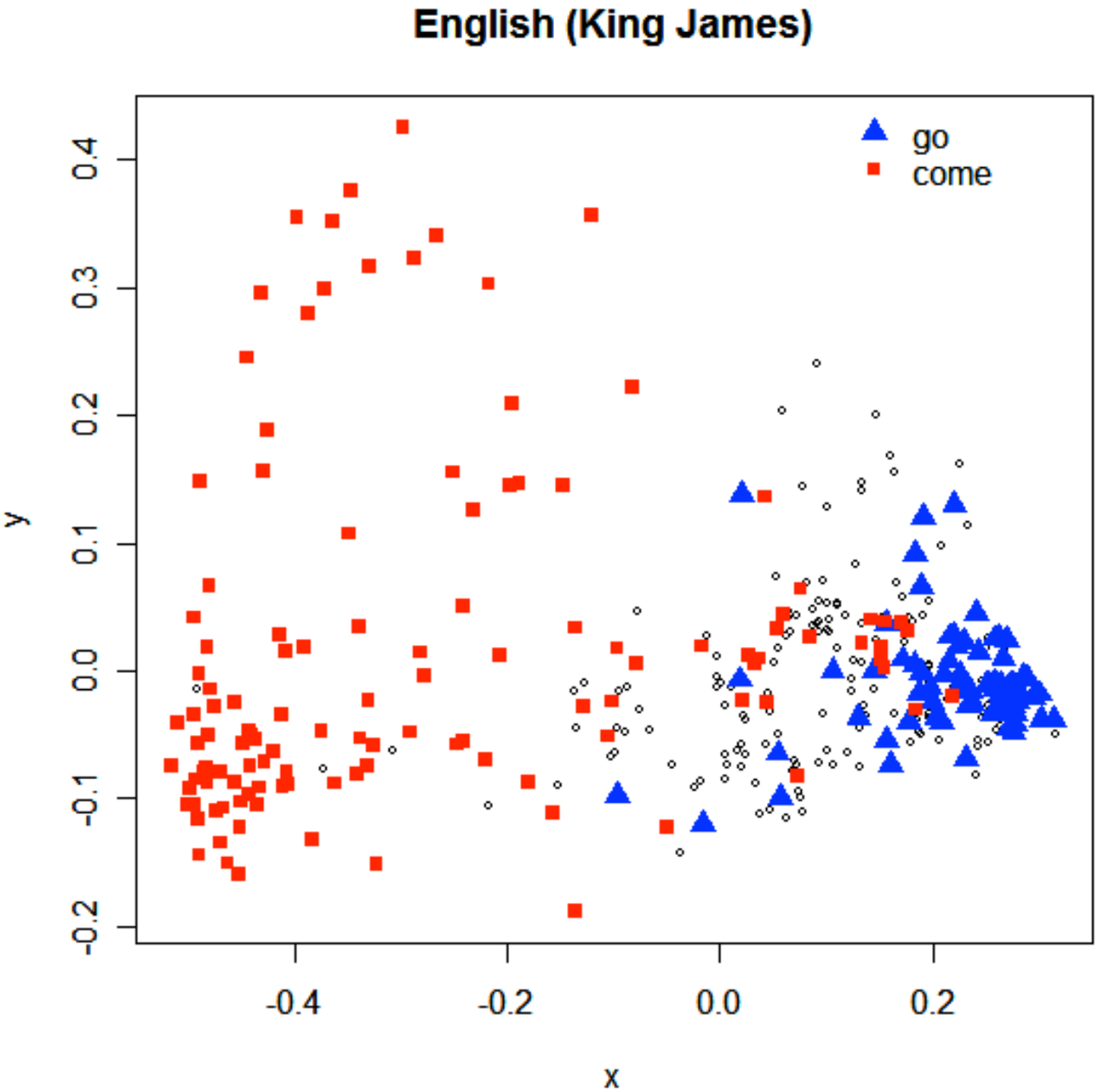
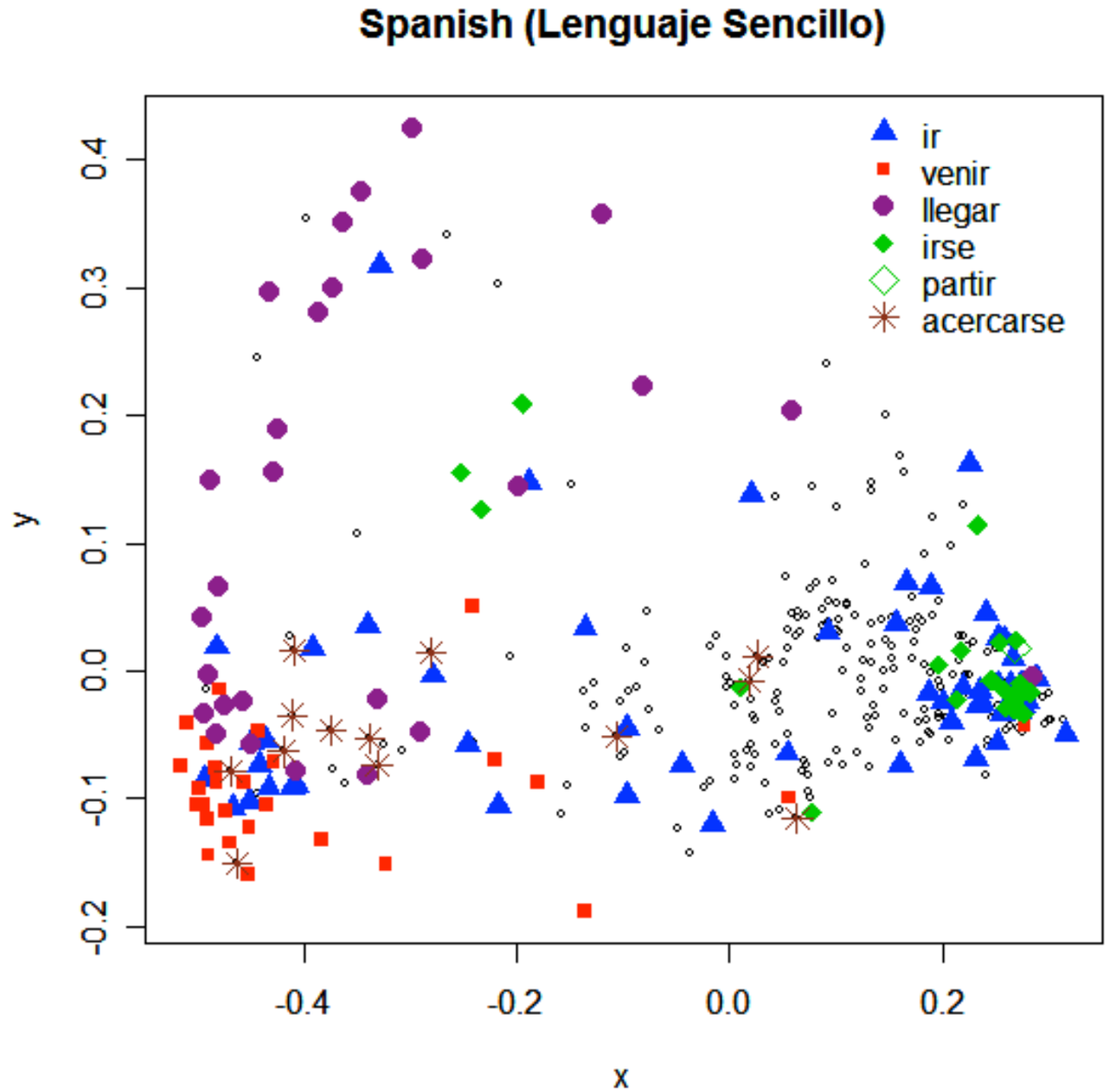
Where is the direction?

