3)	81.25% (13)	88.46% (23)	66.67% (28)
<i>Verb</i> ₁ + <i>verb</i> ₂	0%	0%	3.85% (1)	0%
<i>Neutral verb</i>	25.00% (1)	6.25% (1)	0%	16.67% (7)
<i>Adverb</i> + <i>de</i> + <i>verb</i> _r	0%	0%	0%	2.38% (1)
<i>Verb</i> _r + <i>object</i>	0%	0%	0%	14.29% (6)
Indirect Speech				
<i>Shuō</i> 'say'	100.00% (3)	88.89% (8)	60.00% (3)	75.00% (6)
<i>Non-generic verb</i>	0%	11.11% (1)	40.00% (2)	12.50% (1)
<i>Verb</i> ₁ + <i>verb</i> ₂	0%	0%	0%	12.50% (1)
<i>Neutral verb</i>	0%	0%	0%	0%
<i>Adverb</i> + <i>de</i> + <i>verb</i> _r	0%	0%	0%	0%
<i>Verb</i> _r + <i>object</i>	0%	0%	0%	0%

Note. The percentages were calculated for each form by each age group. Raw frequencies are shown in parentheses. *Verb*_r = Representing verb.

introducers that only contained source speakers but lacked representing signals, were excluded from the analysis.

As shown in Table 3, all groups used *shuō* in their direct speech most of the time, whereas the percentages roughly dropped with age, from 88.89% of three-year-olds to 77.23% of six-year-olds. While the two younger groups of three- and four-year-olds stuck with single representing verbs, the older groups added extended verb constructions, with six-year-olds displaying the most widely distributed pattern. For voice representation, the non-generic verb was the largest category of signals across all ages. Five-year-olds relied on non-generic verbs most heavily (88.46%), whereas six-year-olds used them comparatively less (66.67%) than the other groups. A similar trend to that of direct speech was evident here: the two younger groups were strictly confined to single speech verbs and the two older groups included the extended verb constructions. Despite the fact that *shuō* was the second-favored signal in voice representation for the age groups of four-year-olds (12.50%) and five-year-olds (7.69%), it was absent in the repertoire of six-year-olds. Instead, the oldest group turned to 'verb_r + object' as the second most frequently used

signal (14.29%), which never appeared in the other groups' data. When introducing indirect speech, all the groups preferred *shuō* the most. Non-generic verbs were used by the age groups of four, five and six, and 'verb₁ + verb₂' was used by six-year-olds, but the frequencies were quite low.

Discussion

Our study investigated how Mandarin Chinese-speaking children (aged three to six) apply SR by examining the representing forms and signals that they used in their fictional narratives. It sought to provide evidence for developmental patterns in their use of SR and test our hypotheses.

Overall use of SR

Congruent with Ely and McCabe's (1993) findings, our results revealed that Mandarin-speaking children's use of SR significantly increases with age. This trend is associated with the progression embodied in the representing forms and signals that the children produced, which will be explored in the following two sections. Additionally, children's increasing ability to understand storybook characters may boost their representing practice. When children tell stories based on wordless picture books, they represent characters' speech by apparently relying on their imagination and reasoning rather than on memory-based recall. This requires children to understand the mental states of the characters (e.g., their thoughts, feelings, and intentions) through the ability of theory of mind or social imagination (Lysaker & Miller, 2012), and then to give voice to the characters by bearing in mind their conscious understanding of them. There is evidence among children in preschools or kindergartens showing that their capacity to interpret characters' internal states develops with age (Curenton, 2011; Lai *et al.*, 2010), and that this capacity positively influences their tendency to represent more characters' speech when narrating wordless picture books (Lysaker, Shaw & Alicia, 2016).

Development of representing forms

Similarities among groups: Dominant direct representation

Our first hypothesis concerns the dominant representing form among children. There are conflicting results in this regard: some studies have shown that children use unframed direct speech the most (Hickmann, 1993; Özyürek, 1996); others have suggested that the indirect form overshadows the others (Goodell & Sachs, 1992); and still others have demonstrated that the direct form dominates (Ely & McCabe, 1993; Köder, 2016; Nordqvist, 2001a, 2001b; Zhang *et al.*, 2018), which our findings echoed. The predominant role of direct speech across all ages has confirmed our prediction, and it can be attributed mainly to the fact that Mandarin direct speech requires fewer processing loads for children than other forms (e.g., indirect speech) in two regards.

