Human-human, human-machine communication: on the HuComTech multimodal corpus

Laszlo Hunyadi, Tamás Váradi, István Szekrényes,
György Kovács, Hermina Kiss, Karolina Takács

CLARIN Annual Conference

Pisa, 9 October, 2018

Plan of the talk

On the multimodal HuComTech Corpus

The challenge of identifying multimodal patterns

From data to discovery: a cognitive model of behavioural pattern recognition

Challenges of analysis using *Theme*

Results: (hidden) patterns in selected pragmatic situations

On the multimodal HuComTech Corpus

Language: Hungarian

Kind: dialogues (formal, informal)

Subjects: 110 university students, single agent

Duration: ~ 50 hours

Words: ~ 500.000

Single annotations: ~ 1.5 million

On the multimodal HuComTech Corpus

Modalities: verbal, nonverbal

Verbal:
textual transcription
time aligned word segmentation (CLARIN resource: WebMouse)
syntax (shallow syntax, descriptive syntax for selected relations)
morphology
phonetics (phonetic transcription)
prosody (F0, intensity – both absolute values and stylization, voiced/unvoiced, simultaneous speaking)

Nonverbal:
formal properties vs. functional interpretation
gaze, face, hand, head, posture

On the multimodal HuComTech Corpus

Functional annotations:

pragmatic – multimodal, unimodal (dialogue management, intentions, agreement etc.)

turn management

communicative events

supporting acts

perceived emotions (both multimodal and unimodal)

The challenge of identifying multimodal patterns

within and between subject variability

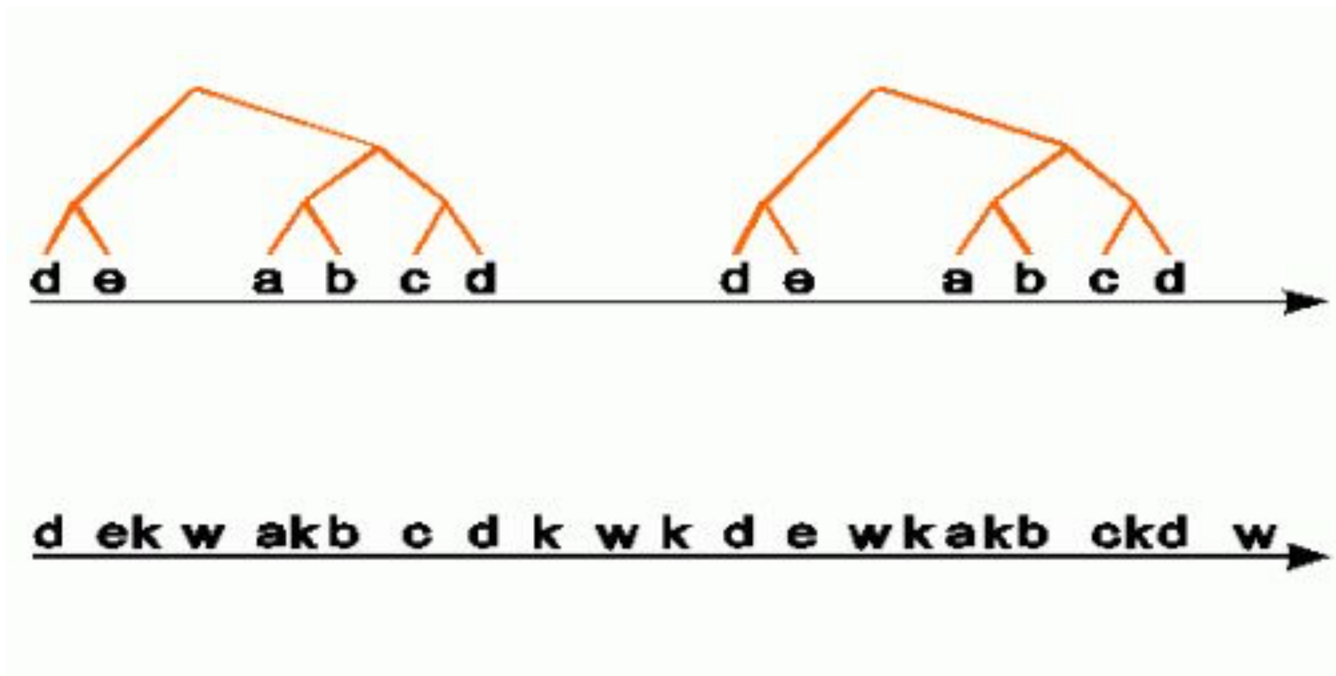
no strict adjacency requirement for events

optionality of events

temporal patterns determined by probability

patterns determined by either high or low frequency

A pattern can be hidden from the naked eye:



From data to discovery

Theme: a computer model of cognitive pattern recognition

Magnusson, M.S. (2000). Discovering hidden time patterns in behavior: T-patterns and their detection. Behavior Research Methods, Instruments and Computers 32(1): 93-110.

Magnusson, M.S., Burgoon, J., Casarrubea, M. (Eds.) (2016) Discovering Hidden Temporal Patterns in Behavior and Interaction: T-Pattern Detection and Analysis with THEME. Springer, 2016.

http://hbl.hi.is

Key concepts:

Event: single time value (no duration)

Critical interval

Detection: bottom-up, level-by-level,
competition and evolution, elimination of redundancies

Suitability of the HuComTech corpus for analysis using *Theme*

time aligned behavioural data

high degree of within subject and between subject variability

large number of subjects

Challenges of analysis presented by the HuComTech corpus

Size:

1.5 million individual annotations not counting word time alignment, morphology and syntax

Challenges of analysis presented by the HuComTech corpus

Complexity:

The subset of the corpus submitted to Theme consists of a maximum of 19 classes,112 items and 266 event types across 222 files with a maximum of 59134 data points within a maximum 1830355 sec. of observation time for a file.

Challenges for computation within *Theme*

Even with 64 Gb base memory and 20 dual-core processors it was practically beyond the system's capacity to simultaneously process all categories and all items even for three levels of search.

Reasonably reducing the number of classes and items was of course helpful, but by doing so we had to give up the chance to discover some of the “hidden hidden” patterns.

Another possibility for us was to shorten the files: we cut the files into shorter pieces to analyse, based on well defined turning points in the dialogues.

Challenges for computation within *Theme*

Another challenge was that within some of the patterns the standard deviation of corresponding critical intervals was expected to be rather small (for event types containing physical parameters), whereas for event types of interpretative data (emotions, pragmatics etc.) it was larger. Accordingly, in our search for these two kinds of patterns (or their combinations) we had to accommodate both of them in setting the maximum and the maximum critical interval.

Challenges for computation within *Theme*

Yet another challenge was that each subject's behavior was in a way unique, that was manifest in their difference in the amount of data points (average: 17695, stdev: 9082). Accordingly, whereas in order to reduce the amount of patterns found (and, accordingly, the load of processing) it was a good idea to limit the amount of patterns discovered at one level to be passed over to the next one (actually, our limit was a uniform 75% for all files), it could also cause a disproportionate reduction of possible patterns in files with fewer data points.

Challenges for computation within *Theme*

In any case, the maximum number of different patterns was 60037 (average: 4338, stdev: 6703) with respective pattern occurrences maximum 228194 (average: 16525, stdev: 25052). After better understanding the option “Adapt to data” we expect to solve this specific issue deriving from the above natural characteristic of individual behavior.

