The Acquisition of In-Situ WH-Questions and WH-Indefinites in Jakarta Indonesian

Peter Cole*, David Gil®, Gabriella Hermon*, Uri Tadmor®

*University of Delaware, °Max Planck Institute for Evolutionary Anthropology

25th Boston University Conference on Language Development, Boston, MA, USA, 3.11.2000

1. Interrogatives and Indefinites

Haspelmath (1997, in preparation)

(1) **NORTHERN PAIUTE**

\[ \text{uso } hano mi\hat{a}T \]
he where go-PERF.PRTC

(i) "Where did he go?"
(ii) "He went somewhere"

(2) **YIDIPA**

\[ \text{wa}\text{ju } \text{walba } ya\text{gi\hat{a}} \]
INDEF/INTERROG-ERG rock-ABS split-PAST

(i) "Who split the rock?"
(ii) "Somebody split the rock"

"In common with most or all other Australian languages, Yidipa has a set of forms that can have either 'indefinite' or 'interrogative' force. In fact, the most appropriate translation would normally involve both notions; thus for [the above sentence] the storyteller provided a translation 'someone must have cut that rock — who did it?' Speakers of Australian languages usually consider vagueness to be a social fault (Dixon 1972:30), so that it is not unnatural for any indefinite specification to, at the same time, enquire for definite information about what is being referred to."
2. **What is Jakarta Indonesian?**

 Jakarta Indonesian is the general colloquial language of Jakarta, used in most everyday contexts for inter-ethnic and ethnically-neutral communication, and increasingly also for intra-ethnic communication; acquired naturally and completely at a young age by most children growing up in Jakarta; described by Wouk (1989, 1999).

 Jakarta Indonesian is not:

 - Standard Indonesian, used in more formal contexts in Jakarta and throughout Indonesia; acquired by children at a later age, often 'imperfectly', largely from the media and via conscious schooling.
 - Betawi Malay, the native dialect of the indigenous ethnic community of Jakarta, now a small minority of the total the population of Jakarta — described by Kähler (1966), Abdul Chaer (1976), Ikranagara (1980), Muhadjir (1981), Grijns (1991) and others.

 However, there exists a continuum of language varieties between Jakarta Indonesian and Standard Indonesian, and between Jakarta Indonesian and Betawi Malay.

3. **Some WH Forms in Jakarta Indonesian**

3.1. **Partial Lexicon of WH Forms**

(3) (a) *apa* 'what', 'do what'
(b) *apaan* 'what' -an AUGMENTATIVE
(c) *apanya* 'what of' -nya ASSOCIATIVE
(d) *siapa* 'who' si- PERSONAL
(e) *kenapa* 'undergo what', 'why', 'how' *kena* 'undergo'

(4) (a) *mana* 'which', 'where'
(b) *gimana* 'what circumstance','how' *< bagai* 'like'

(5) *kapan* 'when'

3.2. **Polyfunctionality: Some Non-Question Interpretations of WH Forms**

(6) (a) all WH forms indefinite ← *this paper*
(b) *apa* disjunction
(c) *apa* YN question
4. Indefinite Interpretations of WH Forms

4.1. The Basic Pattern

(7) The Basic Structure of WH Indefinites:

\[
\begin{array}{ccc}
\text{environment} & \text{form marker} & \text{WH companion} \\
\end{array}
\]

(a) Some Types of Licensing Environments:
(i) \(\emptyset\)
(ii) modal
(iii) negation
(iv) attributive modifier

(b) Types of Companion Markers:
(i) \(\emptyset\)
(ii) reduplication
(iii) conjuctive operator: \textit{aja} "just"

4.2. Adult Examples

4.2.1 Conjunctive Operator / No Licenser

(8) WH Question
Binatangnya apa aja?
star-ASSOC what CONJ.OP
[Experimenter asking child to list the animals that are inside a make-believe house]
"What animals are there?"

(9) Indefinite
Siapa aja
who CONJ.OP
[Asked who receives injections when sick, speaker answers]
"Just anyone"

4.2.2 No Companion Marker / Attributive Modification

(10) Indefinite
Siapa yang cepat juara
who REL fast champion
[Experimenter encouraging two children to finish their food]
"Whoever's fastest is the champion"
4.2.3 No Companion Marker / Pragmatics of Doubt

(11) **WH Question**
Terus ada apa lagi?
continue exist what CONJ.OP
[Experimenter asking child about things she sees on TV] "And what else is there?"