First, in terms of personal pronoun acquisition, direct speech is easy for children who can retain the use of a first-person pronoun as the speaker, albeit in a represented context. On the one hand, representing speech directly or indirectly in Mandarin contexts largely depends on personal pronoun choices, as linguistic constraints (e.g., tense shifts) are absent from Mandarin. Our results confirmed Mandarin-speaking children's significant

tendency to use personal pronouns in direct speech. On the other hand, due to the egocentric nature of children's thoughts, first- and second-person pronouns may be easier to produce and comprehend than third-person forms (Clark, 1978; Zhu, Cao & Zhang, 1986). As can be expected, the first person is the most frequently used personal pronoun among Mandarin-speaking children aged three to six, followed by the second person and then the third person (Zhang, 2010). Notably, all of our subjects assumed the perspective of a third-person narrator in their fictional narrative settings; they could not use an egocentric perspective featuring first- and second-person pronouns. The only exception was in their use of direct speech, which had a represented message anchored to the viewpoint of the represented speaker, allowing the representors to maintain a first-person perspective in their representations. In contrast, the represented content in indirect speech still featured the narrator's viewpoint, which multiplied the representor's processing efforts. When referencing, for example, the represented speaker, representors were blocked from using a first-person system and instead kept using a less-sophisticated third-person perspective.

Second, regarding syntactic adaptations, the direct speech of certain sentence types (such as commands) is easier to produce than their indirect forms. As mentioned above, the indirect representation of commands in Mandarin usually involves a pivotal construction, and children undergo significant development in this regard between the ages of three and six years (Zhang, 2019). Finding a proper structure to adequately relay the syntactic information between the two verbs requires more effort and is consequently determined by children's general grammatical development (Cheung, 2009). Given the higher processing loads entailed, it is expected that children are more likely to represent commands in direct speech than in indirect speech. A case in point is the command representation featured in one page of our eliciting book, wherein a little boy, based on his mouth shape and body posture, seems to ask a waitress to return his pet to him. Among the 26 instances of SR that the children performed, only two were in indirect forms with a pivotal construction, and the rest were in direct forms.

Group differences

In our study, a significant univariate age effect was observed in children's use of direct speech, which is in line with previous research (Ely & McCabe, 1993; Zhang *et al.*, 2018). According to our data, children's use of direct representation develops critically during the ages of three to six, since significant differences existed among almost all the adjacent age groups. However, the marked trend stops after age five, as six-year-olds slightly outperformed five-year-olds in our study. Conversely, five-year-olds' use of direct speech decreased in Ely and McCabe's (1993) study, with boys producing no examples (compared to all of our five-year-old boys producing some examples). This may indicate that Mandarin-speaking children achieve the cognitive and linguistic maturity required to use this form by the ages of five and six, which is earlier than their English-speaking peers, who are constrained by more grammatical rules. Interestingly, the Mandarin-speaking three-year-olds in Zhang *et al.*'s (2018) research did not produce any instances of direct form use, and significant differences only remained between them and the oldest children. Perhaps a similar explanation of the presence of elicitation differences may hold, too. Compared with the narrative topics that Ely and McCabe (1993) and Zhang *et al.* (2018) used, the picture book used in our study – with its large proportion of pages that evidently depicted characters' verbal activities – triggered more children to produce SR.

Our study found that children's use of voice representation, the second major form, is significantly susceptible to an age influence, which corroborates the existing research on English-speaking children (Ely & McCabe, 1993). Our findings suggest that Mandarin-speaking children undergo a critical period of development between the ages of three and six in terms of their use of this highly summarizing form that serves to recapitulate less important information and provide a background for narration (Semino & Short, 2004). Furthermore, children's significant development in this respect indicates their growing awareness of their listener's vantage point by using voice representation for unimportant speech and saving more vivid forms (such as direct speech) for story highlights. This developmental process demonstrates that their narratives shift from being "reporter-centered" to "listener-centered" (Goodell & Sachs, 1992, p. 417).