Indefinite's semantic map revisited

Michael Cysouw
LMU Munich





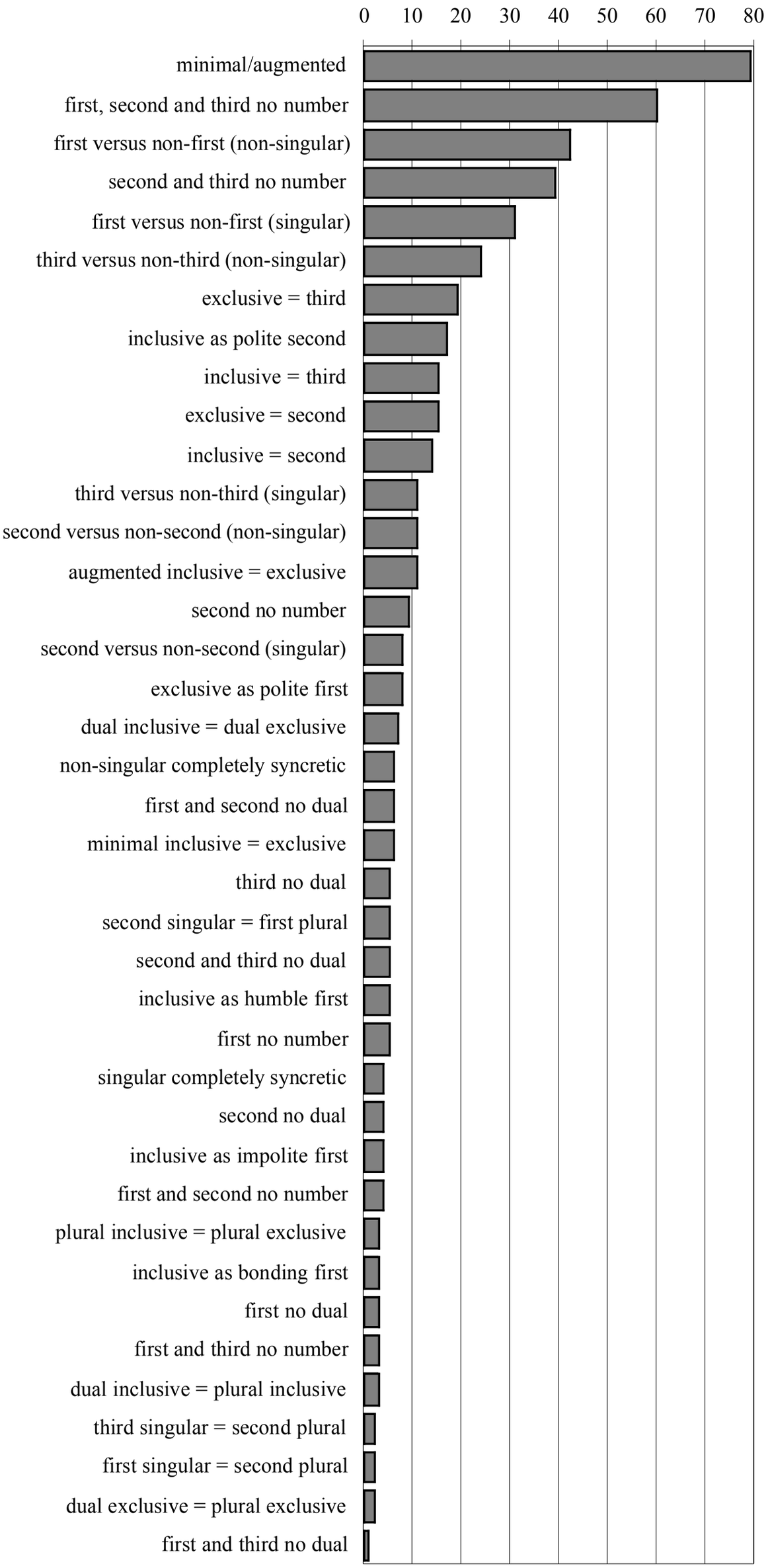
Haspelmath, Martin (1997) *Indefinite Pronouns*. OUP.

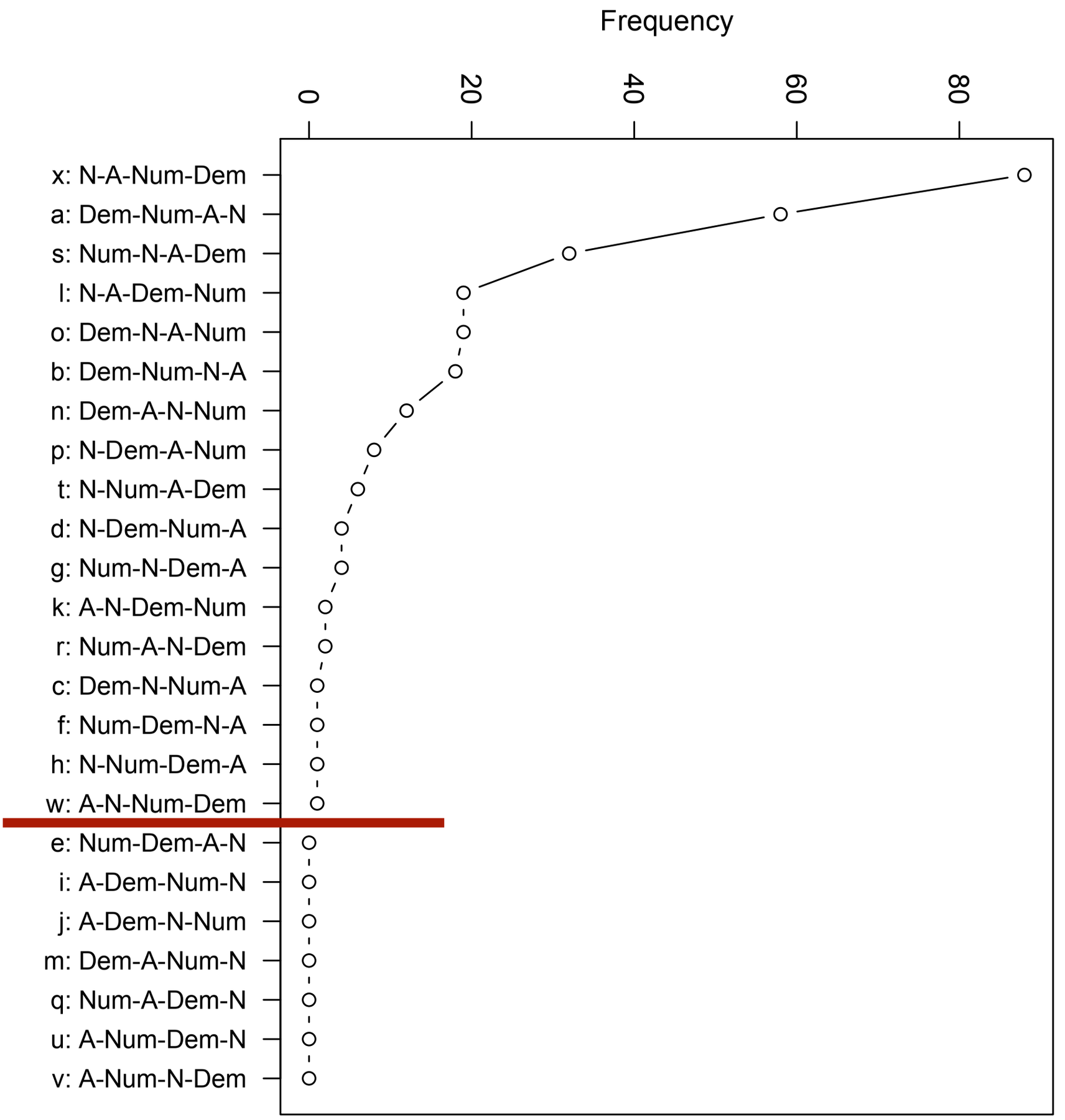
What is the purpose of such semantic maps ?

- Description of cross-linguistic variation
- Prediction of possible vs. impossible linguistic structures
- Modeling language change

Possibility or Probability ?

