1. Preliminaries

*Wer, wie, was?*
*Der, die, das!*
*Wieso, weshalb, warum?*
*Wer nicht fragt, bleibt dumm!*
(German Sesame Street)

This talk is a progress report on an investigation into the semantic distinctions that are made in questionwords. This investigation can be seen as a cross-linguistic *Wortfeld* analysis (Trier 1931) of a closed class of items.

The subject of this talk will be the minimal number of questionwords a language might have (‘how low can you go?’). Do all languages have questionwords? Which categories are distinguished within the set of questionwords? Which distinctions are minimally made? For a preliminary report of the diversity of possible categories asked for by questionwords among the world’s linguistic diversity, see Cysouw (2004a)

1) Kinds of interrogative ‘words’

   – uninflected free form
   – inflected form (inflecting like nominal, verb, or adjective)
   – bound forms (always clitic-like?)

2) Typology of morphological structure of interrogative words

1a) unanalyzable morpeme (*how*, but also *who*)
1b) diachronically analyzable (*why, weshalb*)
2a) analyzable but nonsense semantically (*wofür*)
2b) partly analyzable (‘raspberry’-type)
3a) transparantly derived from another questionword (*how much*)
3b) identical to another questionword (*how long? vs. how did you do it?*)
2. **Transparancy**

Some semantic domains are more commonly expressed by an unanalyzable morpheme than others.

These empirically major categories overlap, but are not identical to most proposed ontologies (e.g. Aristotle,\(^1\) Heine *et al.* 1991a: 159; Jackendoff 1983: 52-3)

In general, there appears to be a high innovation rate for interrogative pronouns, also resulting in many near synonyms (e.g. German *wieso, weshalb, warum*).

(3) Major interrogative categories (estimates of cross-linguistic frequency based on a preliminary, not well-sampled set of languages, see Cysouw 2004a).

- PERSON (*who*) ± 95% unanalysable
- THING (*what*) ± 95% unanalysable
- PLACE (*where*) ± 90% unanalysable
- SELECTION (*which*) ± 60% unanalysable
- QUANTITY (*how much*) ± 60% unanalysable
- MANNER (*how*) ± 40% unanalysable
- TIME (*when*) ± 40% unanalysable
- all others fall below 10%

(4) Major pathways of derivation of questionwords

---

\(^1\) See Aristotle’s *Categories*, part 4, as translated by E. M. Edghill – available at [http://classics.mit.edu/Aristotle/categories.1.1.html](http://classics.mit.edu/Aristotle/categories.1.1.html)
3. **Extreme Transparancy?**

Also called ‘pure transparancy’ (Muysken & Smith 1990). In such systems, there is only one questioning element, all others are derived from it.

(5) 18th Century Sranan (English based Creole, Surinam, Bruyn 1993)

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hu-NOUN</td>
<td>'which', 'what sort of'</td>
</tr>
<tr>
<td>(hu)-summa</td>
<td>'who'</td>
</tr>
<tr>
<td>(hu)-sanni</td>
<td>'what'</td>
</tr>
<tr>
<td>hu-dissi</td>
<td>'which'</td>
</tr>
<tr>
<td>hu-sorte</td>
<td>'what sort of'</td>
</tr>
<tr>
<td>hu-peh</td>
<td>'where'</td>
</tr>
<tr>
<td>hu-sei</td>
<td>'where'</td>
</tr>
<tr>
<td>hu-tem</td>
<td>'when'</td>
</tr>
<tr>
<td>hu-menni</td>
<td>'how much'</td>
</tr>
<tr>
<td>hu-fasi</td>
<td>'how'</td>
</tr>
<tr>
<td>hu-fa</td>
<td>'how, why'</td>
</tr>
<tr>
<td>va-hu-heddi</td>
<td>'why'</td>
</tr>
</tbody>
</table>