What is also noteworthy is that indirect representation, the third most common SR form, does not exhibit significant effects of age or gender or any significant age group differences. Nevertheless, some points merit attention, following our tentative comparisons between the use of indirect speech of our subjects and that of children in previous studies, who spoke European languages. First, with indirect speech consistently occurring in the data of each gender across all age groups, the Mandarin-speaking children in our study showed a more stable use of the indirect form than their Western peers of the same age range (aged three to six). For example, this form was absent in Turkish-speaking five-year-olds, the only group in this age range that Özyürek (1996) investigated. Similarly, among the three Swedish-speaking age groups (three to five years) in the study of Nordqvist (2001a), neither the three- nor the five-year-olds applied the indirect form. In addition, none of the English-speaking boys in the three age groups (four to six) that Ely and McCabe (1993) observed included indirect speech in their narratives. Second, in contrast to Mandarin indirect speech, which the four-year-olds, the second youngest group in our study, used most, the indirect form in European languages appeared most frequently among the oldest groups of children in the respective studies: e.g., the 13-year-old Turkish-speaking group (Özyürek, 1996), the 15-year-old Swedish-speaking group (Nordqvist, 2001a), and the eight-year-old English-speaking group (Ely & McCabe, 1993).

A possible explanation for these discrepancies is that, except for pronoun shifts (which also occur in Mandarin SR), indirect speech in European languages is subject to higher syntactic complexity than Mandarin, which develops over a long period. For example, eight-year-old English-speaking children still struggle with incorrect tense shifts in their indirect speech (Goodell & Sachs, 1992), and five-year-olds are not yet able to correctly use nominalization, as required in the Turkish indirect form (Slobin, 1986). In other words, Mandarin SR forms lack some rigid conditions, such as tense markers, making the production of indirect speech less challenging for Mandarin Chinese representors than for their counterparts who speak European languages. Additionally, given that a clear developmental trajectory seemed to emerge in the previous findings when a wide age range of children was observed (Ely & McCabe, 1993; Nordqvist, 2001a; Özyürek, 1996), the comparatively narrow age range in our study might account for the non-significant age effect we found. Meanwhile, Mandarin's lack of complex syntactic constructions that are correctly used only at later ages may partly explain Mandarin-speaking children's successful production of the indirect form at an early age and their lack of a significant developmental pattern in this regard. Therefore, it would be advisable in the future to extend research to include more age groups to identify the existence of this apparent trend in Mandarin-speaking children's indirect speech.

The results of this study confirmed our hypothesis that there were no significant gender effects regarding children's use of representing forms. This corroborated the findings of Zhang *et al.* (2018), which revealed no gender differences among Mandarin-speaking children when they used direct and indirect speech. Furthermore, our analysis expanded on their findings from two kinds of forms to all the Mandarin SR forms we investigated. Given that personal pronouns are an important diagnostic criterion for classifying Mandarin representing forms, the similar performance of boys and girls (aged 3;5 to 5;5) in their use of personal pronouns (Zhu *et al.*, 1986) might partly account for the observed non-significant gender effect on forms. Although there is, to our knowledge, no specific research on Mandarin-speaking children's development in speech verbs or other representing signals, our research might provide some quantitative evidence that Mandarin-speaking boys and girls share more similarities than differences in their choices of SR introducers.

Development of representing signals

The displayed predominance of *shuō* 'say' in our findings can support our hypothesis and echo previous work on its equivalents used by children speaking other languages (e.g., Ely & McCabe, 1993; Ely *et al.*, 1995; Goodell & Sachs, 1992; Özyürek, 1996). However, Mandarin-speaking children did not show as strong a propensity for it as their peers speaking English and other European languages. For example, four-year-old English-speaking children used *say* in about 80% of all the representing verbs (Goodell & Sachs, 1992), and five-year-old Turkish-speaking children showed an almost exclusive use of *dedi* 'said' (Özyürek, 1996); the proportions in our corresponding age groups were 76.71% and 71.03%. We suggest that the speech activities (such as yelling, criticizing, and requesting) demonstrated in our material may have encouraged a greater use of non-generic verbs. Such speculation, of course, requires further study. The prevalence of *shuō* is largely attributed to its salient semantic neutrality, which makes it a versatile substitute for any speech verb and reduces the representors' efforts. The neutral status of such signals can minimize their interference with the represented utterance and maximize the utterance's pragmatic force (Lucy, 1993). Consequently, *shuō* is a particularly ideal signal for direct speech that features explicit vividness (Tannen, 2007), with its dominance in signalling direct speech evident in Table 3.

By extending the range of representing signals for observation, this study provided evidence for how Mandarin-speaking children shift from the exclusive use of single speech verbs at the early age range of three through four years to more varied use, including complicated verb constructions, from the age of five years onwards; this confirmed our final hypothesis about older children using more diversified signals than their younger peers. Statistical analysis showed that age significantly constrained every signal, and significant differences remained between the youngest and oldest groups in the use of every signal, suggesting that the period between ages three and six is important for Mandarin-speaking children's development of using representing signals. The age of five marks the onset of children's progression toward more complicated verb constructions that presumably require further improvement of their grammatical capacity. For example, the emergence of adverbs accompanying speech verbs can be associated with children's significant development in adverb use between the ages of five and six years (Zhang, 2019). However, considering the empirical evidence that shows that children gain their stable ability to use serial verb constructions at the age of three-and-a-half years

(Cheung, 2009), our findings indicate there is a lag between children's general usage abilities and the specific application of speech verbs.