- Can we really distinguish possible from impossible languages ?
- No: because attested vs. unattested is a very fragile observation !
- Frequent vs. infrequent is a much more robust observation


0
10
20
30
40
50
60
70
80
minimal/augmented
first, second and third no number
first versus non-first (non-singular)
second and third no number
first versus non-first (singular)
third versus non-third (non-singular)
exclusive = third
inclusive as polite second
inclusive = third
exclusive = second
inclusive = second
third versus non-third (singular)
second versus non-second (non-singular)
augmented inclusive = exclusive
second no number
second versus non-second (singular)
exclusive as polite first
dual inclusive = dual exclusive
non-singular completely syncretic
first and second no dual
minimal inclusive = exclusive
third no dual
second singular = first plural
second and third no dual
inclusive as humble first
first no number
singular completely syncretic
second no dual
inclusive as impolite first
first and second no number
plural inclusive = plural exclusive
inclusive as bonding first
first no dual
first and third no number
dual inclusive = plural inclusive
third singular = second plural
first singular = second plural
dual exclusive = plural exclusive
first and third no dual


Frequency
0
20
40
60
80
x: N-A-Num-Dem
a: Dem-Num-A-N
s: Num-N-A-Dem
l: N-A-Dem-Num
o: Dem-N-A-Num
b: Dem-Num-N-A
n: Dem-A-N-Num
p: N-Dem-A-Num
t: N-Num-A-Dem
d: N-Dem-Num-A
g: Num-N-Dem-A
k: A-N-Dem-Num
r: Num-A-N-Dem
c: Dem-N-Num-A
f: Num-Dem-N-A
h: N-Num-Dem-A
w: A-N-Num-Dem
e: Num-Dem-A-N
i: A-Dem-Num-N
j: A-Dem-N-Num
m: Dem-A-Num-N
q: Num-A-Dem-N
u: A-Num-Dem-N
v: A-Num-N-Dem