Such systems are rare overall (cf. Indo-Pakestani Sign Language, Zeshan 2000: 155-7; Zeshan 2004: 23; and Kenya Swahili Pidgin, Heine et al. 1991b: 57). Almost complete transparancy is more common. Only one questionword is not transparent in e.g. Ewe (Westermann 1930: 164-5; Pasch 1995: 79-80), Nambikuara (Kroeker 2001: 17-9), and Pirahà (Everett 1991: 80-3). Widespread transparancy is not confined to Pidgins/Creoles, though it might be more common in such languages than in others. Note that the English system does not count as transparent, because *wh-* is not a morpheme

(6) English

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>wh-o</td>
<td>*th-o</td>
</tr>
<tr>
<td>wh-at</td>
<td>th-at</td>
</tr>
<tr>
<td>wh-ich</td>
<td>*th-ich</td>
</tr>
<tr>
<td>wh-ere</td>
<td>th-ere</td>
</tr>
<tr>
<td>wh-en</td>
<td>th-en</td>
</tr>
<tr>
<td>how</td>
<td>(th-us)</td>
</tr>
<tr>
<td>why</td>
<td>*th-y</td>
</tr>
</tbody>
</table>
4. Extreme Ambiguity?

Givón claims there to be only one questionword in Asheninka, an Arawakan language from Peru. This questionword *tsica* appears to be used in the same contexts as all different English questionwords.

(7) Asheninka (Givón 2001: 304-5, citing D. Payne, p.c.)

a. *tsika* *i-tim-i-ka* *iri-_ta*

   WH 3MASC-be-TAM-Q 3MASC-there
   ‘Who is he (that one)?’

b. *tsika* *o-pait-a-ka* *h-a-ake-ri*

   WH 3FEM-name-TAM-Q 3MASC-take-TAM-REL
   ‘What did he take?’

c. *tsika* *p-a-ake-ro-ka*

   WH 2-get-TAM-3FEM-Q
   ‘Where did you get it (fem)?’

d. *tsika* *i-ka_t-aaiti-ro-ka* *choclo*

   WH 3MASC-say-TAM-3FEM-Q corn
   ‘How do you say “corn”?’

e. *tsika* *i-karat-i-ka* *iri-ka*

   WH 3MASC-be-TAM-Q 3MASC-here
   ‘How much is it?/How many are there?’

In fact, the structure of Asheninka turns out to be a special case of transparency. There is hardly any ambiguity possible. I investigated all 206 content questions that appeared in an Asheninka text collection (Anderson 1985/1986). It turns out that the verb following the questionword *tsica* plays a determining role in the meaning of the question (for details, see Cysouw 2004b)
### Asheninca content questions (from Cysouw 2004b)

<table>
<thead>
<tr>
<th>Structure</th>
<th>Meaning</th>
<th>No. of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>tsica + NP</td>
<td>be where</td>
<td>9</td>
</tr>
<tr>
<td>tsica + V</td>
<td>where</td>
<td>4</td>
</tr>
<tr>
<td>tsica + -saic-</td>
<td>live where</td>
<td>4</td>
</tr>
<tr>
<td>tsica + -iyaat-</td>
<td>go where</td>
<td>7</td>
</tr>
<tr>
<td>tsica + -a(g)-</td>
<td>take from where</td>
<td>8</td>
</tr>
<tr>
<td>tsica + -quen-</td>
<td>go where</td>
<td>5</td>
</tr>
<tr>
<td>tsica + -quen-</td>
<td>will do how</td>
<td>3</td>
</tr>
<tr>
<td>tsica + -quen-</td>
<td>can do how</td>
<td>5</td>
</tr>
<tr>
<td>tsica + -cara-</td>
<td>how much/many</td>
<td>3</td>
</tr>
<tr>
<td>tsica + -tzim- + NP</td>
<td>be who</td>
<td>2</td>
</tr>
<tr>
<td>tsica + -tzim- + relative clause</td>
<td>which of them</td>
<td>2</td>
</tr>
<tr>
<td>tsica + -cant-</td>
<td>say what</td>
<td>4</td>
</tr>
<tr>
<td>tsica + -cant-</td>
<td>do what/what happened</td>
<td>8</td>
</tr>
<tr>
<td>tsica + -cant- + relative clause</td>
<td>how</td>
<td>12</td>
</tr>
<tr>
<td>tsica + -cant- + relative clause</td>
<td>why</td>
<td>5</td>
</tr>
<tr>
<td>tsica + -pait-</td>
<td>call how</td>
<td>8</td>
</tr>
<tr>
<td>tsica + -pait- + relative clause</td>
<td>who</td>
<td>4</td>
</tr>
<tr>
<td>tsica + -pait- + relative clause</td>
<td>what/which</td>
<td>12</td>
</tr>
<tr>
<td>tsica + -pait- + relative clause</td>
<td>why</td>
<td>15</td>
</tr>
<tr>
<td>tsica + -pait- + relative clause</td>
<td>how</td>
<td>1</td>
</tr>
<tr>
<td>-pait- + relative clause</td>
<td>who</td>
<td>6</td>
</tr>
<tr>
<td>-pait- + relative clause</td>
<td>what</td>
<td>16</td>
</tr>
<tr>
<td>-pait- + relative clause</td>
<td>why</td>
<td>40</td>
</tr>
<tr>
<td>-pait- + relative clause</td>
<td>how</td>
<td>1</td>
</tr>
<tr>
<td>-pait- + relative clause</td>
<td>how could (exclamation)</td>
<td>4</td>
</tr>
<tr>
<td>ipaitaca</td>
<td>how are you/what’s up</td>
<td>4</td>
</tr>
<tr>
<td>tsica + relative clause</td>
<td>(various meanings)</td>
<td>4</td>
</tr>
<tr>
<td>Other interrogatives</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>No translation</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>206</strong></td>
</tr>
</tbody>
</table>
(9) Summary of interrogative constructions in Asheninca