Through form–signal preference analysis, our study depicted children's signal preferences when representing different forms. According to Lucy (1993), *say* – featuring semantic unmarkedness – can provide minimal interference with the represented message and allow representors to focus differentially on the form of the represented message (i.e., direct speech) or the referential content (i.e., indirect speech). Our findings have provided further cross-linguistic evidence for this. What's more, children of all ages showed a strong propensity for signalling voice representation with non-generic verbs. Voice representation is used to summarize relatively unimportant speech, and non-generic verbs that specify illocutionary force or an intended purpose of speech are perfectly suited to the minimal reference to speech acts and, possibly, foreground the emergent plot. Finally, form–signal preference analysis can present form-based evidence for our prediction about the greater variety of signals used by older children.

Limitations and future research

There are certain limitations to this study. Firstly, in the Discussion, the cross-linguistic comparisons between the Mandarin-speaking children in our study and those speaking other languages in previous literature sources are largely speculative, since the studies are based on, for example, different age groups, and dissimilar methods in terms of the elicitation tasks and of the classifications of the representing forms. As an anonymous reviewer suggested, we should consider these potentially influential factors when drawing comparisons. Future investigations require comparable and reliable methods applied uniformly across datasets sourced from different language speakers to uncover developmental patterns relating to representation. Secondly, this study used a wordless picture book or storytelling alone to elicit narratives. Future investigations should adopt different elicitation methods, such as story retelling and personal narratives, to gain a better understanding of the development of SR skills in children across different communicative contexts. Thirdly, our study relied primarily on caregiver reports to exclude children with language impairments and other disorders, so the findings cannot be generalized to TD Mandarin-speaking children. Therefore, future work should implement a more reliable screening procedure that includes, for example, standardized tests for assessing children's verbal/non-verbal abilities, to investigate the general developmental pattern of Mandarin SR that TD children use, as well as that children with language disorders use.

Conclusions

This study was conducted to investigate how Mandarin Chinese-speaking children, from the ages of three to six years, employ SR in their story narratives by analyzing the representing forms and signals that they use. On the one hand, our findings, such as the dominant representing form and signal, are congruent with other explorations of similar populations who speak European languages (e.g., Ely & McCabe, 1993; Köder, 2016; Nordqvist, 2001a, 2001b; Özyürek, 1996). On the other hand, our study revealed that Mandarin speakers aged three to six years demonstrated a more stable use of indirect speech than their Western peers of the same age range (e.g., Ely & McCabe, 1993; Nordqvist, 2001a; Özyürek, 1996). Based on the analysis of SR used by

Mandarin-speaking children, this study contributes to a cross-linguistic perspective for children's developmental features relating to their use of this narrative skill.

Acknowledgements. We thank the anonymous reviewers and the editors for their thoughtful comments and suggestions. This research was supported by Ministry of Education Project of Humanities and Social Sciences, China (Grant No. 20YJA740014), and Zhejiang Education Department Project for Teachers' Development (Grant No. FX2019026). We are grateful to all the children who participated in this study.