Challenges of *post-hoc* analysis

With such a large amount of different patterns in a file and a hierarchical complexity of patterns with sequentially identical or similar composition it was a difficult task to determine the set of characteristic patterns

both under within subjects and between subjects conditions

Challenges of *post-hoc* analysis

István Szekrényes built an application that was highly useful to meet these challenges.

In the following slides we'll demonstrate patterns identified in our subset of the corpus using István's work.

Results: (hidden) patterns in selected pragmatic situations

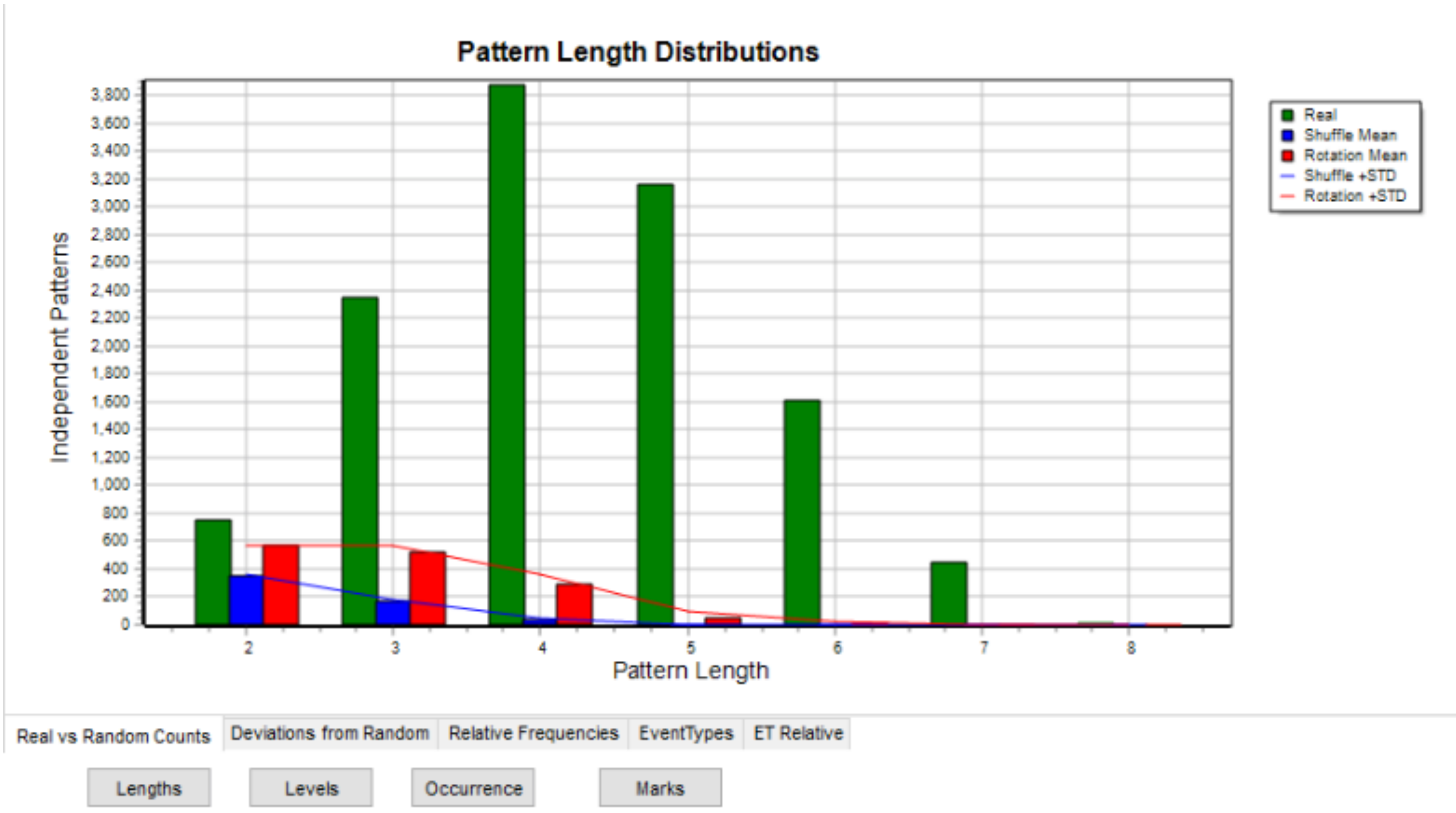
Results: (hidden) patterns in selected pragmatic situations

Event types in project

Results: (hidden) patterns in selected pragmatic situations

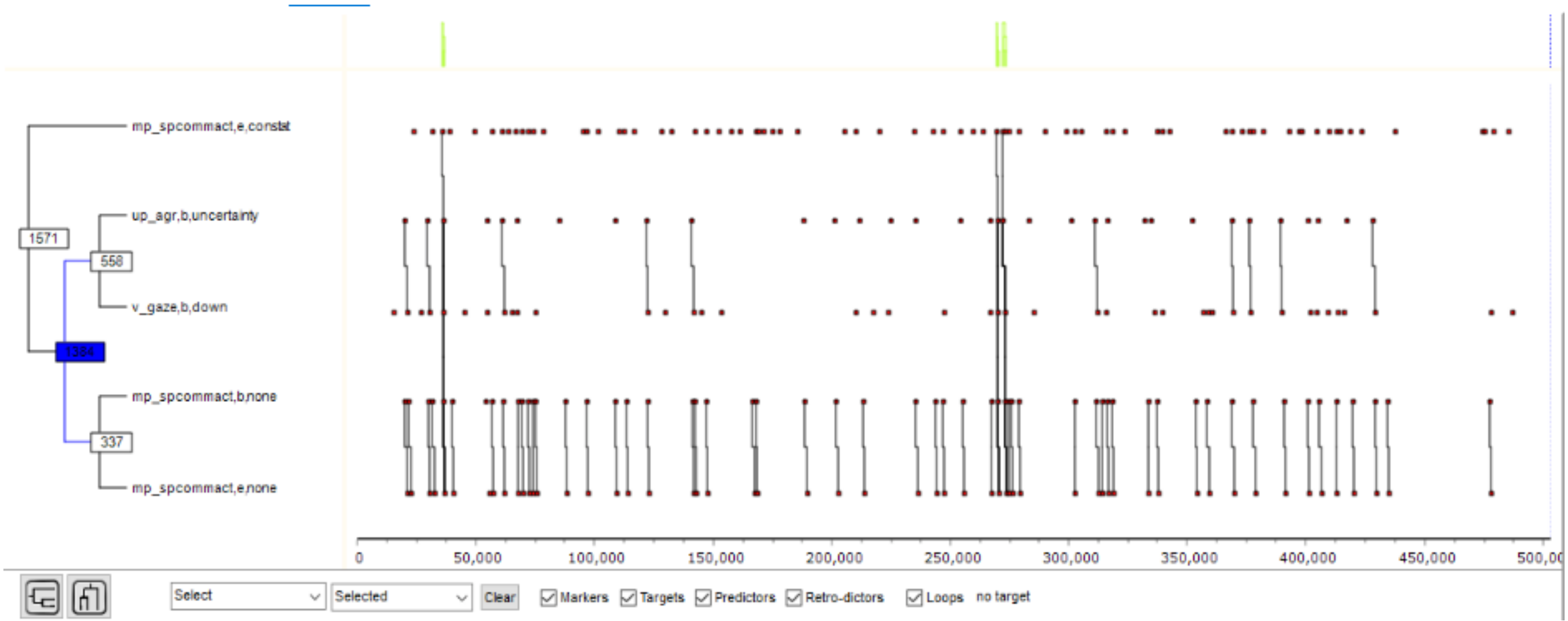
Event types in within and outside patterns