<table>
<thead>
<tr>
<th>Construction</th>
<th>Interrogative category</th>
</tr>
</thead>
<tbody>
<tr>
<td>tsica</td>
<td>PLACE ‘where’</td>
</tr>
<tr>
<td>tsica -tzim-</td>
<td>PERSON ‘who’</td>
</tr>
<tr>
<td>(tsica) -pait-</td>
<td>relative clause VERB ARGUMENT ‘who/what/which’</td>
</tr>
<tr>
<td></td>
<td>REASON ‘why’</td>
</tr>
<tr>
<td>tsica -quen-</td>
<td>MANNER ‘how’</td>
</tr>
<tr>
<td>tsica -cant-</td>
<td>MANNER ‘how’</td>
</tr>
<tr>
<td>tsica -cara-</td>
<td>QUANTITY ‘how much’</td>
</tr>
<tr>
<td>tsicapaite</td>
<td>TIME ‘when’</td>
</tr>
</tbody>
</table>


*tsica* pi-pait-a-ca
WH 2-call-TAM-Q
‘¿Cómo te llamas?’ (How are you called?/What is your name?)


a. tsica ipaitaca poc atsi-ri iroñaaca
WH come-TAM-REL now
‘¿Quién viene?’ (Who is coming?)

b. ipaitaca pi-caim i ri incaaran-qui
WH 2-call-TAM-REL LOC
‘¿A quién estabas llamando?’ (Who are you calling?)


a. paitaca p-amemana-tzi-ri-ca jaca no-yovite-qui
WH 2-search-TAM-REL-Q here 1-casserole-LOC
‘¿Qué estabas buscando en mi olla de barro?’
(What are you looking for in my casserole?)

b. tsica opaitaca ant-aque-mi-ri
WH do-TAM-1-REL
‘¿Qué te ha pasado?’ (What happened to you?)
5. **Attested Ambiguities**

‘The number and kind of distinctions which QWs [Questions Words, MC] may or may not reflect in terms of those existing elsewhere in a give language vary considerably from language to language, but at least one constrast appears to be nearly universal: Q-pronouns show a human/nonhuman or, in a few cases, an animate/inanimate dichotomy.’ (Ultan 1978: 229)