References

- Astington, J. W., & Jenkins, J. M.** (1995). Theory of mind development and social understanding. *Cognition and Emotion*, *9*(2-3), 151-165.
- Bamberg, M., & Damrad-Frye, R.** (1991). On the ability to provide evaluative comments: further explorations of children's narrative competencies. *Journal of Child Language*, *18*(3), 689-710.
- Berman, R. A., & Slobin, D. I.** (eds.). (1994). *Relating events in narrative: a crosslinguistic developmental study*. Hillsdale, NJ: Erlbaum.
- Chang, C.** (2000). *Narrative performance across contexts and over time: preschool Chinese children and mothers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9968293).
- Chang, C.** (2003). Talking about the past: how do Chinese mothers elicit narratives from their young children across time. *Narrative Inquiry*, *13*(1), 99-126.
- Chang, C., & Huang, C.** (2016). Mother-child talk during joint book reading in two social classes in Taiwan: interaction strategies and information types. *Applied Psycholinguistics*, *37*(2), 387-410.
- Cheung, H.** (2009). Grammatical characteristics of Mandarin-speaking children with specific language impairment. In S. Law, B. S. Weekes, & A. M. Wong (eds.), *Language disorders in speakers of Chinese* (pp. 33-52). Bristol: Multilingual matters.
- Clark, E. V.** (1978). From gestures to word: on the natural history of deixis in language acquisition. In J. S. Bruner & A. Garton (eds.), *Human growth and development* (pp. 85-120). Oxford: Clarendon Press.
- Coulmas, F.** (1986). Represented speech: some general issues. In F. Coulmas (ed.), *Direct and indirect speech* (pp. 1-28). Berlin: Mouton de Gruyter.
- Curenton, S. M.** (2011). Understanding the landscapes of stories: the association between preschoolers' narrative comprehension and production skills and cognitive abilities. *Early Child Development and Care*, *181*, 791-808.
- Dong, X.** (2008). Shiji yupian zhong zhijie yinyu yu jianjie yinyu de hunyong xianxiang [Blending of direct discourse and indirect discourse in actual texts]. *Linguistic Sciences*, *35*(4), 367-376.
- Drijbooms, E., Groen, M. A., & Verhoeven, L.** (2017). Children's use of evaluative devices in spoken and written narratives. *Journal of Child Language*, *44*(4), 767-794.
- Ely, R., & McCabe, A.** (1993). Remembered voices. *Journal of Child Language*, *20*(3), 671-696.
- Ely, R., Gleason, J. B., Narasimhan, B., & McCabe, A.** (1995). Family talk about talk: mothers lead the way. *Discourse Processes*, *19*(2), 201-218.
- Evans, N.** (2012). Some problems in the typology of quotation: a canonical approach. In D. Brown, M. Chumakina, & G. G. Corbett (eds.), *Canonical Morphology and Syntax* (pp. 66-98). Oxford: Oxford University Press.
- Fang, M.** (2006). Beijinghua li shuo de yufahua: cong yanshuo dongqi dao congju biaoji [Grammaticalization of *shuo* (say) in Beijing Mandarin: from lexical verb to subordinator]. *Journal of Chinese Dialects*, *1*(1), 107-121.
- Gao, X.** (2013). *Han ying baozhi xinwen yupian zhong zhuanshu yanyu de bijiaoyan jiu* [A comparative analysis of the use of reported speech in Chinese and English newspapers]. (Unpublished doctoral dissertation). Nanjing Normal University, Nanjing, China.
- Goodell, E. W., & Sachs, J.** (1992). Direct and indirect speech in English-speaking children's retold narratives. *Discourse Processes*, *15*(4), 395-422.

