



What do Statistics Say about Discourse Production ? A Textometric study of Drafts

E. Née, E. MacMurray, S. Fleury, F.
Pugnière-Saavedra

SYLED

SYstèmes Linguistiques, Enonciation et Discours (EA2290)





2 work hypothesis

1. Observation of repetition in the writing process reveals both discursive routines which highly mark the reports as well as local changes that affect the writing process
2. These two phenomena display the strategies that characterize the processing of each child-report.



Two Points of Entry

- Corpus as whole unit
 - The frequency of the vocabulary
 - Structure of the different sections
- All-encompassing words that explain “what’s wrong”
 - Problems, difficulties, symptom, disorders



Overview

1. Textometric methods and their application to a corpus of drafts
2. Discursive routines, local changes and common strategies of the social writing/writers



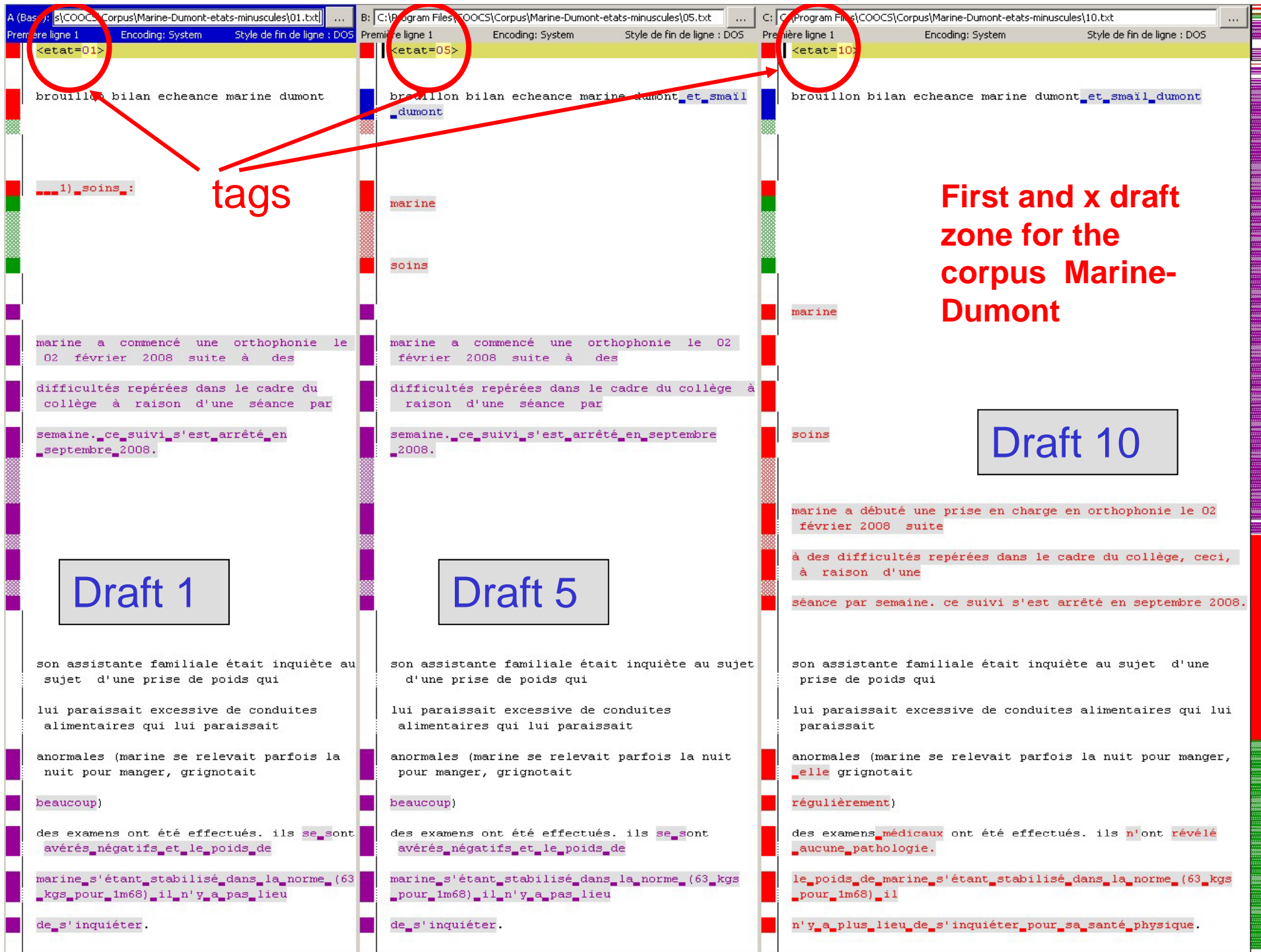
Textometry

Textometry has its foundations in the statistical study of linguistic data, originally used for lexical counts of words. In this approach, the text is considered as possessing its own internal structure, comparable to a prism of numbers and figures, producing information on the frequency counts of words which allows for the analysis of larger amounts of text than could be analyzed by manual means alone.