Results: (hidden) patterns in selected pragmatic situations



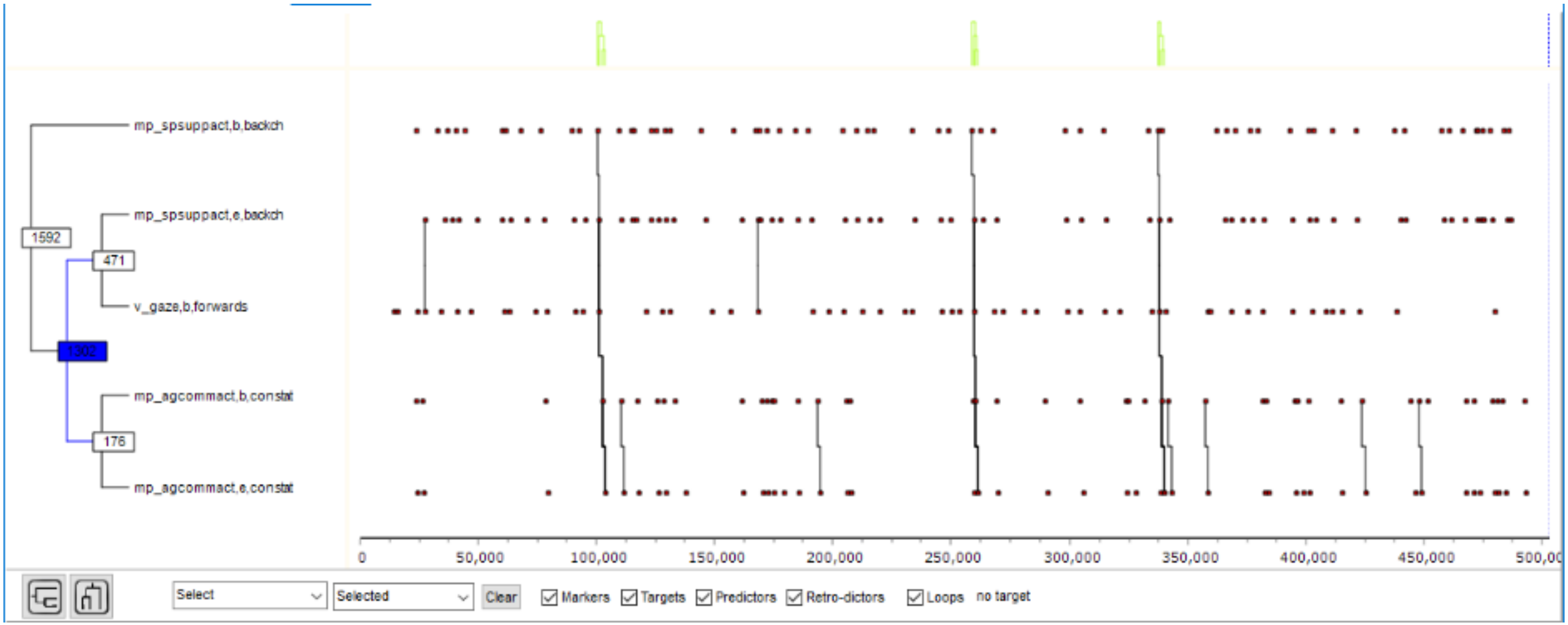
Reliability of pattern formation: real vs. random data