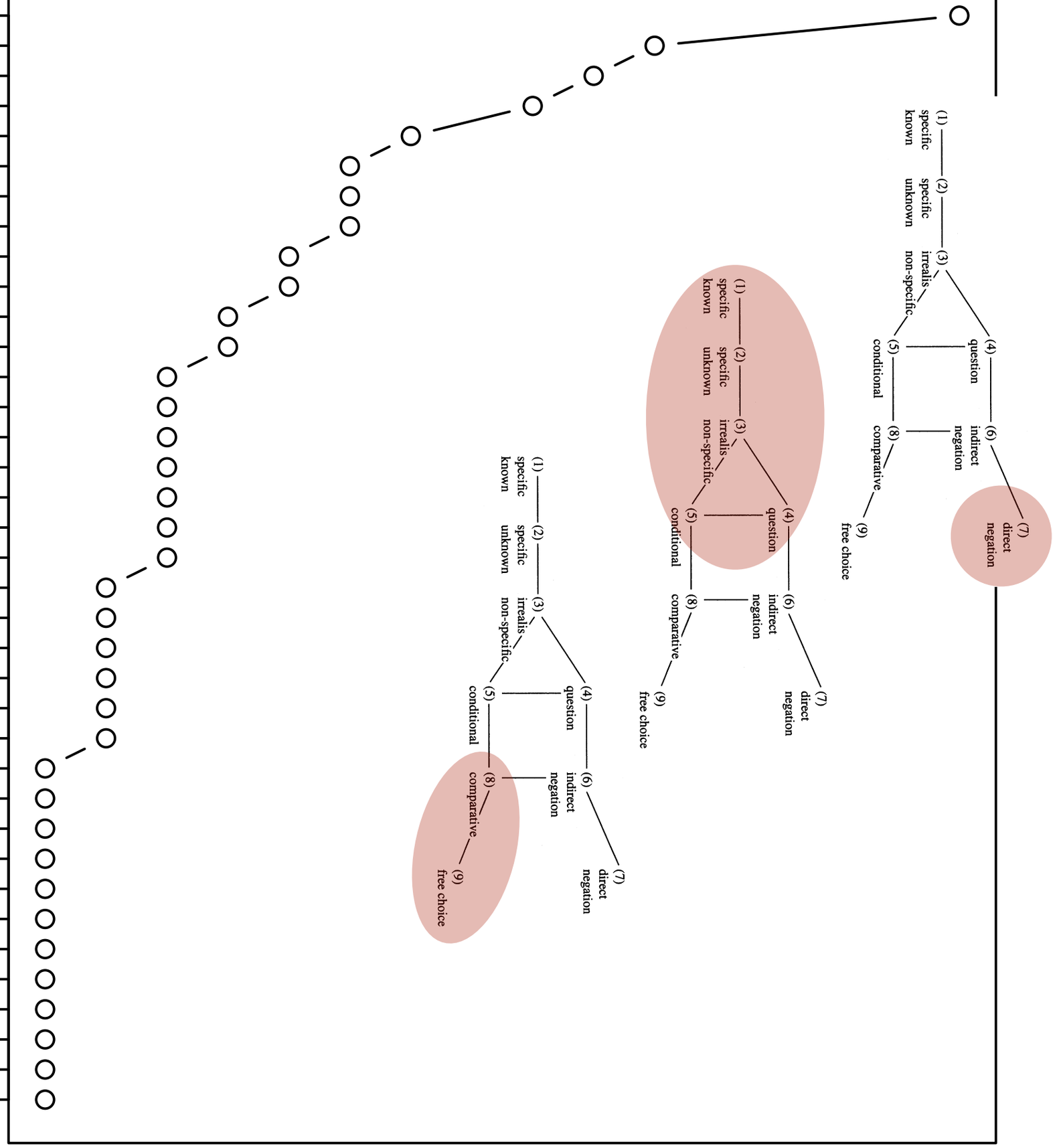
Frequency

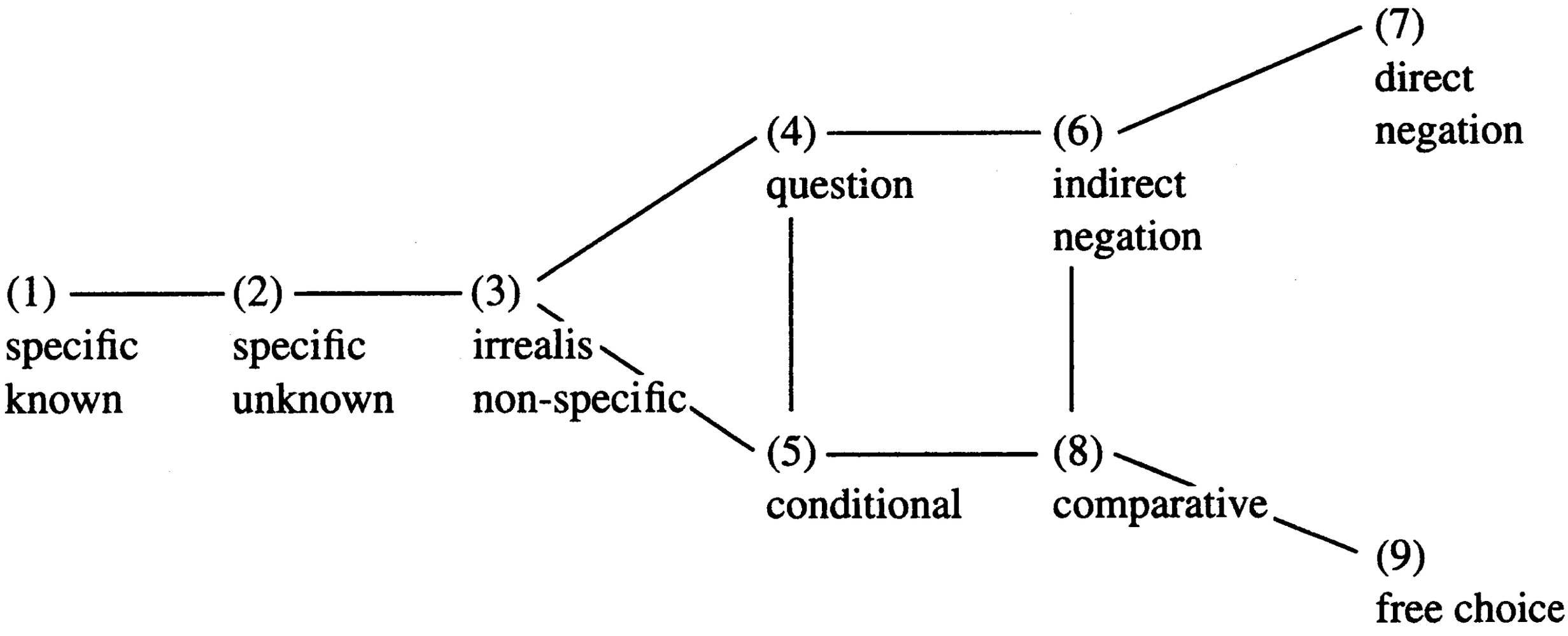
5

10

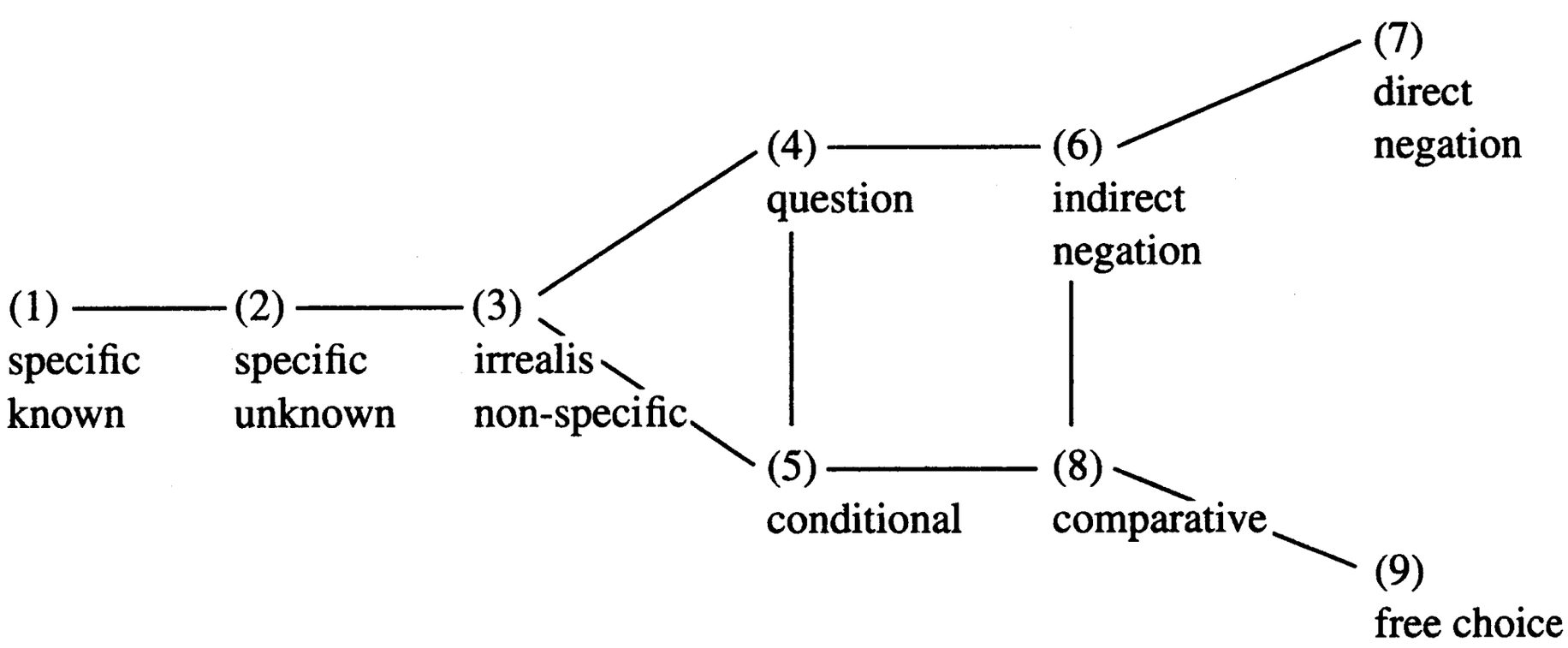
15

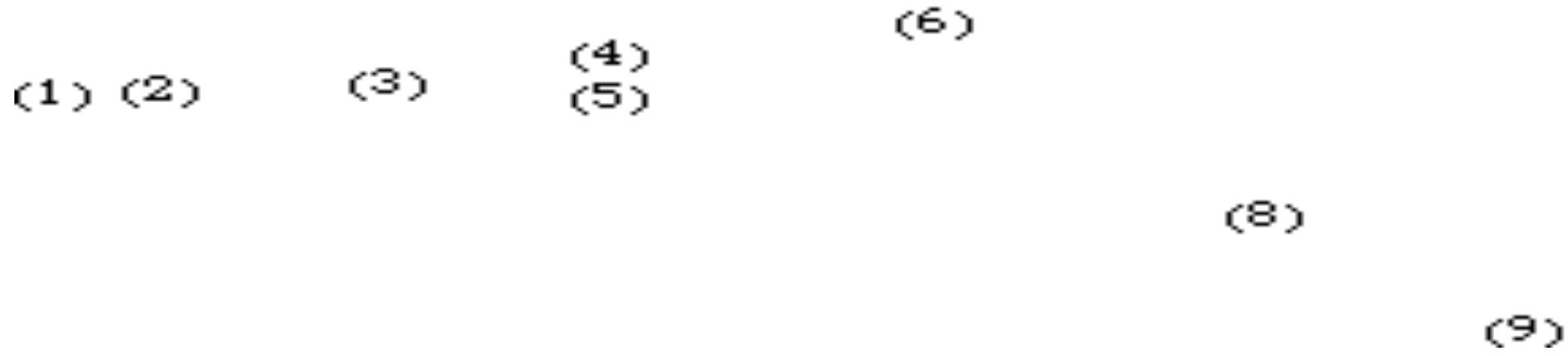
000000100
111110000
000000011
111111000
000000001
000001100
100000000
110000000
000111111
011110000
000001111
111111100
000001110
000101100
000111110
001110000
001111010
001111100
001111111
000000000
000010011
000011011
000111010
000111011
111000000
000001011
000011010
000101010
000101110
000110011
001111000
001111011
010000000
011000000
011111011
011111111
111111110





Haspelmath, Martin (1997) *Indefinite Pronouns*. OUP.