(12) **Indefinite**
Terus kan ada apa gitu, nggak tahu
continue NEG exist what like-DEM:DIST NEG know
[From a narrative about somebody who saw a ghost] "And then there was something else; I don’t know”

4.2.4 Reduplication / Negation, Conditional, ← this paper
Restrictive Conjunctive Operator, Yes-No Question, Surprise, Modal

(13) **Indefinite / Negation**
Gua nggak suka sama siapa-siapa
1:SG NEG like with RED-who
[Elicited] “I don’t like anybody”

(14) **Indefinite / Conditional**
Kalo lu punya apa-apa, kasi gua
if 2 have RED-what, give 1:SG
[Elicited] “If you have anything, give it to me.”

(15) **Indefinite / Restrictive Conjunctive Operator**
Dia makan apa-apa sendiri
3 eat RED-what REST.CONJ.OP
[Elicited] “He eats anything himself”

(16) **Indefinite / Surprise**
Kok lu ngasi tahu apa-apa ama polisi
SURPRISE 2 AG-give know RED-what with police
[Elicited] “I am surprised that you told anything to the police.”

(17) **Indefinite / Modal**
Dia boleh kawin ama siapa-siapa; gua nggak peduli
3 can marry with RED-who, 1:SG NEG care
[Elicited] “He can marry anyone, I don’t care”
4.3. Generalizations

(18) WH+aja is interpreted either as an indefinite or as a question, and does not require a licensing syntactic environment.

(19) A bare WH form may receive an indefinite interpretation licensed by an attributive modifier or when the context indicates doubt or uncertainty.

(20) Reduplicated WH forms with indefinite interpretations are ← this paper (Negative) Polarity Items, requiring an appropriate licensing environment (negation, conditional, conjunctive operator, yes-no question, surprise, modal, focusing etc.).

Coyaud and Hamou (1971)

(21) (a) quisquis  LATIN  Haspelmath (1997:179)
"who-RED" > "whoever"
(b) ka Leyka Ley  KHASI  Rabel (1961)
"RED-who" > "someone"
(c) ngana-ngana  WIRANGU  Hercus (1999:91)
"what/who-RED" > "anything/anyone"

6. Licensing Environments Cross-Linguistically

(22) *John ate any cake

(23) (a) John will eat any cake
(b) John can eat any cake
(c) [IMP] Eat any cake
(d) If John ate any cake (he'll regret it later)

(24) John didn't eat any cake
(25) John ate any cake he was given

A Semantic Map

<table>
<thead>
<tr>
<th>specific</th>
<th>specific</th>
<th>irreals</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>unknown</td>
<td>non-specific</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>question</th>
<th>indirect</th>
<th>direct</th>
</tr>
</thead>
<tbody>
<tr>
<td>negation</td>
<td>negation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>conditional</th>
<th>comparative</th>
<th>free choice</th>
</tr>
</thead>
</table>

7. Theories of How WH Questions and Indefinites are Related


STAR In situ WH forms with question interpretations are question operators which move to scopal position in LF.

STAR WH forms with indefinite interpretations and WH forms with question interpretations are not claimed to be related to each other synchronically.


STAR In situ WH forms with question interpretations are bound variables (not question operators).

STAR Question interpretations result from these variables being bound by null or overt question operators.

STAR Indefinite interpretations result from these same variables being bound by other null or overt indefinite operators, or by existential closure.
7.3. The Underspecification Analysis

✯ WH forms are neither operators nor bound variables. They are associated with a question feature $F$. The interpretation of $F$ is "properties in question", or "properties unknown and which need to be filled in by a contextually appropriate person".

✯ The feature $F$ is inherently projective. Accordingly, in the simple case, the $F$ feature projects upwards, to the root node of the sentence, thereby characterizing the sentence as a question.

✯ However, when the $F$ feature fails to project, it results in an indefinite interpretation. Depending on the properties of the indefinite in question, this may be due to the presence of an element like negation or modality, which constitutes the licensing environment for the indefinite interpretation.

(27) (a) Dia makan apa-apa?
            question interpretation
            3  eat   RED-what
            "What (all) did he eat?"

            F

            Dia     maken     apa-apa

(b) Dia nggak makan apa-apa
            indefinite interpretation
            3  NEG    eat    RED-what
            "He didn't eat anything"

            Dia    nggak    maken    apa-apa
8. Predictions for Acquisition

8.1. Predictions of Previous Theories

The LF WH-Movement Analysis

- In general, indefinite NPs present little difficulty for acquisition (issues of relative scope aside, see Kramer 2000)
- In contrast, LF movement (which, according to this analysis, underlies the question interpretations of in-situ WH forms), may be acquired late. WH-in situ may come in later than moved WH in French (Crisma 1992) and children are known have problems with wide scope interpretation of quantified NPs under negation (but perhaps for different reasons, see Musolino 1988)
- Therefore, the LF WH-Movement Analysis may be interpreted as predicting that the indefinite interpretations of WH forms will be acquired before their question interpretations.