- Güldemann, T.** (2012). Thetic speaker-instantiating quotative indexes as a cross-linguistic Type. In I. Buchstaller & I. van Alphen (eds.), *Quotatives: cross-linguistic and cross-disciplinary perspectives* (pp.117–142). Amsterdam: John Benjamins.
- Hagenaar, E.** (1996). Free indirect speech in Chinese. In T. Janssen & W. van der Wurff (eds.), *Reported speech: forms and functions of the verb* (pp. 289–298). Amsterdam: John Benjamins.
- Hasund, I. K., Opsahl, T., & Svennevig, J.** (2012). In I. Buchstaller & I. van Alphen (eds.), *Quotatives: cross-linguistic and cross-disciplinary perspectives* (pp.37–70). Amsterdam: John Benjamins.
- Hemphill, L., Picardi, N., & Tager-Flusberg, H.** (1991). Narrative as an index of communicative competence in mildly mentally retarded children. *Applied Psycholinguistics*, **12**, 263–279.
- Hemphill, L., Uccelli, P., Winner, K., Chang, C. J., & Bellinger, D.** (2002). Narrative discourse in young children with histories of early corrective heart surgery. *Journal of Speech Language & Hearing Research*, **45**(2), 318–331.
- Hickmann, M.** (1993). The boundaries of represented speech in narrative discourse: some developmental aspects. In J. Lucy (ed.), *Reflexive language: represented speech and metapragmatics* (pp. 63–90). Cambridge: Cambridge University Press.
- Huang, Y.** (2009). *Zhuan shu huayu yanjiu [Study on the reported speech]* (Unpublished doctoral dissertation). Fudan University, Shanghai, China.
- Klewitz, G., & Couper-Kuhlen, E.** (1999). Quote-unquote: the role of prosody in the contextualization of represented speech utterances. *Pragmatics*, **9**, 459–485.
- Köder, F.** (2016). *Between direct and indirect speech: the acquisition of pronouns in represented speech* (Doctoral dissertation). Retrieved from <https://www.rug.nl/research/portal/files/28231090/Complete_thesis.pdf>.
- Köder, F., & Maier, E.** (2016). Children mix direct and indirect speech: evidence from pronoun comprehension. *Journal of Child Language*, **43**(4), 843–866.
- Köder, F., & Maier, E.** (2018). The advantage of story-telling: children's interpretation of reported speech in narratives. *Journal of Child Language*, **45**(2), 1–17.
- Lai, W., Lee, Y., & Lee, J.** (2010). Visiting doctors' offices: a comparison of Korean and Taiwanese preschool children's narrative development. *Early Education & Development*, **21**(3), 445–467.
- Leech, G. N., & Short, M. H.** (1981). *Style in fiction: a linguistic introduction to English fictional prose*. London/New York: Longman.
- Levey, S.** (2003). Represented dialogue and pragmatic particles in the narratives of preadolescents. *World Englishes*, **22**(3), 305–321.
- Li, C. N.** (1986). Direct and indirect speech: a functional study. In F. Coulmas (ed.), *Direct and indirect speech* (pp. 29–45). Berlin: Mouton de Gruyter.
- Li, C. N., & Thompson, S. A.** (1981). *Mandarin Chinese: a functional reference grammar*. Berkeley and Los Angeles: University of California Press.
- Lucy, J.** (1993). Metapragmatic presentational: reporting speech with quotatives in Yucatec Maya. In J. Lucy (ed.), *Reflexive language: reported speech and metapragmatics* (pp. 91–125). Cambridge: Cambridge University Press.
- Luo, G.** (2013). *Fating hudong zhong de lichang yanjiu [A study of stance-taking in courtroom interaction]*. (Unpublished doctoral dissertation). East China Normal University, Shanghai, China.
- Lysaker, J. T., & Miller, A.** (2012). Engaging social imagination: the developmental work of wordless book reading. *Journal of Early Childhood Literacy*, **13**(2), 147–174.
- Lysaker, J. T., Shaw, K., & Alicia, Z. A.** (2016). Emergent reading comprehension: social imagination and kindergarteners' readings of a wordless picture book. *Australian Journal of Language and Literacy*, **9** (3), 245–255.
- Manhardt, J., & Rescorla, L.** (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics*, **23**, 1–21.
- Mathis, T., & Yule, G.** (1994). Zero quotatives. *Discourse Processes*, **18**(1), 63–76.
- Mayes, P.** (1990). Quotation in spoken English. *Studies in Language*, **14**(2), 325–363.
- Norbury, C. F., & Bishop, D. V. M.** (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders*, **38**(3), 287–313.
- Nordqvist, Å.** (2001a). *Speech about speech: a developmental study on form and function of direct and indirect speech* (Doctoral dissertation). Retrieved from <<https://gupea.ub.gu.se/handle/2077/15182>>.