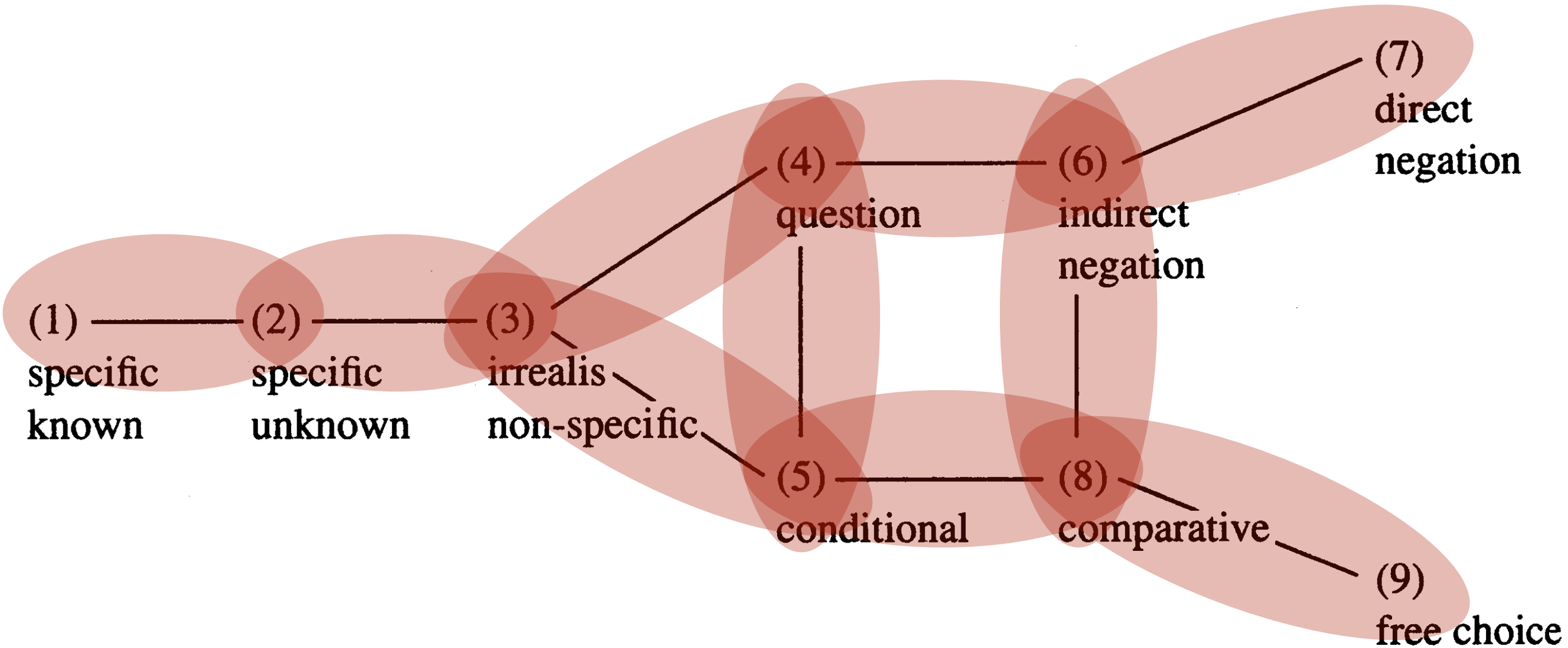




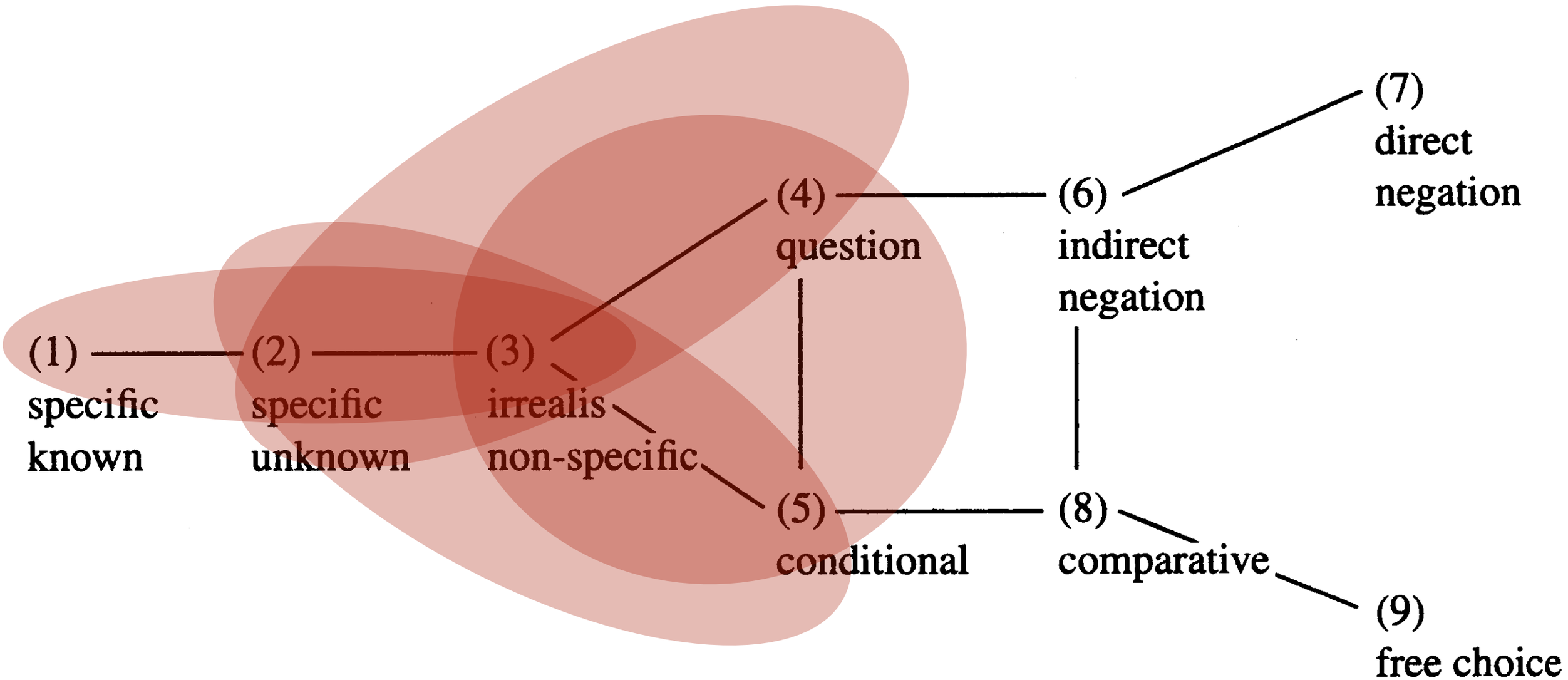
Cysouw, Michael (2001) Review of "Indefinite Pronouns". *Journal of Linguistics* 37(3), 607-612.

Adding lines?

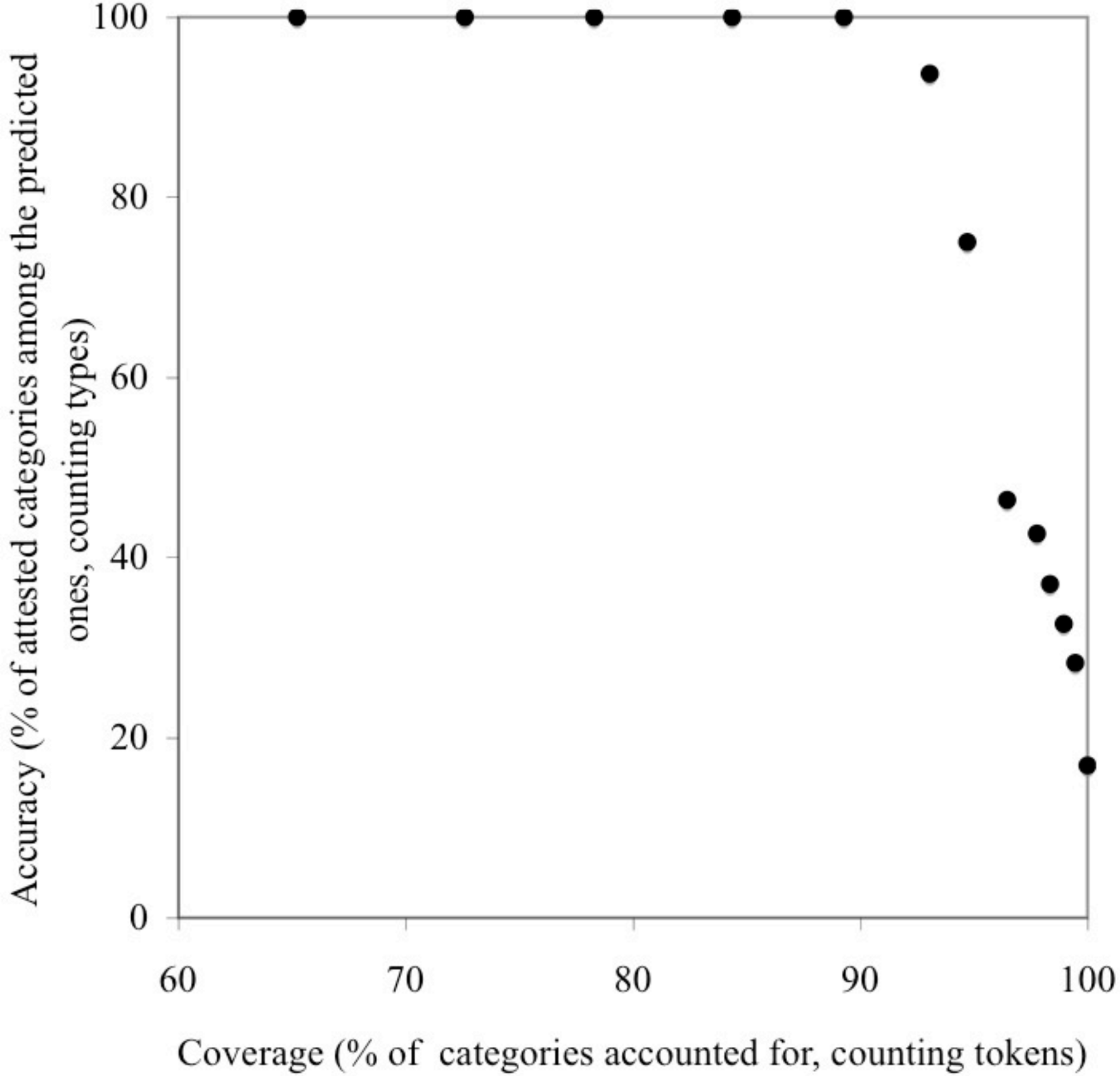
- Lines predict categories
- The more lines, the more spurious predictions are made
- Lines should be added balancing *coverage* (all attested should be included) vs. *accuracy* (all predicted should be attested)