(13) PERSON=THING

- Baltic:
  - Latvian (Nau 1998, 1999)
  - Lithuanian (Lindström 1995: 314)
- Khasi (Austro-Asiatic, India, Rabel 1961: 68-9)
- !Xóo (Tuu, Namibia/Botswana, Güldemann forthcoming)
- Mekens, Sak_rabit dialect (Tupi, Brazil, Galucio 2001: 166-8)
- Paumarí (Arawá, Brazil, Chapman & Derbyshire 1990: 203-16)
- Ika (Chibchan, Colombia, Frank 1990: 82-6)
- Arawakan:
  - Achaguna (Colombia, Wilson 1992: 26, 125-7)
  - Apurína (Brazil, Facundes 2000: 365)
  - Bare (Venezuela, Aikhenvald 1995: 25)
  - Nomatsiguenga (Peru, Shaver 1996: 37, 40, 169)
  - Terena (Brazil, Eastlack 1968: 7-8; Ekdahl & Butler 1979: 190-4)
  - Warekena (Venezuela, Aikhenvald 1998: 261, 325-6)
- Mataco-Guaicuruan:
  - Maká (Paraguay, Gerzenstein 1994: 178)
  - Mataco (Paraguay, Viñas Urquiza 1974: 106-7)
  - Mocoví (Argentina, Grondona 1998: 162)
  - Toba (Argentina, Klein 2001: 23)

(14) MANNER=QUANTITY

- Asmat (Central and South New Guinea, Irian Jaya, Voorhoeve 1965: 157-161)
- Nabak (Huon Finisterre, Papua New Guinea, Fabian *et al.* 1998: 32)
- Wardaman (Gunungguan, Australia, Merlan 1994: 156, 590)
- Sanuma (Yanomam, Brazil, Borgman 1990: 66-72)
- Apalai (Carib, Brazil, Koehn & Koehn 1986: 56-60)
- Mekens (Tupi, Brazil, Galucio 2001: 166-180)
- Urubu-Kaapor (Tupí-Guaraní, Brazil, Kakumasu 1986: 353-5)

---

2 Sango is quoted by Ultan (1978: 229), but this is not substantiated by the sources (Samarin 1967: 74-5, 217; Thornell 1997: 76).
(15) Asmat (Voorhoeve 1965: 160)

a. *tirîf ucîm emamis*
   letters WH lie
   ‘How many letters are lying there?’

b. *makîk ucîm êmatamcêm*
   fish line WH will you throw it
   ‘How will you throw out the fish line?’

(16) Urubu-Kaapor (Kakumasu 1986: 354)

a. *myja ne tapi’ir ne ere-juka katu*
   WH 2SG.PRON tapir 2SG.PRON 2SG-kill well
   ‘How do you kill a tapir?’

b. *myja ere-juka*
   WH 2SG-kill
   ‘How many did you kill?’

(17) THING=REASON
   Migili (Niger-Congo/Platiod, Nigeria, Stofberg 1978: 131, 134)

a. *á gbé (i)bô_*
   3SG.PRON buy WH
   ‘What did he buy?’

b. *ó bé (i)bô_*
   2SG.PRON come WH
   ‘What have you come for?’

(18) THING=QUALITY
   Manam (Austronesian/Oceanic, Papua New Guinea, Lichtenberk 1983: 403)

a. *ra_ána stûa-lo _u-zâza-i*
   WH store-LOC 2SG-buy-3SG
   ‘What did you buy in the store?’

b. *ra_ána warîge áne-_a _u-wáur-i*
   WH rope INSTR-FOC 2SG-fasten-3SG
   ‘What (kind of) rope was it that you fastened it with?’
(19) Some other possible ambiguities

THING=MANNER
– Mocovi (Mataco-Guaicuran, Argentina, Grondona 1998: 162)

PLACE=TIME
– Terêna (Arawakan, Brazil, Ekdahl & Butler 1979: 190-4)

QUANTITY (MASS)=TIME
– Hup (Maku, Brazil; P. Epps, p.c.)

6. **Conclusions:**

_ synchronically, transparency of questionwords is ubiquitous
_ diachronically, there is a high ‘turn-over rate’ of questionwords
_ extreme transparency is rare
_ extreme ambiguity does not seem to exist
_ ambiguities among the major categories are rare
_ the ambiguity *who* = *what* appears to be the most common among them
7. References