The text is separated into the different versions

	Draft 1	Draft 2	Draft 3	Draft 4	Draft 5	Draft 6	Draft 7	Draft 8	Draft 9	Draft 10
14	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine a des difficultés en espagnol, matière qu'elle vient de commencer cette année mais de façon générale elle se montre volontaire même si son assistante familiale doit parfois lui rappeler ses obligations scolaires.	Marine éprouve des difficultés en espagnol, matière qu'elle vient d'appréhender cette année.	Marine éprouve des difficultés en espagnol, matière qu'elle vient d'appréhender cette année.	Marine éprouve des difficultés en espagnol, matière qu'elle vient d'appréhender cette année.

Different stages of writings



tags

First and x draft zone for the corpus Marine-Dumont

Draft 1

Draft 5

Draft 10

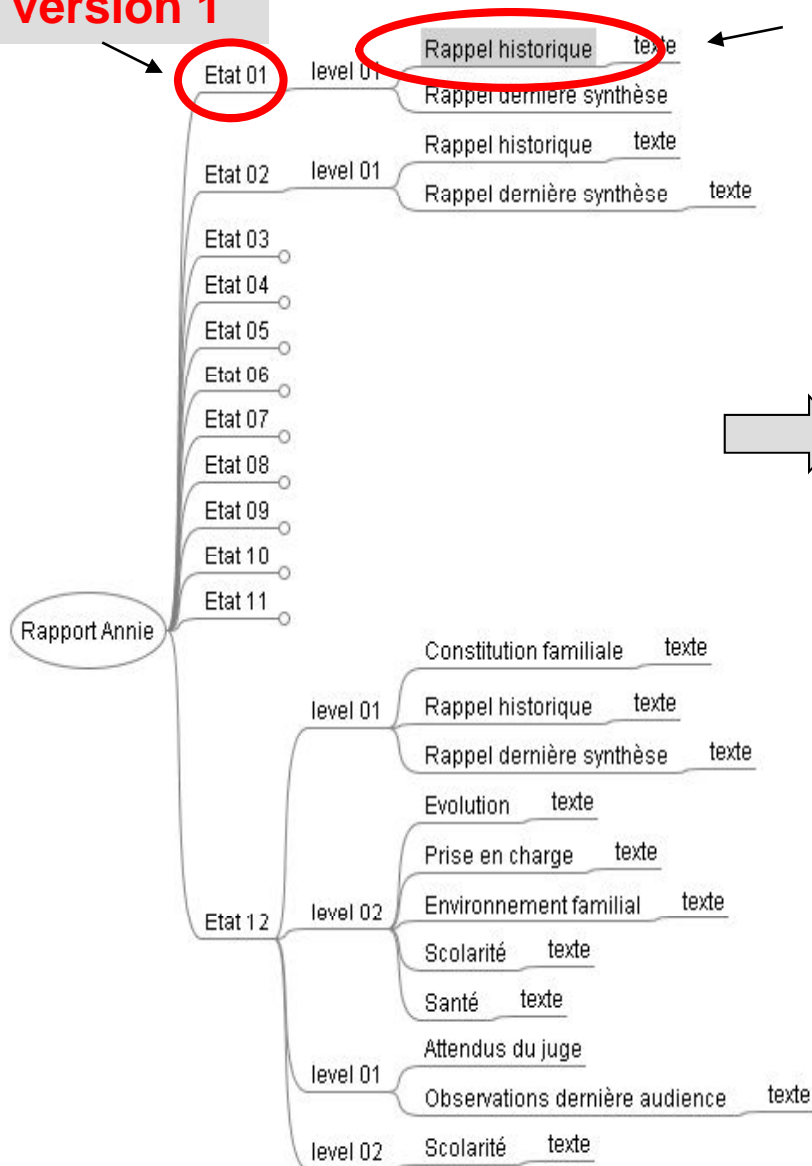
<etat=01>
brouillon bilan echeance marine dumont
---1)_soins_:
marine a commencé une orthophonie le 02 février 2008 suite à des difficultés repérées dans le cadre du collège à raison d'une séance par semaine. ce suivi s'est arrêté en septembre 2008.
son assistante familiale était inquiète au sujet d'une prise de poids qui lui paraissait excessive de conduites alimentaires qui lui paraissait anormales (marine se relevait parfois la nuit pour manger, grignotait beaucoup)
des examens ont été effectués. ils se sont avérés négatifs et le poids de marine s'étant stabilisé dans la norme (63 kgs pour 1m68) il n'y a pas lieu de s'inquiéter.