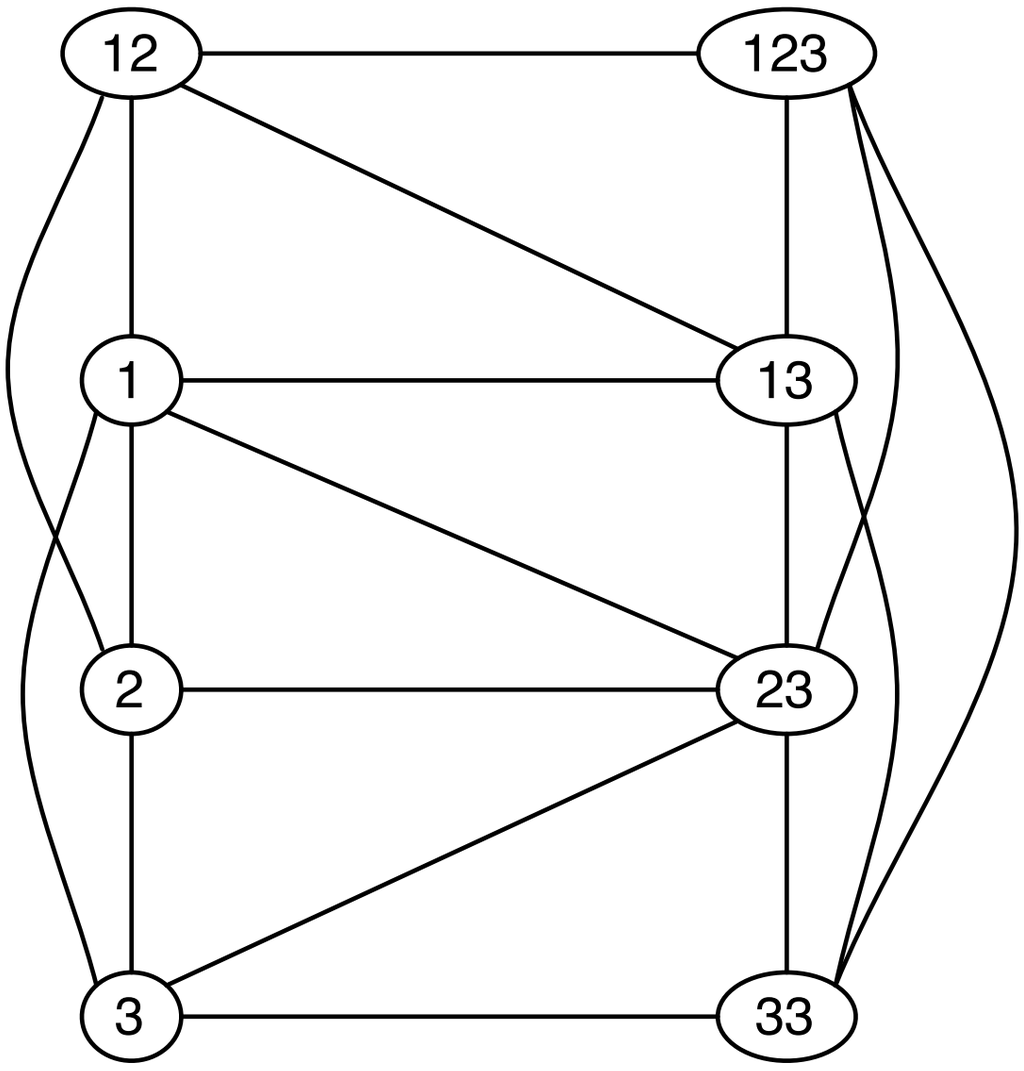


Haspelmath, Martin (1997) *Indefinite Pronouns*. OUP.

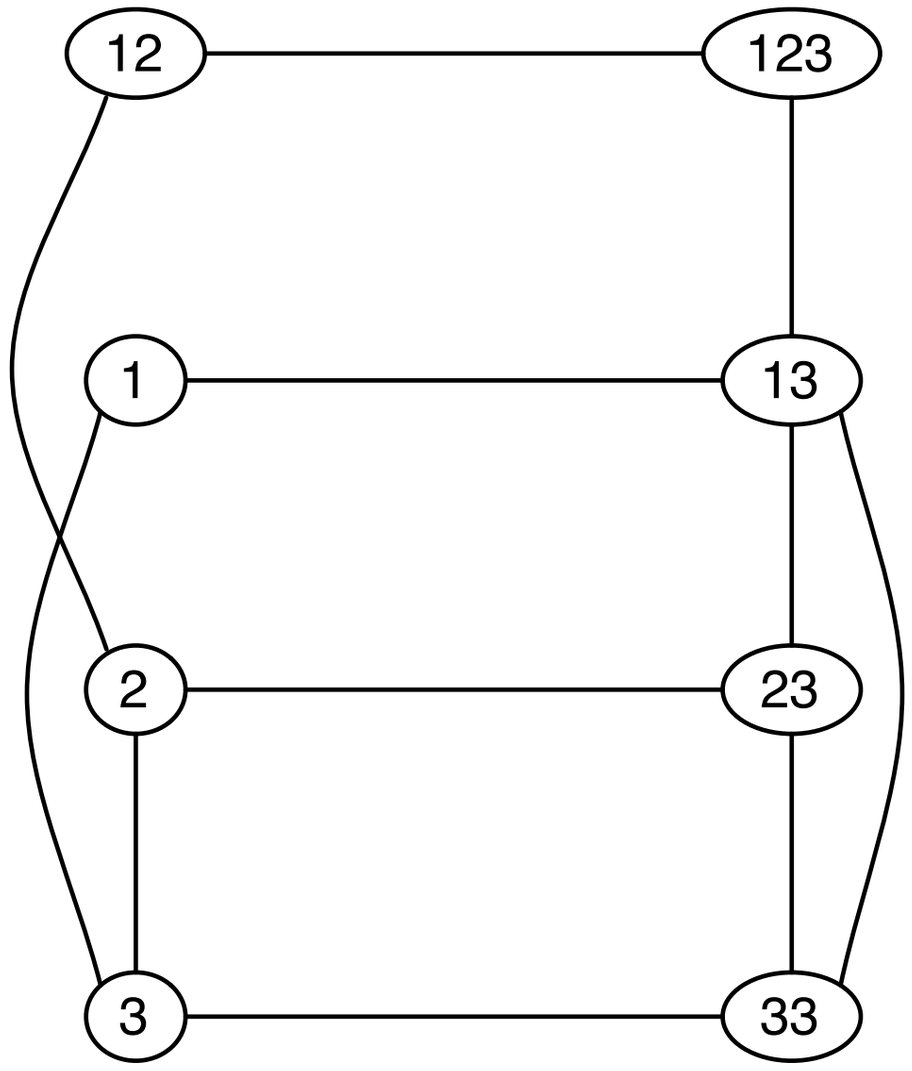


Haspelmath, Martin (1997) *Indefinite Pronouns*. OUP.