The Bound Variable Analysis

- Question and indefinite interpretations of WH forms result from the presence of different binding operators for WH variables. Neither of these binding operators is intrinsically more complex than the other.
- Therefore, the Bound Variable Analysis predicts that, all other things being equal, indefinite interpretations of WH forms will be acquired at the same time as question interpretations.

8.2. Predictions of the Underspecification Analysis

- WH questions derive from the projection of the feature $F$ to clausal level.
- Indefinite interpretations of WH forms may require the presence of a licensing environment when the feature $F$ fails to project to the clausal level.
- Accordingly, indefinite interpretations of WH forms subject to licensing requirements are of greater complexity than their question counterparts, since the conditions on licensing must be learned by the child.
- Therefore, the Underspecification Analysis predicts that the question interpretations of WH forms will be acquired before their indefinite interpretations.
8.3. The Facts

Question interpretations of WH forms are acquired before the corresponding indefinite interpretations.

These facts do not corroborate the LF WH-Movement Analysis (though they are not necessarily inconsistent with that hypothesis).

These facts do not provide corroboration for the Bound Variable Analysis (though they are not necessarily inconsistent with that hypothesis).

These facts, while possibly explicable by other factors, provide corroboration for the Underspecification Analysis.

9.0 Indefinite forms in child language

9.1. The Naturalistic Data Collection Project

- Eight children (four boys, four girls), longitudinal study
- Weekly or bi-weekly one-hour video-recording sessions.
- 55200 records, containing 7664 WH-forms

Table 1: Age ranges and number of utterances
(total of 33226 child utterances)

<table>
<thead>
<tr>
<th>subjects</th>
<th>number of child utterances</th>
<th>number of sessions</th>
<th>age range in months</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIM</td>
<td>327</td>
<td>4</td>
<td>18-19</td>
</tr>
<tr>
<td>HIZ</td>
<td>6247</td>
<td>28</td>
<td>19-28</td>
</tr>
<tr>
<td>MIC</td>
<td>1049</td>
<td>22</td>
<td>24-29</td>
</tr>
<tr>
<td>RIS</td>
<td>8554</td>
<td>33</td>
<td>20-33</td>
</tr>
<tr>
<td>PRI</td>
<td>4618</td>
<td>8</td>
<td>31-33</td>
</tr>
<tr>
<td>LAR</td>
<td>3573</td>
<td>17</td>
<td>34-40</td>
</tr>
<tr>
<td>IDO</td>
<td>4950</td>
<td>15</td>
<td>40-46</td>
</tr>
<tr>
<td>PIT</td>
<td>3903</td>
<td>11</td>
<td>52-57</td>
</tr>
</tbody>
</table>
## 9.2 The acquisition of indefinites and WH-questions

Table 2: WH Forms with question and indefinite Interpretations

<table>
<thead>
<tr>
<th>form</th>
<th>meaning</th>
<th>adult</th>
<th>child</th>
<th>age of first appearance (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEG (a)pa-(a)pa</td>
<td>a. as a frozen expression: 'it's OK, doesn't matter'</td>
<td>69</td>
<td>35</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>b. 'nothing', 'anything'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEG ...apa-apa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEG... siapa-siapa</td>
<td>'nobody', 'anybody'</td>
<td>2</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td>NEG... ke mana-mana</td>
<td>'to nowhere, anywhere'</td>
<td>4</td>
<td>8</td>
<td>31</td>
</tr>
<tr>
<td>NEG...kenapa-kenapa</td>
<td>'for no reason'</td>
<td>0</td>
<td>2</td>
<td>54</td>
</tr>
<tr>
<td>NEG ngapa-ngapain</td>
<td>'not do anything'</td>
<td>0</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>apa aja</td>
<td>'whatever'</td>
<td>2</td>
<td>5</td>
<td>54</td>
</tr>
<tr>
<td>siapa aja</td>
<td>'whoever'</td>
<td>1</td>
<td>3</td>
<td>55</td>
</tr>
<tr>
<td>ngapain aja</td>
<td>'do anything, do whatever'</td>
<td>0</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>siapa (plus YANG attributive clause)</td>
<td>'whoever'</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>apa without reduplication or AJA</td>
<td>'something'</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total indefinite</td>
<td></td>
<td>83</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>total question</td>
<td></td>
<td>4896</td>
<td>1864</td>
<td>18</td>
</tr>
<tr>
<td>% of indefinite use</td>
<td></td>
<td>1.7%</td>
<td>3.0%</td>
<td></td>
</tr>
<tr>
<td>total number of utterances counted</td>
<td></td>
<td>27340</td>
<td>27860</td>
<td></td>
</tr>
</tbody>
</table>
Children acquire indefinites with reduplication at a later age than they acquire WH-questions.