- Nordqvist, Å.** (2001b). The use of direct and indirect speech by 1/2- to 4-year-olds. *Psychology of Language and Communication*, 5(1), 21–45.
- Oliveira, M., & Cunha, D. A. C.** (2004). Prosody as marker of direct represented speech boundary. In *Proceedings of Speech Prosody* (pp. 263–266). Nara, Japan.
- Özyürek, A.** (1996). How children talk about a conversation. *Journal of Child Language*, 23, 693–714.
- Premack, D., & Woodruff, G.** (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4, 515–526.
- Semino, E., & Short, M.** (2004). *Corpus stylistics: speech, writing and thought. presentation in a corpus of English writing*. London: Routledge.
- Serratrice, L., Hesketh, A., & Ashworth, R.** (2015). The use of reported speech in children's narratives: a priming study. *First Language*, 35(1), 68–87.
- Shen, D.** (1991). Ye tan zhongguo xiaoshuo xushu zhong zhuan shu yu de dute xing [On the idiosyncrasies of modes of speech presentation in Chinese narrative fiction]. *Journal of Peking University (Philosophy and Social Science Edition)*, 27(4), 76–82.
- Slobin, D. I.** (1986). The acquisition and use of relative clauses in Turkic and Indo-European languages. In D. I. Slobin & K. Zimmer (eds.), *Studies in Turkish linguistics* (pp.273–294). Amsterdam: John Benjamins.
- Spronck, S.** (2012). Minds divided: speaker attitudes in quotatives. In I. Buchstaller & I. van Alphen (eds.), *Quotatives: cross-linguistic and cross-disciplinary perspectives* (pp.71–116). Amsterdam: John Benjamins.
- Tang, S.** (2005). *Hua yu zhi jiji pianzhang gongneng yanjiu [On discourse specification and its textual function]*. (Unpublished doctoral dissertation). East China Normal University, Shanghai, China.
- Tannen, D.** (2007). *Talking voices: repetition, dialogue, and imagery in conversational discourse* (2nd ed.). Cambridge: Cambridge University Press.
- Ukrainetz, T. A., Justice, L. M., Kaderavek, J. N., Eisenberg, S. L., Gillam, R. B., & Harm, H. M.** (2005). The development of expressive elaboration in fictional narratives. *Journal of Speech Language & Hearing Research*, 48, 1363–1377.
- Wu, H.** (2014). *Repetitive reportings in spontaneous Mandarin conversation*. (Unpublished doctoral dissertation). University of California, Los Angeles.
- Xiao, R., Rayson, P., & McEnery, T.** (2009). *A frequency dictionary of Mandarin Chinese: core vocabulary for learners*. London: Routledge.
- Xu, J.** (1996). Xushu wen zhong zhijie yinyu fenxi [Direct reported speech in narrative discourse]. *Language Teaching and Research*, 1, 52–66.
- Yue, Y.** (2011). *Hanyu de chuanxin fanchou ji qi yu xiangguan yuyan fanchou de hudong yanjiu [Evidentiality and its interaction with other linguistic categories in Mandarin Chinese]*. (Unpublished doctoral dissertation). Peking University, Beijing, China.
- Zhang, T.** (2010). *Jiyu yuliaoku de 3-6 sui hanyu ertong cihui yanjiu [Lexical research based on a 3-6-year-old Mandarin-speaking Chinese spoken corpus]*. (Unpublished doctoral dissertation). Shandong University, Jinan, China.
- Zhang, Y.** (2019). *Jiyu hanyu ertong yuliaoku de yuyan zhangai zhenduan xitong yanjiu [The establishment of a system for language disorder assessment: research based on Mandarin-speaking children's corpus]*. (Unpublished doctoral dissertation). East China Normal University, Shanghai, China.
- Zhang, F., McCabe, A., Ye, J., Wang, Y., & Li, X.** (2018). A developmental study of the narrative components and patterns of Chinese children aged 3–6 years. *Journal of Psycholinguistic Research*, 48(2), 477–500.
- Zhao, Y.** (1987). Xiaoshuo xushu zhong de zhuanshu [Reported speech in narrative fiction]. *Literature and Art Studies*, 5, 78–88.
- Zhu, M., Cao, F., & Zhang, R.** (1986). Youer dui zhishi daici de lijie [Children's understanding of deixis]. In M. Zhu. (ed.), *Ertong yuyan fazhan yanjiu [Research in child language development]* (pp. 104–113). Shanghai, China: East China Normal University Press.

Appendix A. Classification of representing forms with examples and translation

Representing form	Definition	Example
<i>Direct speech</i>	This is marked by a quotative index; its represented message, either a clause or a single word (e.g., an interjection or expression of politeness like <i>zài jiàn</i> ['goodbye']), features the perspective of the source speaker.	<p>ránhòu tā shuō zhè shì wǒ de qīngwā. then he say this be I NMLZ frog ‘Then he said: “This is my frog.”’</p> <p>ránhòu zhègè xiǎo nánhái gěi tā shuō zàijiàn. then this little boy to he say goodbye ‘Then this little boy said “goodbye” to him.’</p>
<i>Unframed direct speech</i>	Its content reflects the source speaker’s point of view but is free from a representing introducer.	<p>zhè wǒ de qīngwā, bùyào rēng rēngdiào duìbùqǐ. this I NMLZ frog not throw throw away sorry ‘“This is my frog. Do not throw it away. Sorry.”’</p>
<i>Indirect speech</i>	Its represented content shows the current speaker’s vantage point, which is marked by a preceding frame of introducer.	<p>tā shuō tā bù shì gùyì de. he say he not be intention NMLZ ‘He said that he did not do it intentionally.’</p>
<i>Voice representation</i>	This prototypically has only one clause that contains a verb of speech with or without a noun phrase following it. It gives a minimal reference to the occurrence of a verbal activity without specifying its details.	<p>nǎinǎi mà nàgè fúwùyuán le granny scold that waiter PFV ‘The old woman scolded that waiter.’</p>
<i>Mixed form</i>	It indicates a mingling of features of both the source speaker’s and the representor’s perspectives and, in most cases, involves mixes of deictic expressions.	<p>ránhòu tā shuō, zhè shì wǒ de then he say this be I NMLZ qīngwā, néng bùnéng huán gěi tā. frog can not can return give he ‘Then he said: “This is my frog”, and that if it was possible to give (it) back to him.’ In this example, the character involved asked the waiter to return his pet frog. <i>Wǒ</i> (‘my’) clearly marks the reported speaker’s point of view, but <i>tā</i> (‘he’), referring to reported speaker in this case, indicates the perspective of the reporting speaker or the narrator.</p>