100
80
60
40
20
0
60
70
80
90
100
Accuracy (% of attested categories among the predicted ones, counting types)
Coverage (% of categories accounted for, counting tokens)



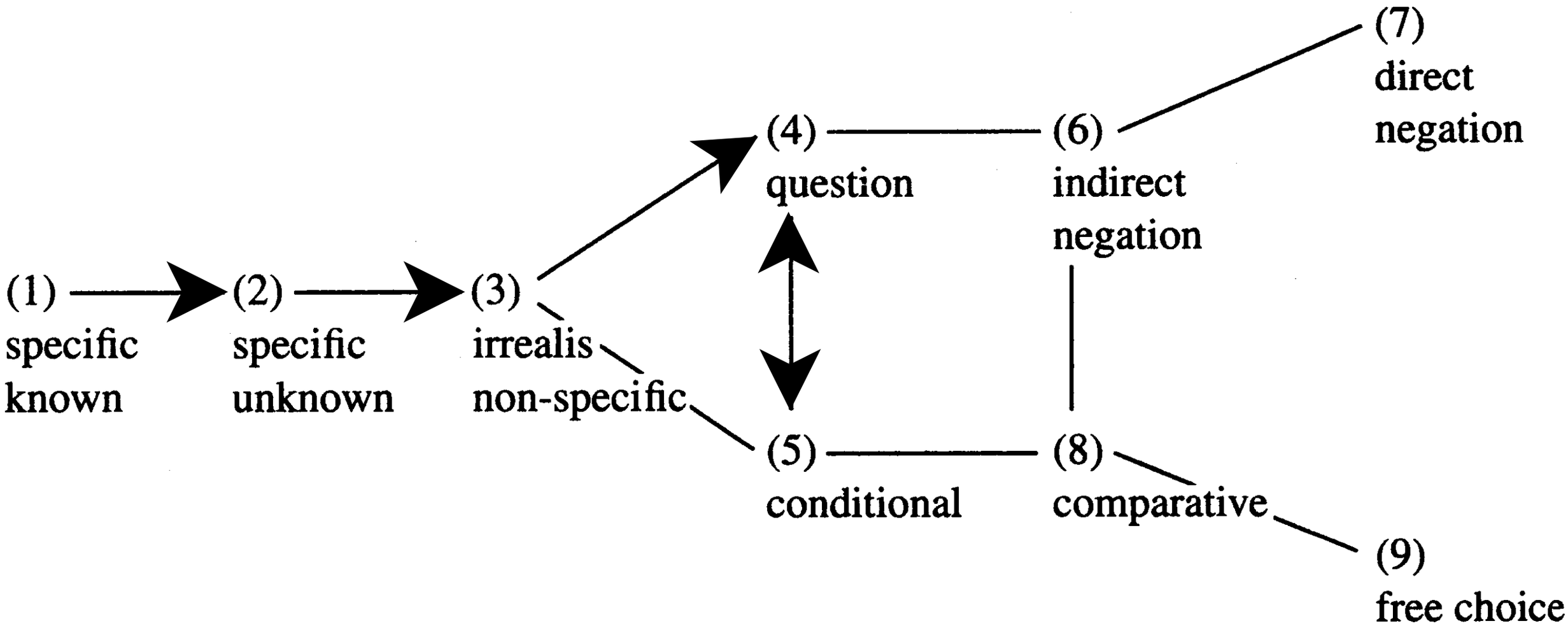
Maximum Coverage



Balancing Coverage and Accuracy

Adding direction

- Lines appear to be easily extended to arrows, adding direction to the map
- But same problem of possible vs. probable applies: many changes are attested, but not all are equally likely
- The more detailed the semantic map, the more problematic the usage of arrows becomes



Haspelmath, Martin (1997) *Indefinite Pronouns*. OUP.

Estimating Directions

- German *irgendjemand* contains *jemand*
- Interpret this asymmetry as a sign of past change *jemand* → *irgendjemand*
- Generalize this as asymmetric relation:
  Overlap / Size of form
- *jemand* → *irgendjemand* (6/6 = 1.0)
  irgendjemand → *jemand* (6/12 = 0.5)
- *nobody* → *anybody* (4/6 = 0.67)
  anybody → *nobody* (4/7 = 0.57)

direct.negation

indirect.negation

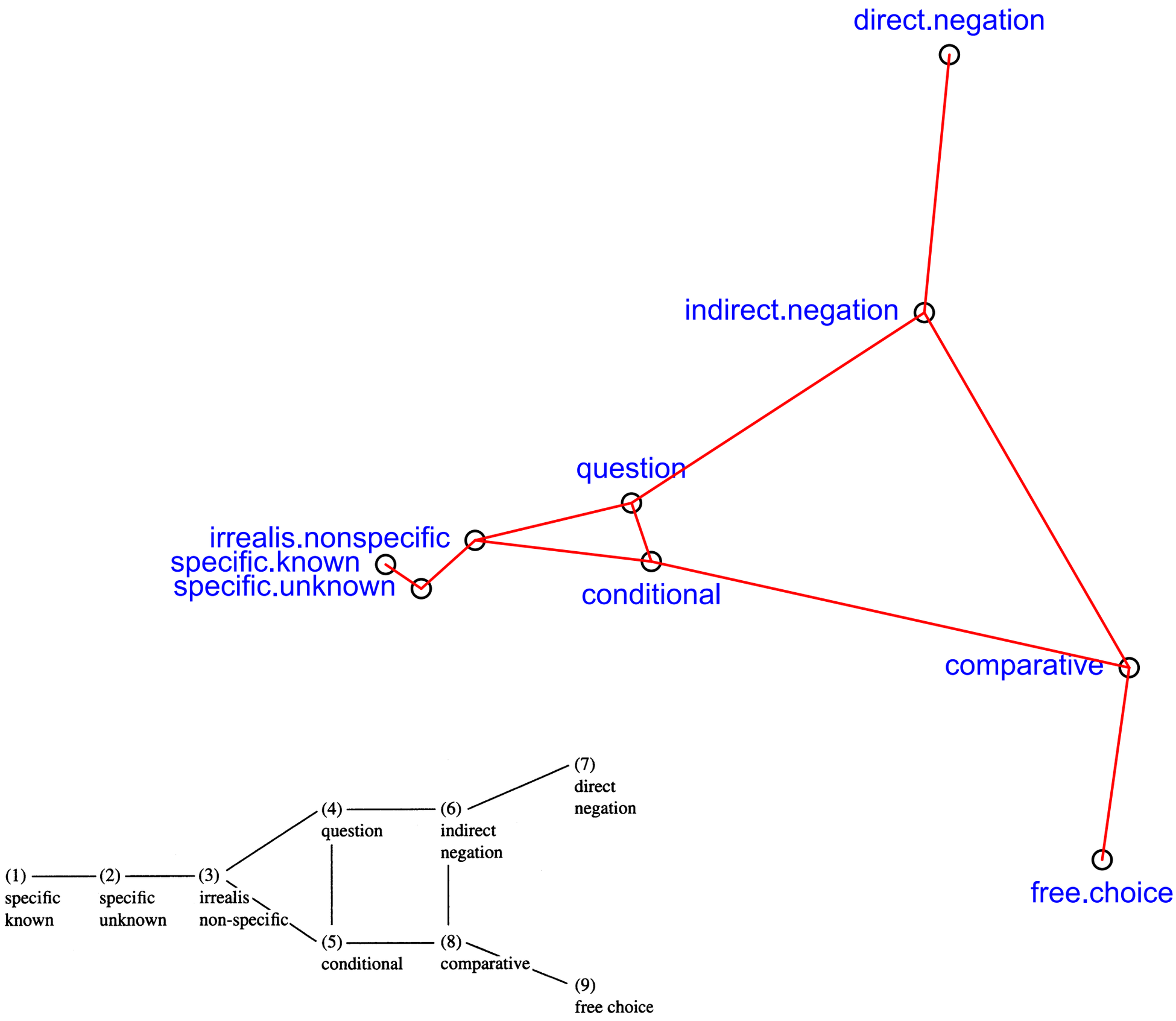
question

irrealis.nonspecific
specific.known
specific.unknown

conditional

comparative

free.choice


direct.negation
indirect.negation
question
irrealis.nonspecific
specific.known
specific.unknown
conditional
comparative
free.choice