Children acquire the correct licensing conditions from the start.

Children seem to be 'cautious' learners, and do not widen the domain in which indefinites are licensed without robust positive evidence from the data. Even at a later age, the only licensor for indefinites is overt negation. (cf., van der Wal, 1996  for the acquisition of NPIs in Dutch)

Table 3: Use of indefinites in younger and older children

<table>
<thead>
<tr>
<th>forms</th>
<th>meaning</th>
<th>26 - 39</th>
<th>&gt; 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEG (a)pa-(a)pa [gapapa]</td>
<td>as a frozen expression: 'it is OK', 'it does not matter'</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>NEG ADA apa-apa / siapa-siapa</td>
<td>'there is not anything/anyone'</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>NEG V apa-apa</td>
<td>'nobody', 'anybody'</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>NEG... ke mana-mana</td>
<td>'to nowhere', 'to anywhere'</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>NEG...kenapa-kenapa</td>
<td>'for no reason'</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>NEG ngapa-ngapain</td>
<td>'not do anything'</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>apa aja</td>
<td>'whatever'</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>siapa aja</td>
<td>'whoever'</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>ngapain aja</td>
<td>'do anything, do whatever'</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>total indefinite</td>
<td></td>
<td>14</td>
<td>42</td>
</tr>
<tr>
<td>total question</td>
<td></td>
<td>844</td>
<td>1020</td>
</tr>
<tr>
<td>% of indefinite use</td>
<td></td>
<td>1.6%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Summary: There are no productive uses of indefinites before age 26 months. Indefinites with RED begin to come in at about 26 months, with 14 examples of indefinites in the 26-39 months age range. Some kids never use indefinites, even by age 39 (like HIZ), others do.

Precursors: at age 25 we have 2 examples of indefinites, but both are total imitations of a preceding adult utterance.
(28) Ndak *apa-apa*, deh!
   NEG red-what EXCL
   [repeating EXP's statement which is unheard by the recorder ]
   "It's okay!"

(29) *Siapa aja*
   who CONJ.OP
   [repeating adult]
   "Anybody"

☆ All examples are grammatical uses, with overt negation licensing the reduplicated indefinites.

☆ Most uses at this age are of the frozen form *nggak apa-apa*, pronounced [ga papa], used with the meaning "It is OK, nothing is the matter".

(30) Ini ama ini *nggak apa-apa*?
   DEM:PROX with DEM:PROX NEG RED-what
   [putting together two parts of a car]
   "Is this okay, this one and this one?"

(31) *Nggak apa-apa*
   NEG RED-what
   [squabbling over a cassette]
   "It's okay"

☆ The same child also uses the construction more productively, with the form *ada* intervening between the NEG and the indefinite or with other indefinites:

(32) *Nggak ada apa-apanya* (repeated twice)
   NEG exist RED-what-ASSOC
   [Mother asks what's in it, child responds]
   "There is nothing in it"

(33) *Nggak ke mana-mana*
   NEG to RED-where
   [Adult asks child where she's going; child answers]
   "Nowhere"

(34) *Enggak, aku nggak pergi ke mana-mana*
   NEG 1:SG NEG go to RED-where
   [Adult asks child where she's going; child answers]
   "No, I'm not going anywhere"

(35) *Nggak mo dibawa xxx mana-mana*
   NEG want PAT-bring xxx RED-where
   [Experimenter asks child where she's taking her toy to; child answers]
   "I'm not taking it anywhere"

☆ Summary: Children under 26 months show no indefinite uses at all (except for 2 repetitions by RIS25), and children under 39 months have a limited range of indefinites.
The syntactic knowledge needed for indefinites (reduplication and negation) are in place by this age. For example, we find 121 examples of negation, with various negative forms and use of both verbal and constituent negation. Reduplication is found in 35 examples, and occurs freely with N+N forms.

In contrast to the limited uses of indefinites, WH-questions are used productively from a very young age. The table shows uses of WH-questions in various grammatical 'frames' from early on.

Table 4: Younger and older children: major uses of WH-questions
Figures are based on 49,407 records (1762 child wh-questions total), and seven kids age ranges 18-65 months.