(Continued)

Representing form	Definition	Example
<i>Ambiguous form</i>	Containing a quotative index, this form can be either direct speech or indirect speech due to the absence of features specifying the speaker's point of view in the represented message.	xiǎo nánhái shuō nǎ gè dōu xiǎng yào. little boy say which CLAS all want have 'The little boy said, (I/he) want(ed) everything.' In the above example, the narrator dropped the subject in the represented content. Due to the missing pronoun, the example becomes ambiguous: it could be interpreted as either direct speech (with the first-person pronoun being added) or indirect speech (with the third-person pronoun being added).
<i>Unclear utterance</i>	With neither a framing introducer nor any features of the source speaker's involvement in conversational exchanges, it is unclear whether the utterance is represented speech or thought, or whether it is a non-representing statement of the current speaker.	zhīhòu, tā yī chī, zěnmē yǒu yī after that she once eat how have one gè qīngwā de wèidào yā. CLAS frog NMLZ taste PRT 'After that, as soon as she started eating it, why did it taste like a frog?' The exclamatory and interrogative form of <i>zěn me yǒu yīgè qīngwā de wèidào yā</i> presents the features of direct speech and can be interpreted – due to the absence of the introductory clause of saying – as either quoted speech or thought.

Appendix B. Classification of representing signals with examples and translation

Representing signal	Definition	Example
<i>Shuō</i> 'say'	This is a typical generic verb and its equivalents in other languages are the most common lexical signals of SR.	ránhòu tā shuō "zhè shì wǒ de qīngwā." then he say this be I NMLZ Frog 'Then he said: "This is my frog."'
<i>Neutral verb</i>	A generic verb other than <i>shuō</i> 'say' that is also semantically unmarked.	xiǎolán gàosù tāmen bù néng qù. xiǎolán tell they not can go 'Xiaolan told them: "Can't go."'
<i>Non-generic verb</i>	A verb that characterizes the represented content by indicating either the illocutionary force or the manner of speaking.	nǎinǎi mà nàgè fúwùyuán le. granny scold that waiter PFV 'The granny scolded that waiter.'
<i>Adverb + de + representing verb</i>	This contains a speech verb and the accompanying construction of an adverb with the particle <i>de</i> , which offers a detailed description of the verb.	nà gè xiǎo, nà gè rén, nà gè that CLAS little that CLAS person that CLAS rén shēngqì de shuō : "zhè zhī qīngwā." person angry PRT say this CLAS frog 'That person angrily said: "This is a frog."'
<i>Verb₁ + verb₂</i>	Serial verb construction, where the representing verb – normally <i>shuō</i> 'say' – has a foregoing verb, either presenting the illocutionary force, or referring to a kinetic behavior just before speaking.	diàn lǐ de lǎobǎn zégùai tā shuō : restaurant inside NMLZ owner rebuke he say "nǐ zěnmé bù xiǎoxīn, bǎ wǒ de you how not careful PASS I NMLZ gǔ gěi nòng huài le." drum give make broken PRT 'The restaurant owner rebuked him by asking him, "Why are you so careless? You broke my drum."'

(Continued)

Representing signal	Definition	Example
		<p>tāmen dōu zhǐ zhe tā shuō: “nǐ they all point IPFV he say you nǐ nǐ wèishénme yào bǎ xiǎo qīngwā you you why want PASS small frog dài huí jiā?” bring back home ‘They all pointed at him and asked: “Why did you bring the small frog back with you?”’</p>
<i>Representing verb + object</i>	This is a verb phrase consisting of a representing verb and its object that is normally the production of speech such as words and sentences.	<p>ránhòu jiù shì tā gēn gǒugǒu shuō le then just be he to dog say PFV yī jù huà, nǐ yào guāiguāi zài jiā. one CLAS utterance you want obedient at home ‘Then he said a few words to the dog: “Be obedient and stay at home.”’</p>