Results: (hidden) patterns in selected pragmatic situations



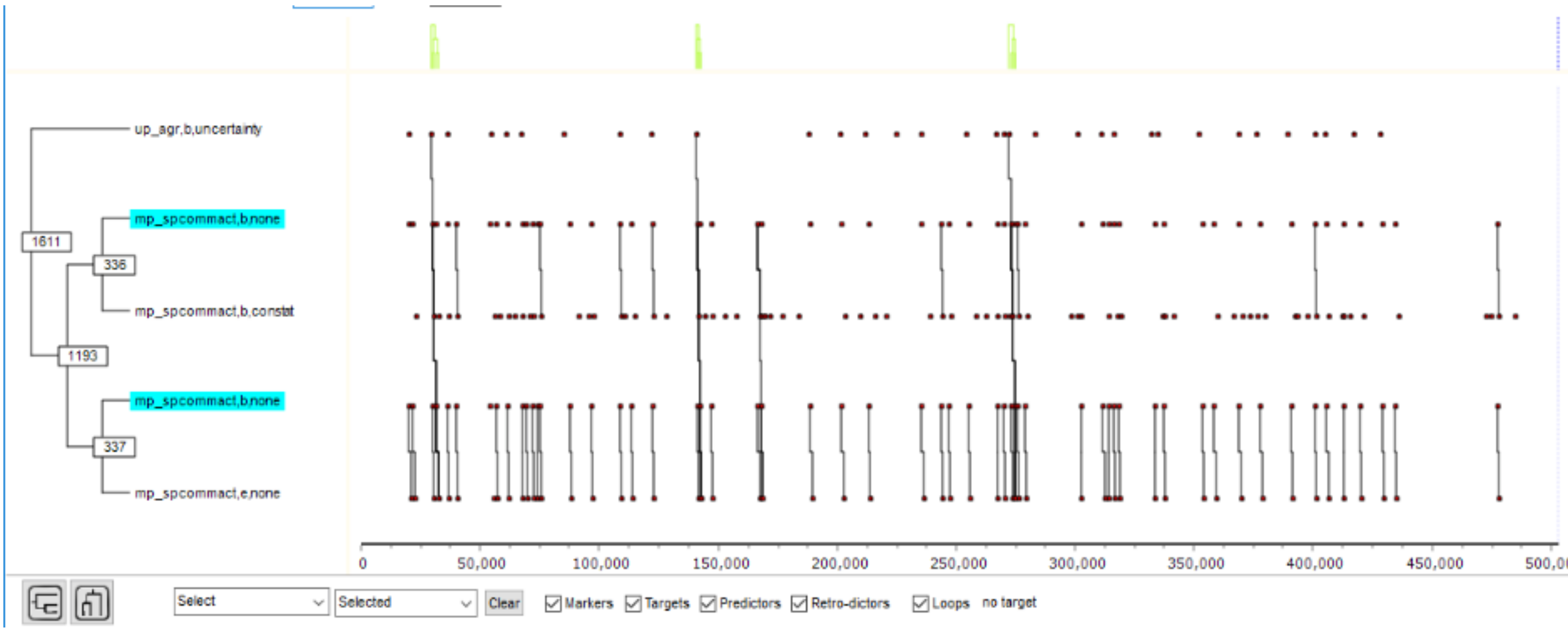
marker

Results: (hidden) patterns in selected pragmatic situations



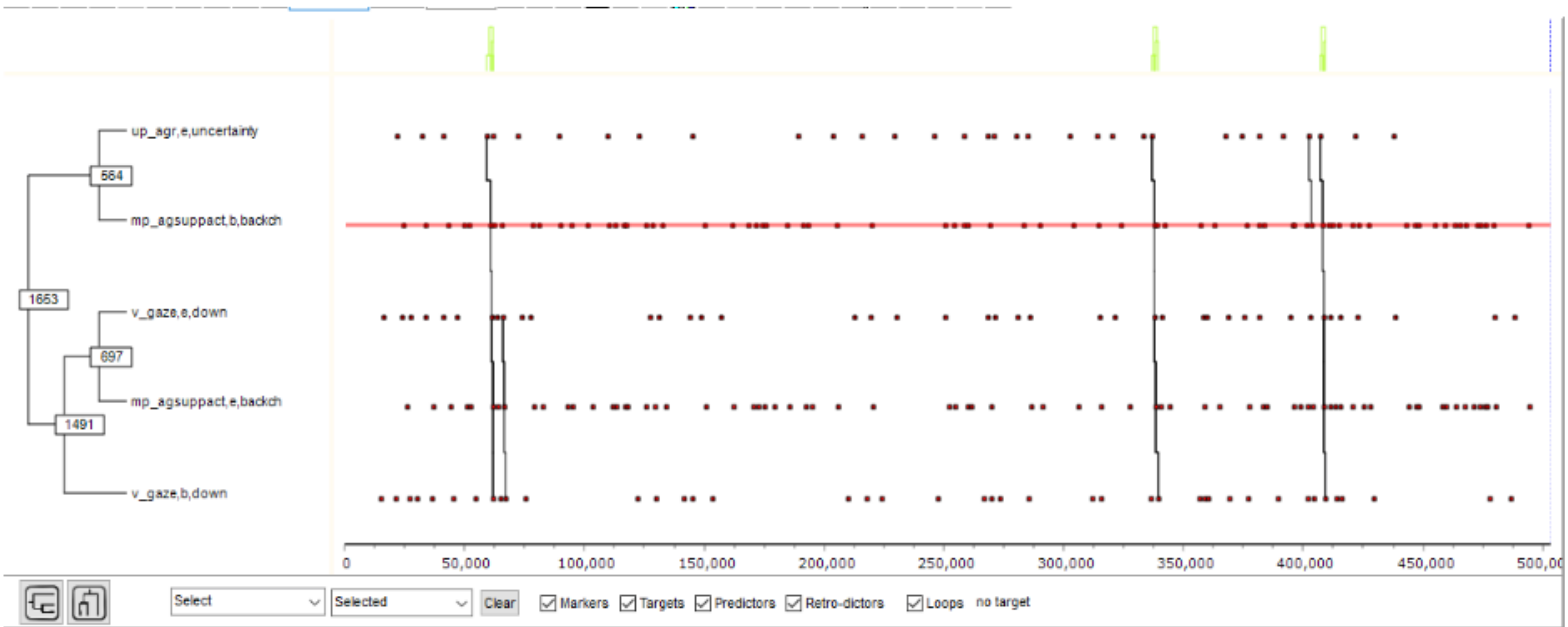
marker

Results: (hidden) patterns in selected pragmatic situations

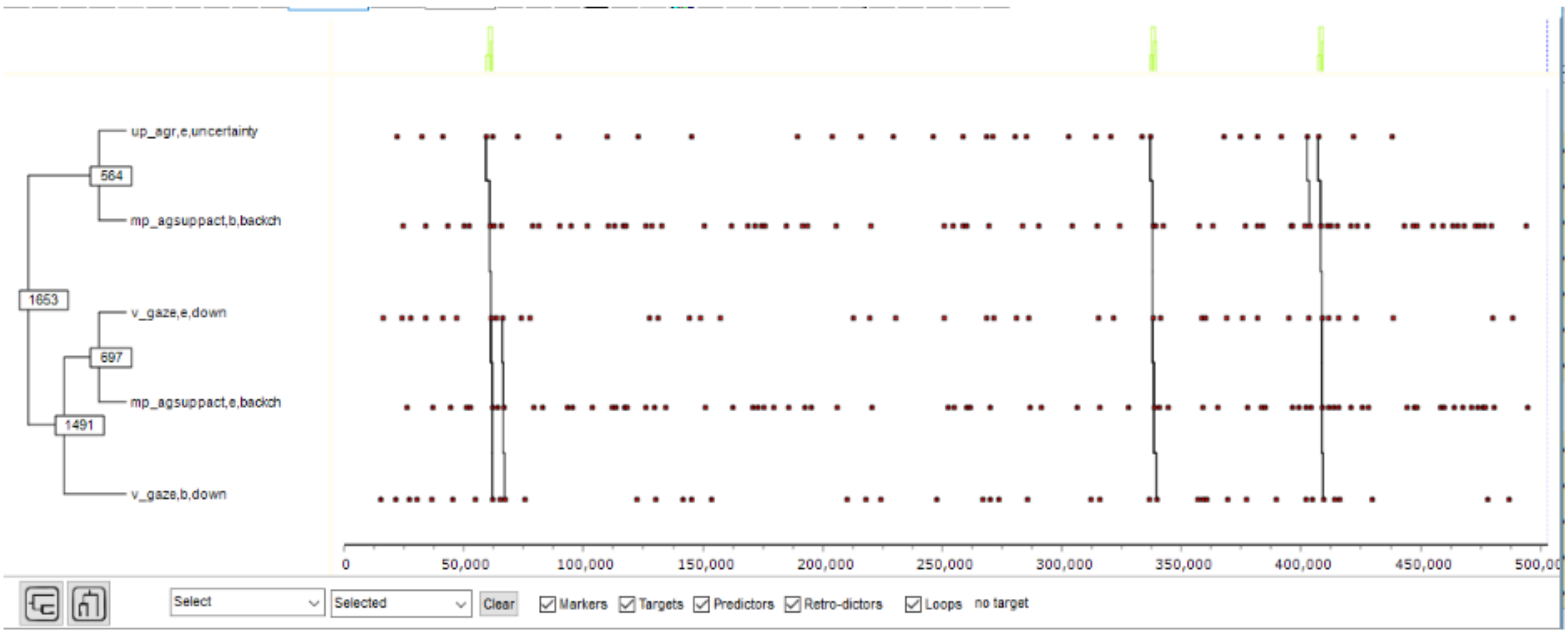


loop/recursion

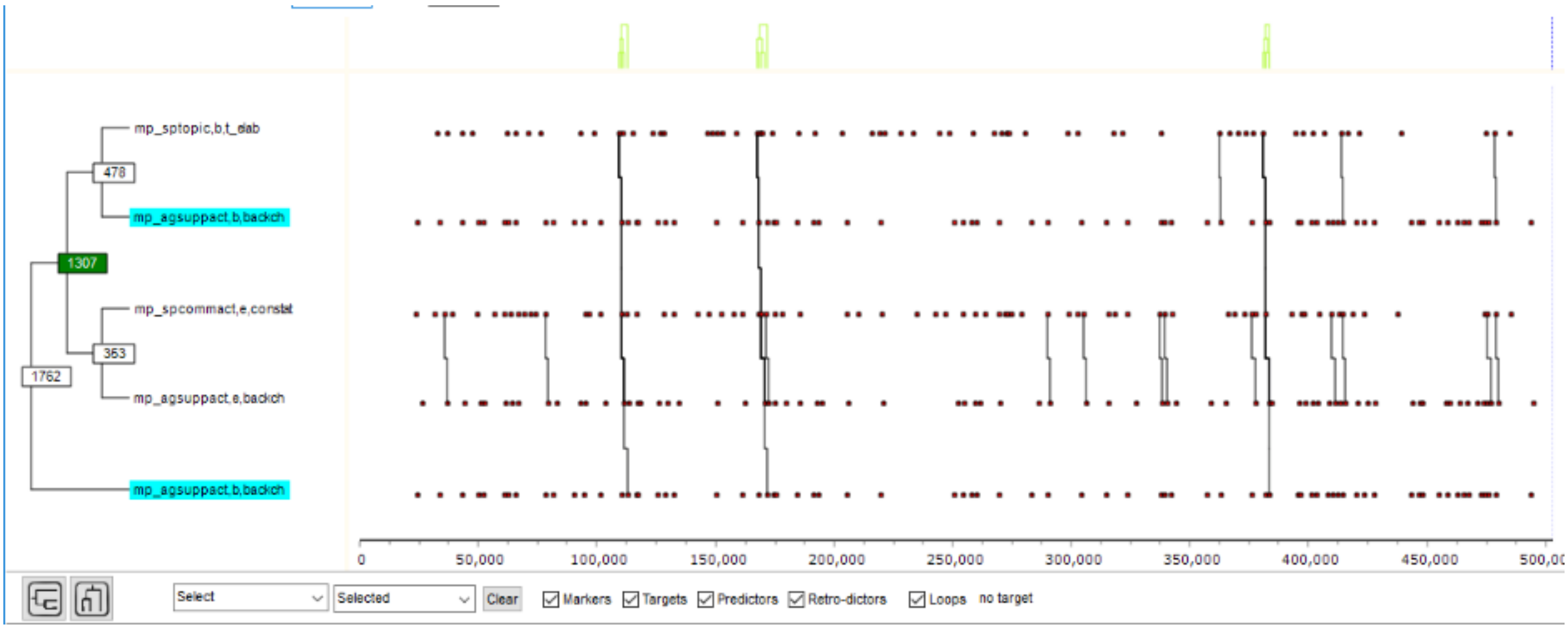
Results: (hidden) patterns in selected pragmatic situations



