Evaluation of cues in L1 Chinese input and output: 
A Competition Model approach to corpus data

Alessia Cherici, Yung-Yung Chang, Nozomi Tanaka (Indiana University)

Mandarin Chinese canonically follows an SVO word order but allows null arguments and flexible word order. It is also known for its sparse morphology. Thus, the input Chinese children receive with respect to verb-argument structure seems indeterminate, unlike for children acquiring languages with rigid word order like English, or languages with rich verbal inflection like Italian.

The present study uses corpus data to examine the information that helps Chinese children learn transitive verbs in their L1, namely, what kind of information is available in child-directed speech (input) and how they use that same information in their own production (output). We focus on the following “cues” that are used to encode transitive events in Chinese: word order (WO, canonically SVO), animacy contrast (AC, an animate noun is typically the subject and an inanimate noun is typically the object), the object marker \( ba \) (optionally marks the object in non-canonical SOV sentences, example 1), and the passive marker \( bei \) (also signals non-canonical word order, example 2).

(1) **Object marker \( ba \)**

\[ \begin{align*}
\text{Ni} & \quad \text{zenme} \quad \text{ba} \quad \text{chezi} \quad \text{nong} \quad \text{dao} \quad \text{le}?
\end{align*} \]

you how BA car make fall ASP

‘Why did you drop the car?’

(ADULT, from CHW)

(2) **Passive marker \( bei \)**

\[ \begin{align*}
\text{ni} & \quad \text{na} \quad \text{ge} \quad \text{pengyou} \quad \text{bei} \quad \text{ni} \quad \text{da}?
\end{align*} \]

you which CL friend BEI you hit

‘Which friend of yours has been beaten by you?’

(ADULT from CHW)

We adopted the methodology developed in previous studies (Kempe & MacWhinney, 1998 for German; Tanaka & Shirai, 2014 for Japanese) that used the Competition Model framework (Bates & MacWhinney, 1989). We estimated cue availability (how often the cue is present), reliability (how often the cue leads to a correct interpretation), and validity (the product of availability and reliability that serves as the primary determinant of cue strength) to evaluate the relative importance of these cues. For example, if both nouns are animate, animacy information itself is not useful to establish a transitive relation, and thus the AC cue is unavailable (3).

(3) **AC cue unavailable**

\[ \begin{align*}
\text{wo} & \quad \text{bu} \quad \text{hui} \quad \text{da} \quad \text{ni}
\end{align*} \]

I not MOD hit you

‘I’m not gonna hit you.’

(CHILD 3;6 from CHW)

If one noun is animate and one inanimate, the contrast is potentially useful and therefore the cue is available; however, the cue is reliable only when the subject is animate and the object inanimate (4). The AC cue is unreliable in the reverse pattern, as the reliance on animacy will lead to an incorrect interpretation (5).

(4) **AC cue available and reliable (inanimate noun bolded)**

\[ \begin{align*}
\text{Apo} & \quad \text{ye} \quad \text{hui} \quad \text{ti} \quad \text{qiu-qiu}
\end{align*} \]

grandma too MOD kick ball-ball

‘Grandma can kick the ball too.’

(CHILD 2;6 from WANG)

(5) **AC cue available but unreliable (inanimate noun bolded)**

\[ \begin{align*}
\text{Duzi} & \quad \text{zhang} \quad \text{le} \quad \text{chong}
\end{align*} \]

belly grow PFV bug

‘The belly grew bugs.’

(CHILD from WU)
We also looked at coalition (multiple cues working in combination) and competition (multiple cues working against each other).

So far, 1,221 child utterances and 2,298 adult utterances from the Taiwan Corpus of Child Mandarin (Cheung & Chang, 2011) in CHILDES (MacWhinney, 2000) have been analyzed. Preliminary results suggest that AC had the highest validity (adults: 76.36%, children: 74.61%), followed by WO (adults: 49.41%, children: 41.11%). Speakers also frequently used these two cues in combination, but AC was the sole reliable cue more often than WO. Ba had low availability (adults: 4.09%, children: 0.90%) but high reliability (adults: 100%, children: 100%). Bei was rarely available (adults: 1.35%, children: 2.70%).

### Table 1. Availability, reliability and validity of Word Order (WO), Animacy Contrast (AC), ba, and bei

<table>
<thead>
<tr>
<th></th>
<th>Children</th>
<th></th>
<th></th>
<th>Adults</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WO</td>
<td>AC</td>
<td>ba</td>
<td>bei</td>
<td>WO</td>
<td>AC</td>
<td>ba</td>
</tr>
<tr>
<td><strong>Cue availability</strong></td>
<td>42.75%</td>
<td>75.59%</td>
<td>2.70%</td>
<td>0.90%</td>
<td>50.93%</td>
<td>77.37%</td>
<td>4.09%</td>
</tr>
<tr>
<td><strong>Cue reliability</strong></td>
<td>96.17%</td>
<td>98.70%</td>
<td>87.88%</td>
<td>100.00%</td>
<td>97.02%</td>
<td>98.70%</td>
<td>100.00%</td>
</tr>
<tr>
<td><strong>Cue validity</strong></td>
<td>41.11%</td>
<td>74.61%</td>
<td>2.37%</td>
<td>0.90%</td>
<td>49.41%</td>
<td>76.36%</td>
<td>4.09%</td>
</tr>
</tbody>
</table>

![Figure 1. The validity of Word Order (WO), Animacy Contrast (AC), ba, and bei](image)

We posit the following cue strength hierarchy: AC>WO>ba>bei; that is, animacy is the most important cue in expressing transitive events in Chinese-speaking children’s input and output. While Li et al. (1993) suggested the hierarchy bei>AC>WO>ba for adult native speakers, our study is the first to report the hierarchy for these four cues for Chinese-speaking children. Ours is also the first study to draw on naturalistic corpus data to evaluate cue strengths in Chinese.

### References