<etat=05>
brouillon bilan echeance marine dumont_et_small_dumont
marine
soins
marine a commencé une orthophonie le 02 février 2008 suite à des difficultés repérées dans le cadre du collège à raison d'une séance par semaine. ce suivi s'est arrêté en septembre 2008.
son assistante familiale était inquiète au sujet d'une prise de poids qui lui paraissait excessive de conduites alimentaires qui lui paraissait anormales (marine se relevait parfois la nuit pour manger, grignotait beaucoup)
des examens ont été effectués. ils se sont avérés négatifs et le poids de marine s'étant stabilisé dans la norme (63 kgs pour 1m68) il n'y a pas lieu de s'inquiéter.

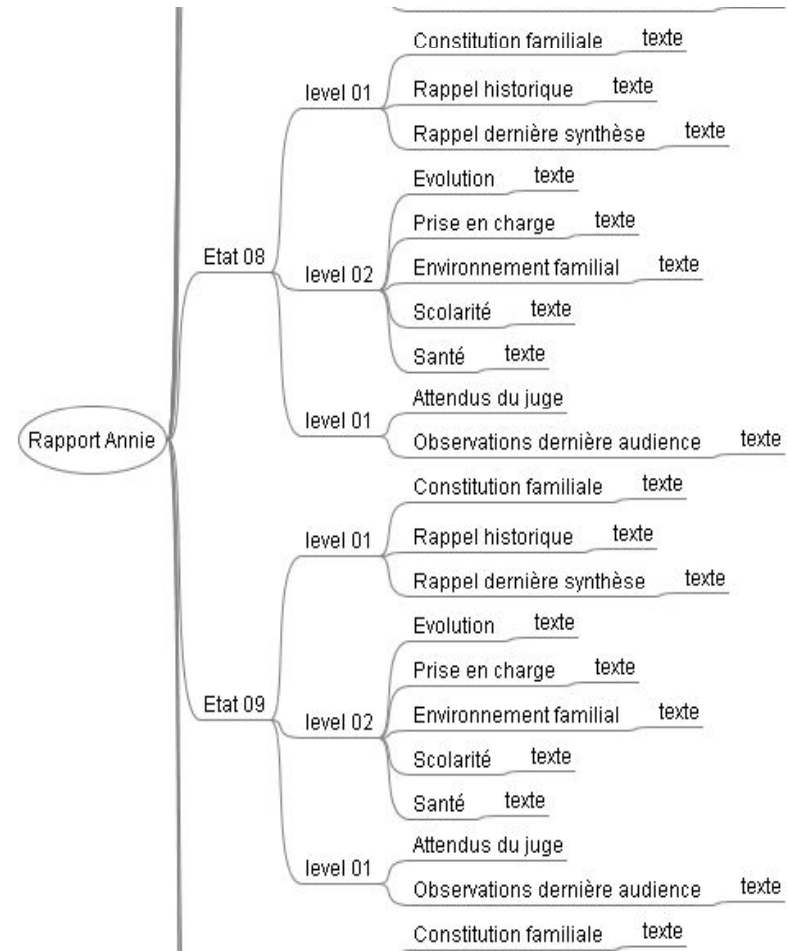
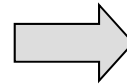
<etat=10>
brouillon bilan echeance marine dumont_et_small_dumont
marine
soins
marine a débuté une prise en charge en orthophonie le 02 février 2008 suite à des difficultés repérées dans le cadre du collège, ceci, à raison d'une séance par semaine. ce suivi s'est arrêté en septembre 2008.
son assistante familiale était inquiète au sujet d'une prise de poids qui lui paraissait excessive de conduites alimentaires qui lui paraissait anormales (marine se relevait parfois la nuit pour manger, elle grignotait régulièrement)
des examens médicaux ont été effectués. ils n'ont révélé aucune pathologie.
le poids de marine s'étant stabilisé dans la norme (63 kgs pour 1m68) il n'y a plus lieu de s'inquiéter pour sa santé physique.