Results: (hidden) patterns in selected pragmatic situations

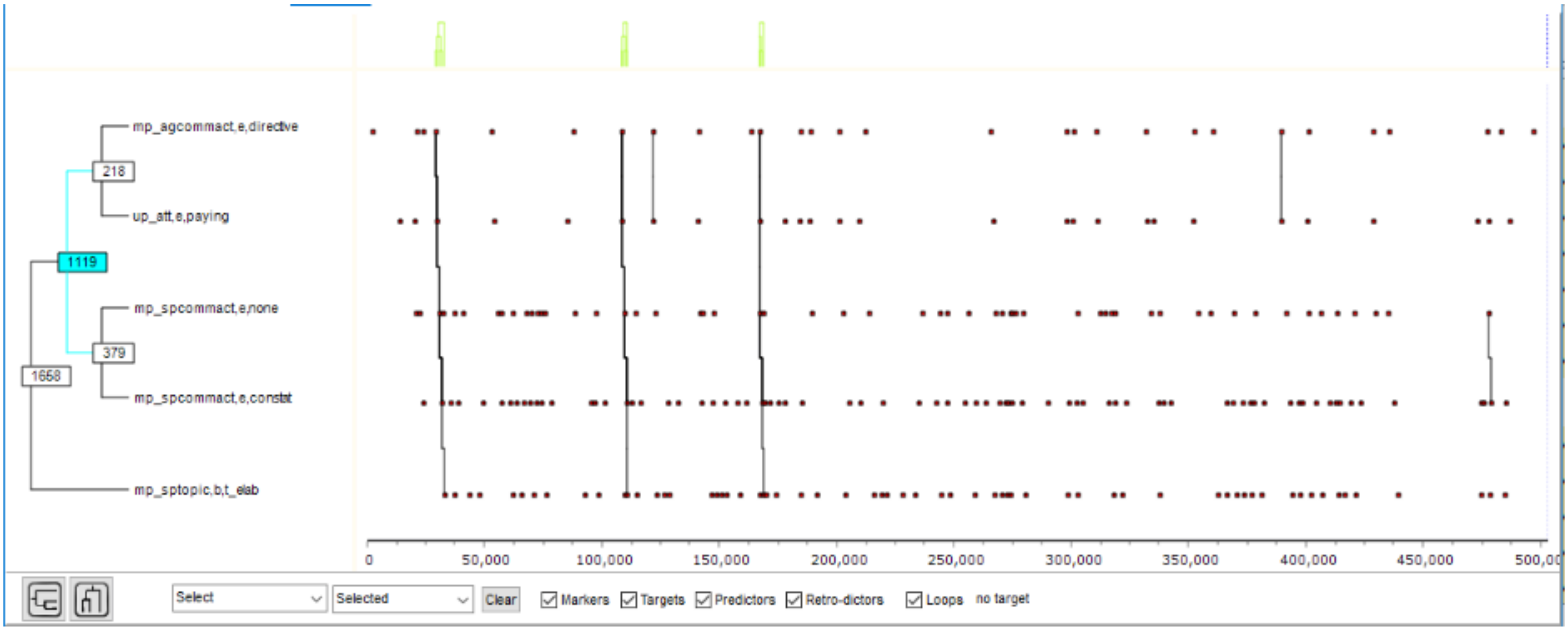


Results: (hidden) patterns in selected pragmatic situations



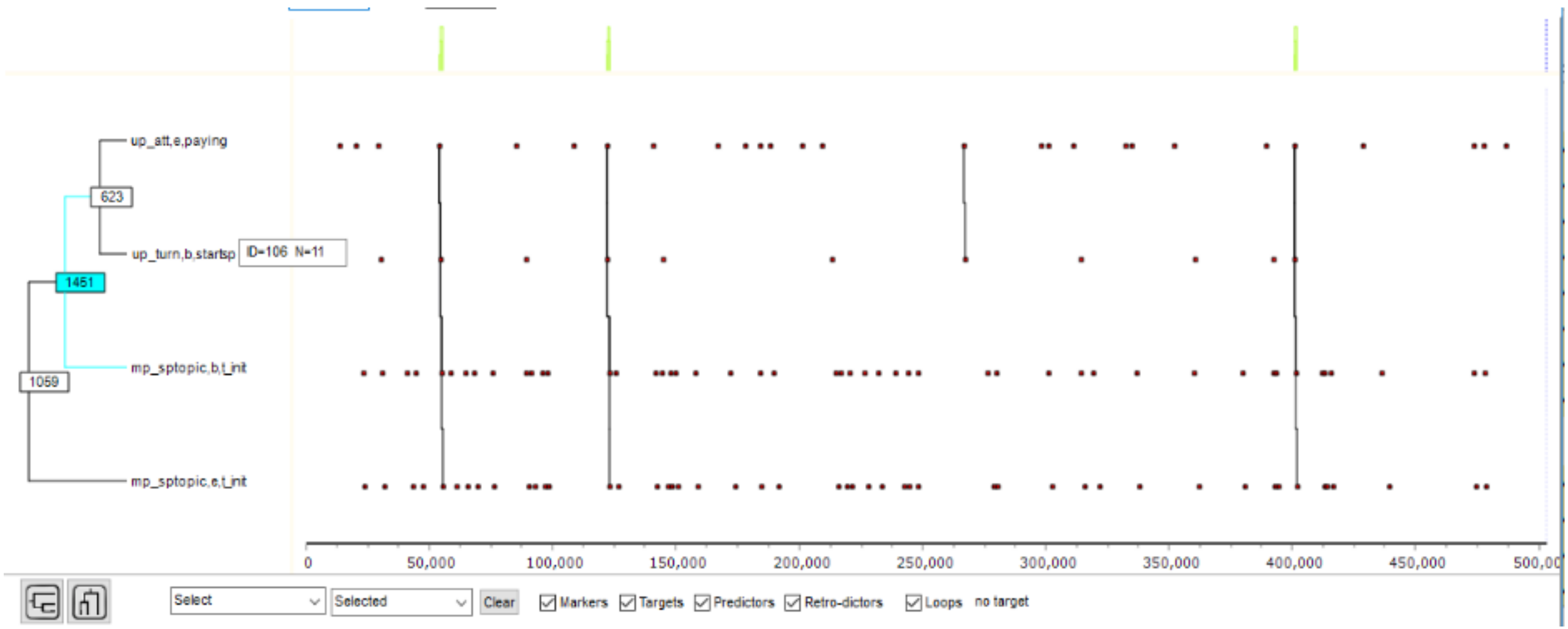
marker

Results: (hidden) patterns in selected pragmatic situations



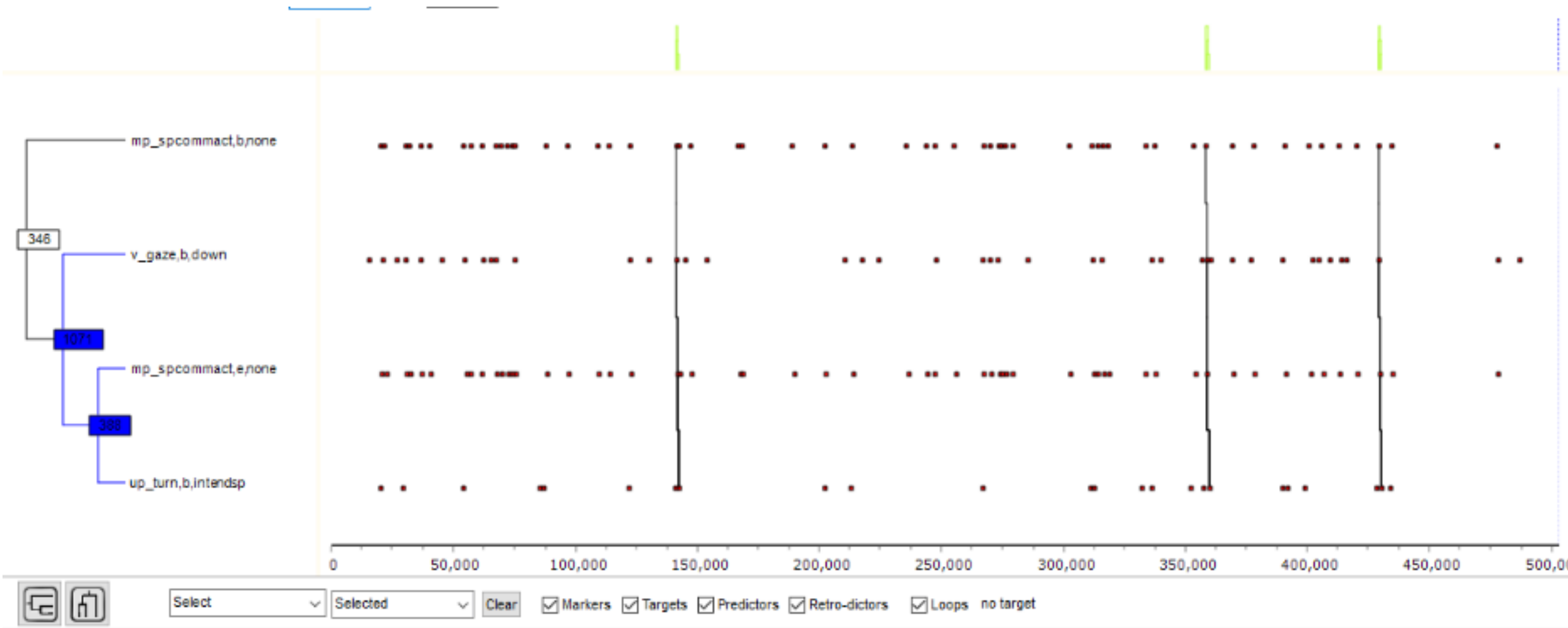
predictor

Results: (hidden) patterns in selected pragmatic situations



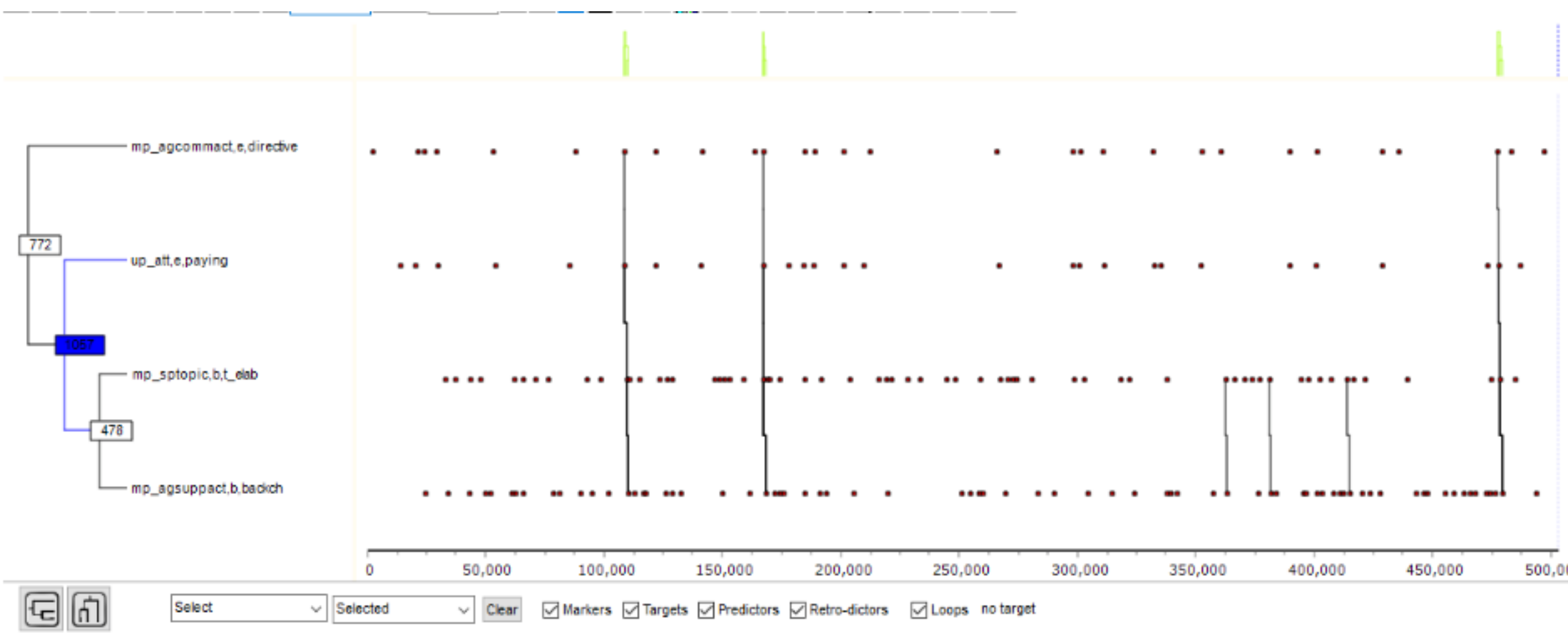
predictor

Results: (hidden) patterns in selected pragmatic situations



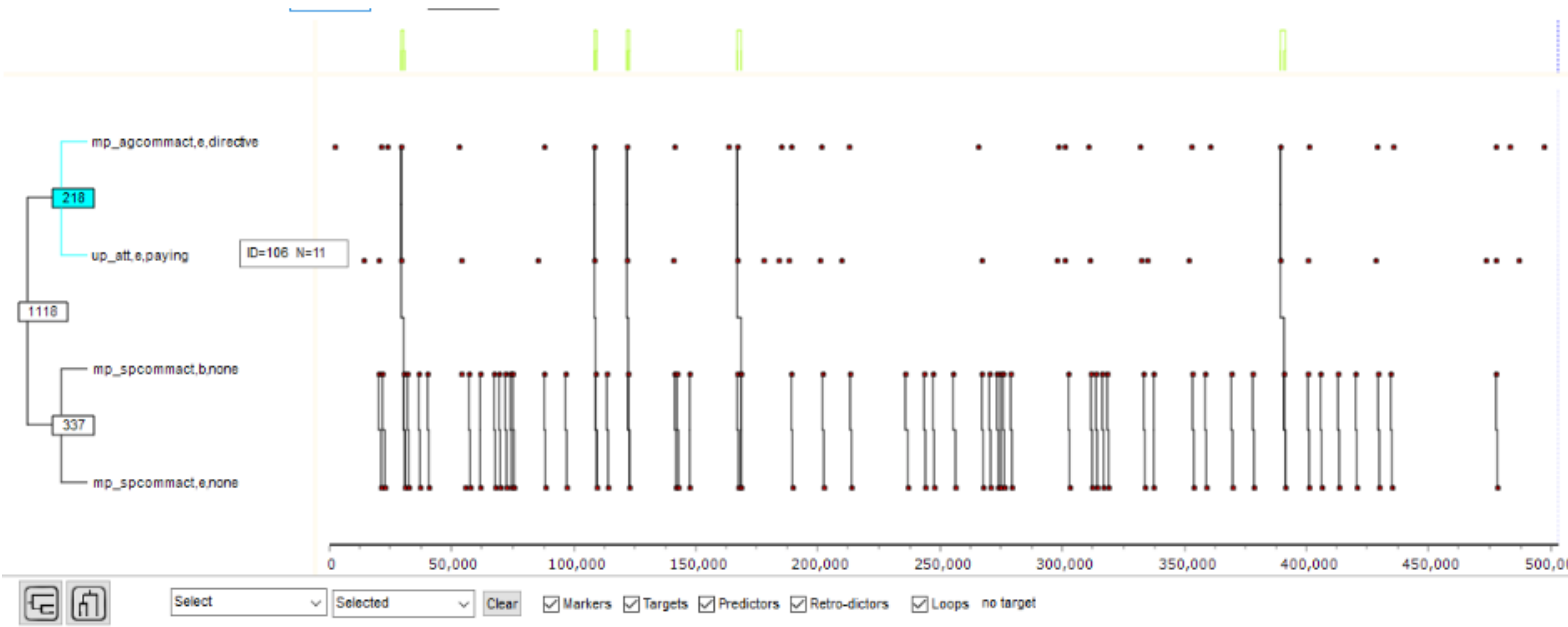
retrodictor

Results: (hidden) patterns in selected pragmatic situations

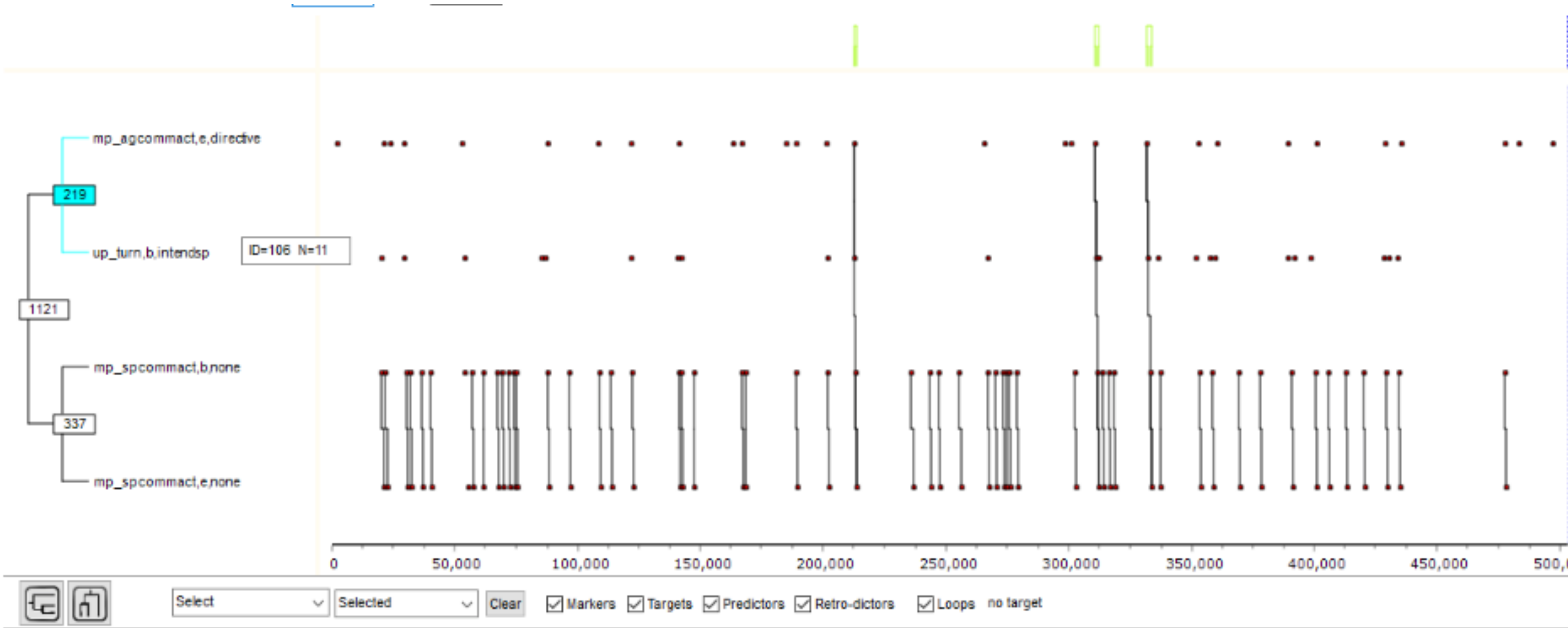


retrodictor

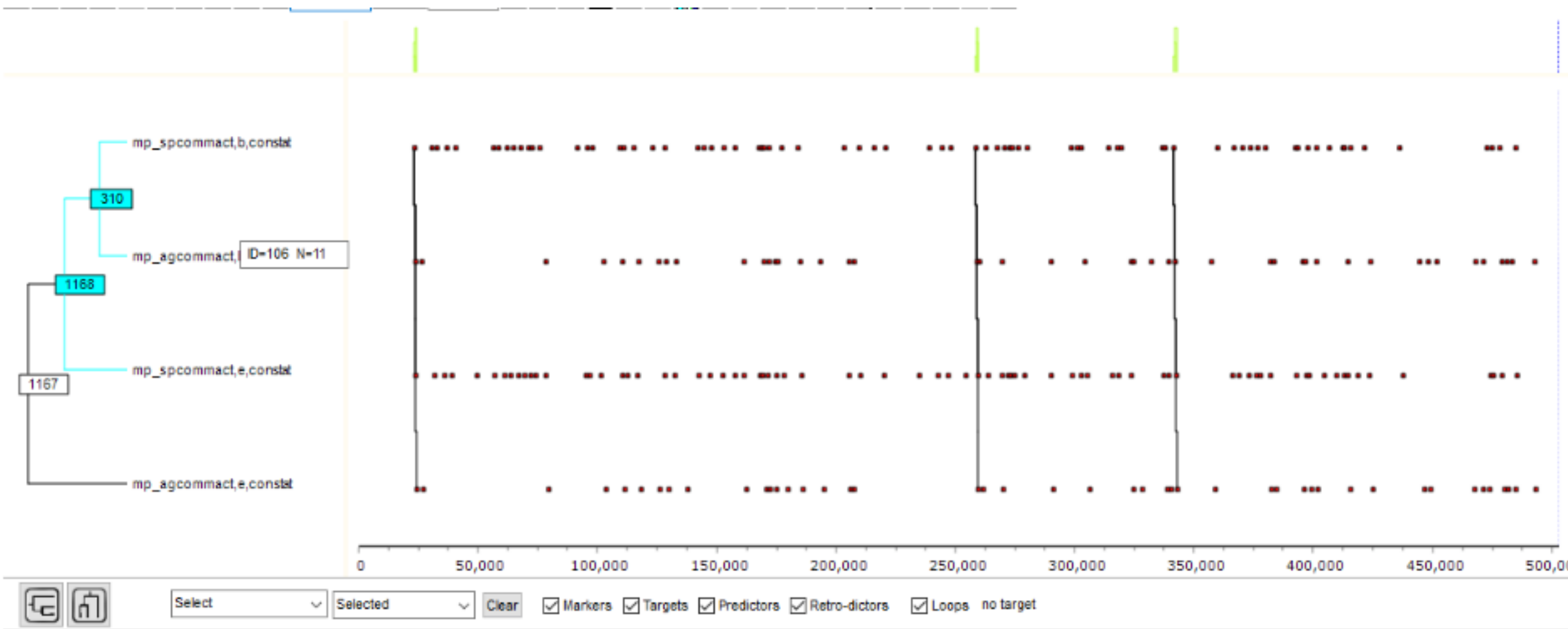
Results: (hidden) patterns in selected pragmatic situations



Results: (hidden) patterns in selected pragmatic situations



Results: (hidden) patterns in selected pragmatic situations



Results: (hidden) patterns in selected pragmatic situations

(((mp_spcommact,b,constat mp_agcommact,b,constat) mp_spcommact,e,constat) mp_agcommact,e,constat)

ID=1167   N= 3   Length= 4   Duration = 4222   %Duration= 1   Log10(P(1))= -6.63   Log10(P(N))= -323.31