<table>
<thead>
<tr>
<th>Type of clause</th>
<th>age &lt; 26 months</th>
<th>age 26-39 months</th>
<th>age 40-65 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single wh</td>
<td>43% (124)</td>
<td>49% (261)</td>
<td>20% (207)</td>
</tr>
<tr>
<td>Dem+wh</td>
<td>30% (87)</td>
<td>22% (117)</td>
<td>29% (300)</td>
</tr>
<tr>
<td>NP+wh</td>
<td>19% (55)</td>
<td>15% (81)</td>
<td>21% (217)</td>
</tr>
<tr>
<td>act/prop*+wh</td>
<td>7% (21)</td>
<td>13% (67)</td>
<td>22% (225)</td>
</tr>
</tbody>
</table>

* this includes WH-qu with various verbal and adjectival forms.

The youngest kids have fewer verbal wh-questions, but we do get at least 7 different verbs: makan ‘eat’, mimik ‘drink’ masak ‘cook’, maem ‘eat’ ndengerin ‘listen’ pake ‘use/wear’, setel ‘turn on’

Productive us of WH-questions can be shown even for kids like HIZ (who does not get indefinites even by age 39) and for children like MIC, who also does not exhibit indefinites:

(36) Minum apa ya? drink what yes [Experimenter asks child what he wants to drink; child reflects] "What will I drink?"

(37) Bikin apa ya? make what yes [Child playing with crayons, wonders what to draw] "What should I make?"

(38) Mana taronya? where put-ASSOC [Child carrying a chair, wondering where to put it] "Where should I put it?"

(39) Mau bikin apa? MIC-29
want make what
[Playing with crayons, child asks adult]
"What are you making?"

9.3 Explanations for the acquisitional 'delay'

✯ Is the late and somewhat restricted use of indefinites due to low frequency in the input?

✯ Other low frequency WH-words can be early acquired (kok and ngapain):

(40) Kok ininya hilang?
how.come how.come DEM:PROX-ASSOC disappear
[Looking at an empty box which should contain one of her books]
"How come this one's thing is missing?"

(41) Oh suster kok pakenya ...
EXCL nurse how.come use-ASSOC
[Looking at a picture book (unfinished utterance)]
"How come the nurse is wearing... ?"

(42) Apa...lagi ngapain?
what more AG-what-APPL
[Looking at a picture of a man praying]
"What's he doing?"

Summary: children under 26 never use indefinites, but already use a variety of WH forms in various productive ways. Between 26 and 39, indefinites begin to come in but are quite restricted: In contrast many different WH words are used, in many productive structures, and even fairly low frequency words (such as kok and ngapain) are picked up the children productively.

9.4 The acquisition of indefinites in older kids

✯ After 39 months, (age range 40 to 65) some more indefinites come in. We now get 52 examples, with some new verb forms are used with indefinites. We also find 9 uses of indefinites with the companion marker aja.

(43) Dek Idonggak jadi apa-apa deh, nggak jadi, naek perahu
TRU-younger.sibling Ido NEG become RED-what EXCL-NEG become go.up boat
[Playing with toy boat; children assuming various roles; child's elder brother says he wants to be captain; experimenter asks child what he wants to be; he answers]
"I don't want to be anyone, just to ride the boat"

(44) Enggak, nggak dikasi apa-apa, nggak dikasi
NEG NEG PAT-give RED-what NEG PAT-give
[Experimenter asks child what he got at a birthday party the day before, child answers]
"No, I didn't get anything, anything"
(45) Minum apa aja
   drink what \textsc{conj.op}
   [Experimenter asks child what she wants to drink; child answers]
   "I'll drink just anything"

9.5 Conclusions:

★ Overall Summary: while WH-questions are productive and widely used from early on, very few examples of indefinite uses of WH-forms can be observed in children under age 40. Even in the >40 group in our study, the indefinite uses are quite limited.

★ The late use of indefinites is to be attributed to the fact that children need to learn the licensing conditions for the indefinite uses of WH-questions.

★ Negation is the only licensor used by children in our data for reduplicated indefinites. Children do not exhibit the use of indefinites with attributive clauses, nor do they have examples of modals or dubitative context as licensors.

★ The fact that indefinites come in after WH-questions does not in itself pose a problem for theories which attempt to derive WH-questions and indefinites from a single source. The data is directly predicted to fall out from any theory which analyses the indefinites as wh-forms which need a special negative licensor.

Acknowledgements:

Data collection:
Dini Andarini, Ferdinand Okki Kurniawan, Liana Husein, Like Raskova Octoberlina, Regina Yanti, Yokebed Trigawati

Data analysis and technical assistance:
Sarah Chakrawati, Guy Sharett, Antonia Soriente, Brad Taylor, Yassir Nasanius
REFERENCES