(1)
specific
known
(2)
specific
unknown
(3)
irrealis
non-specific
(4)
question
(5)
conditional
(6)
indirect
negation
(7)
direct
negation
(8)
comparative
(9)
free choice

direct.negation

indirect.negation

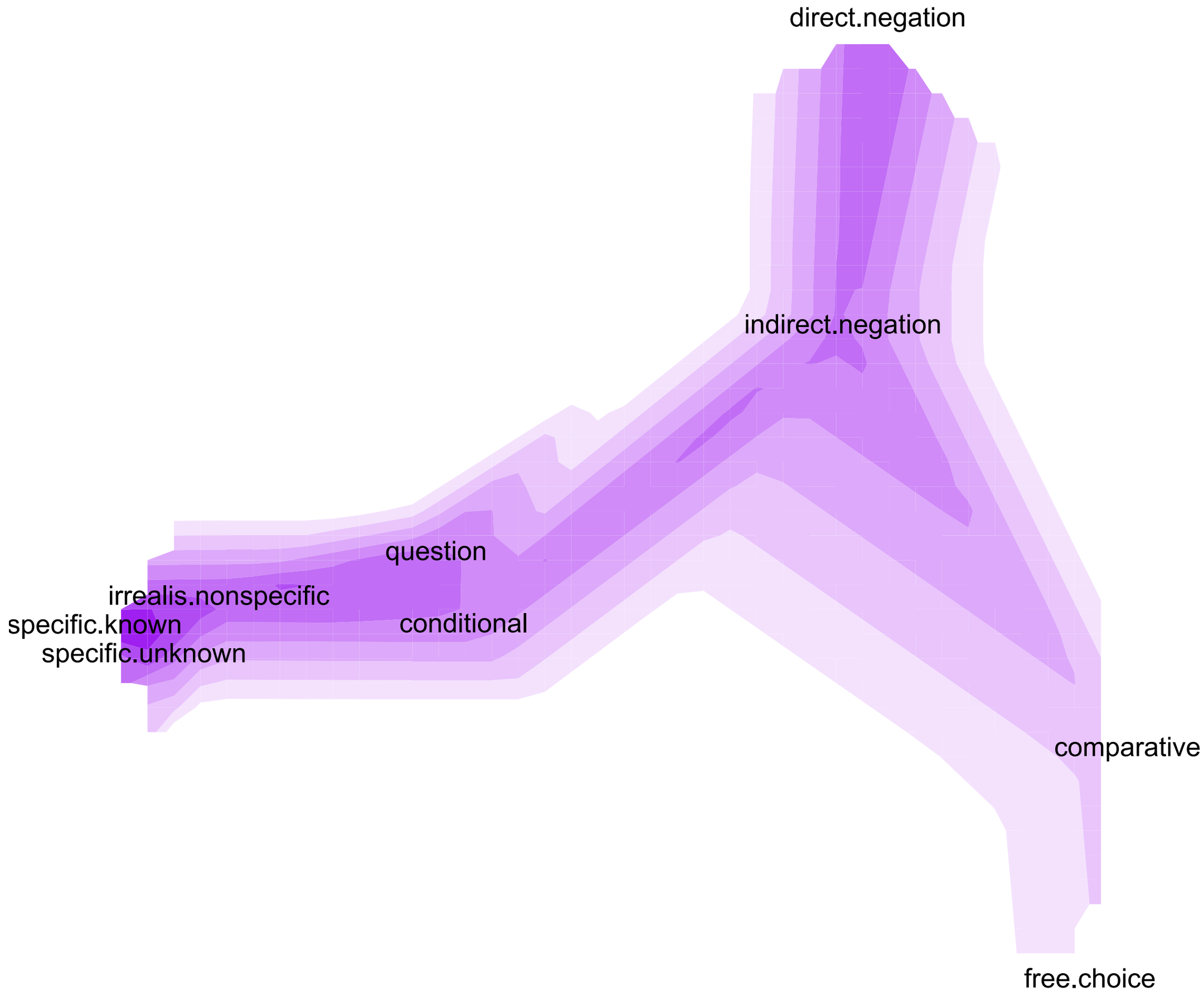
question

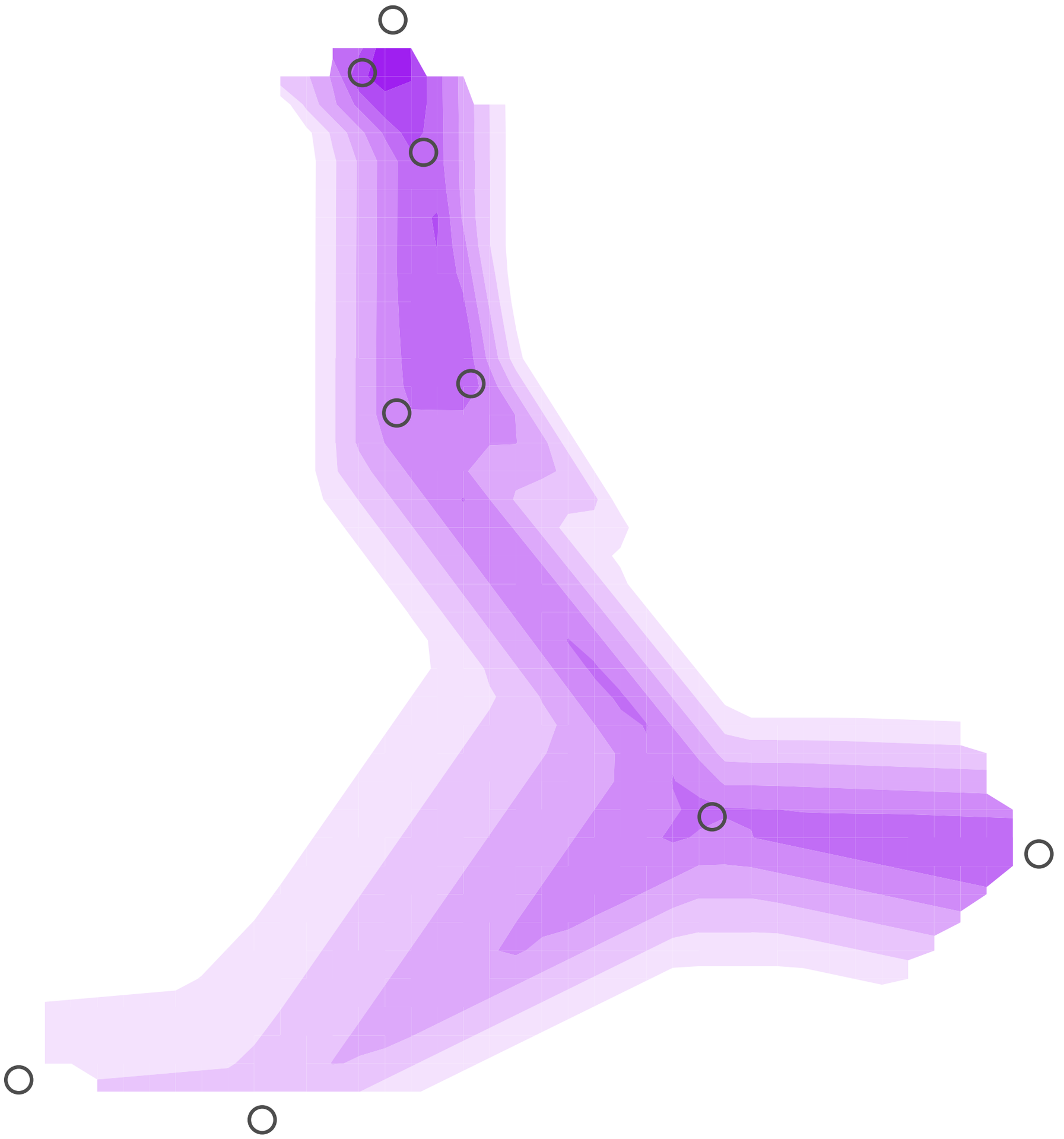
irrealis.nonspecific
specific.known
specific.unknown

conditional

comparative

free.choice


direct.negation
indirect.negation
question
irrealis.nonspecific
specific.known
specific.unknown
conditional
comparative
free.choice



Conclusions

- The world's linguistic diversity does not show a clear difference between attested and unattested structures
- Probabilistic semantic maps can deal much better with this than traditional graph-based maps
- Direction of change can be added as an asymmetric force-field in semantic space