Stats	StatI1	StatI2	StatI3
Min	553	87	233
Max	553	407	873
Mean	553.00	300.33	553.00
Median			
Mode			
StdDev	0.00	184.75	320.00
Max-Min+1	1	321	641

ID=106 N=11

Results: (hidden) patterns in selected pragmatic situations

(((mp_spcommact,b,constat mp_agcommact,b,constat) mp_spcommact,e,constat) mp_agcommact,e,constat

ID=1167 N= 3 Length= 4 Duration = 4222 %Duration= 1 Log10(P(1))= -6.63 Log10(P(N))= -323.31

Sample	Occur	I1	I2	I3
1	1	553	87	553
1	2	553	407	233
1	3	553	407	873

Results: (hidden) patterns in selected pragmatic situations

(((mp_spcommact,b,constat mp_agcommact,b,constat) mp_spcommact,e,constat) mp_agcommact,e,consta

ID=1167   N= 3   Length= 4   Duration = 4222   %Duration= 1   Log10(P(1))= -6.63   Log10(P(N))= -323.31

Sample	Occur	L1	L2	L3	L4
1	1	23040	23593	23680	24233
1	2	258560	259113	259520	259753
1	3	341440	341993	342400	343273

Results: (hidden) patterns in selected pragmatic situations

Frequent functional patterns with blink (selection)

id	length	Level	patstring	number_of_files	total_number
3426	2	1	(mp_spcommact,b,constat v_gaze,b,blink)	67	291
6751	2	1	(v_gaze,b,blink mp_spcommact,e,constat)	64	233
361	2	1	(v_gaze,b,blink mp_agsuppact,e,backch)	56	216
7811	2	1	(v_head,e,nod v_gaze,b,blink)	56	363
365	2	1	(v_gaze,b,blink mp_sptopic,b,t_elab)	54	200
63	2	1	(mp_agsuppact,e,backch v_gaze,b,blink)	53	235
7654	2	1	(mp_spcommact,e,constat v_gaze,b,blink)	52	219
10	2	1	(mp_agcommact,b,constat v_gaze,b,blink)	49	199
308	2	1	(up_att,e,paying v_gaze,b,blink)	49	194
360	2	1	(v_gaze,b,blink mp_agcommact,e,constat)	49	217
367	2	1	(v_gaze,b,blink mp_sptopic,e,t_elab)	49	189
6542	2	1	(mp_agsuppact,b,backch v_gaze,b,blink)	48	199
6752	2	1	(v_gaze,b,blink mp_spsuppact,b,backch)	48	186
7717	2	1	(v_gaze,b,blink mp_agsuppact,b,backch)	48	186
6542	2	1	(mp_agsuppact,b,backch v_gaze,b,blink)	48	199
11686	2	1	(mp_spinf,e,new v_gaze,b,blink)	48	205
6577	2	1	(mp_spsuppact,b,backch v_gaze,b,blink)	47	203
3361	2	1	(mp_agcommact,e,constat v_gaze,b,blink)	46	189
6753	2	1	(v_gaze,b,blink mp_spsuppact,e,backch)	46	184
369	2	1	(v_gaze,b,blink up_att,b,calling)	44	161
7721	2	1	(v_gaze,b,blink v_head,b,nod)	44	277

Results: (hidden) patterns in selected pragmatic situations

Frequent functional patterns with blink (selection)

id	length	Level	patstring	number_of_files	total_number
8627	2	1	(mp_sptopic,b,t_elab v_gaze,b,blink)	44	227
38362	2	1	(v_gaze,b,blink mp_spinf,b,new)	43	176
46575	2	1	(v_gaze,b,blink mp_spinf,e,new)	43	167
336	2	1	(up_turn,b,startsp v_gaze,b,blink)	42	172
6662	2	1	(up_att,b,paying v_gaze,b,blink)	41	177
43427	2	1	(mp_spsuppact,e,backch v_gaze,b,blink)	41	159
46561	2	1	(up_turn,b,endsp v_gaze,b,blink)	41	159
6557	2	1	(mp_agtopic,e,t_elab v_gaze,b,blink)	40	166
6748	2	1	(v_gaze,b,blink mp_agcommact,b,constat)	40	153
280	2	1	(up_att,b,calling v_gaze,b,blink)	39	139
7690	2	1	(up_att,e,calling v_gaze,b,blink)	39	148
8201	3	2	((v_head,b,nod v_gaze,b,blink) v_gaze,b,forwards)	37	174
209	2	1	(mp_sptopic,b,t_init v_gaze,b,blink)	37	152
3600	2	1	(v_gaze,b,blink up_att,b,paying)	36	147
11747	2	1	(mp_sptopic,e,t_elab v_gaze,b,blink)	35	146
362	2	1	(v_gaze,b,blink mp_agtopic,e,t_elab)	34	133
7639	2	1	(mp_agtopic,b,t_elab v_gaze,b,blink)	34	145
11763	2	1	(mp_sptopic,e,t_init v_gaze,b,blink)	33	138
8800	2	1	(v_gaze,b,blink mp_sptopic,b,t_init)	32	124
34931	2	1	(v_gaze,b,blink mp_agtopic,b,t_elab)	31	127
3351	2	1	(mp_agcommact,b,directive v_gaze,b,blink)	30	127
53843	2	1	(mp_aginf,b,new v_gaze,b,blink)	30	106

THANK YOU!