Structure of sections in Marine-Dumont's report

Draft version 1

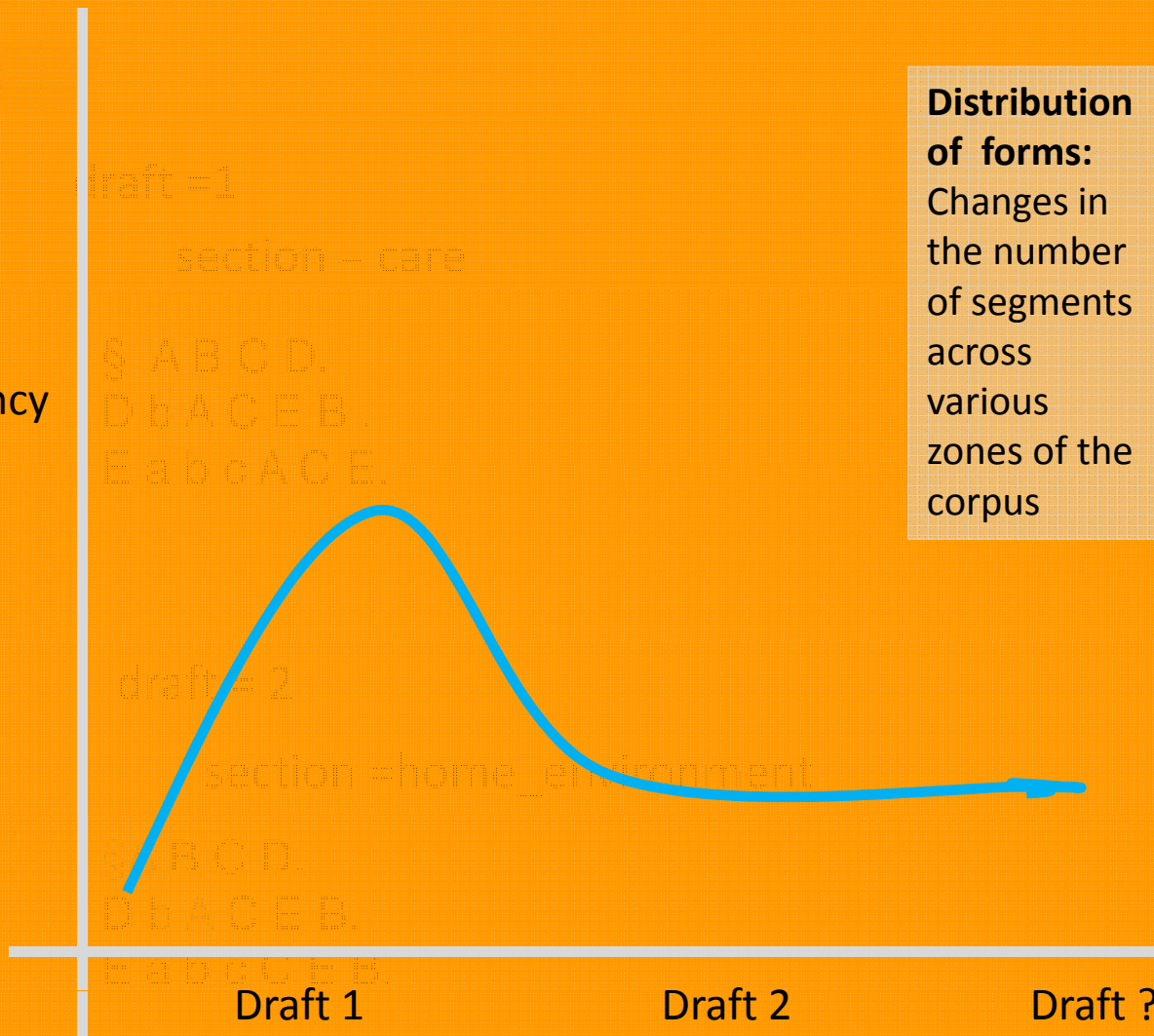


Section « summary of the earlier synthesis »



Textometric Method

Frequency
of A



**Distribution
of forms:**
Changes in
the number
of segments
across
various
zones of the
corpus

Frequency: 3
occurrences
draft 1

Frequency: 2
occurrences
draft 2

The frequency dictionary

The screenshot shows a software interface for a frequency dictionary. On the left, a table lists words and their frequencies. A bracket groups the first five words (de, et, a, sa, d) and is labeled "forms". Another bracket groups the last five words (pour, pas, au, madame, qu) and is labeled "frequencies". The software is identified as "Dictionnaire (Lexico 3 software)".

Formes (ordre lexicometrique)	Frequence
de	709
et	313
a	287
sa	271
d	237
à	225
une	225
marine	210
elle	200
en	199
son	198
le	189
des	179
il	173
est	168
mère	164
avec	159
l	157
la	154
dans	138
les	132
s	125
smail	123
dumont	115
pour	112
pas	110
au	108
madame	106
qu	103
se	102
un	99
été	98
qui	98
lui	94

697 formes

On the right, a text corpus is displayed. The word "<etat=01>" is highlighted in yellow. A speech bubble points to this word with the text "text of the corpus".

brouillon bilan echeance marine dumont

1) soins :

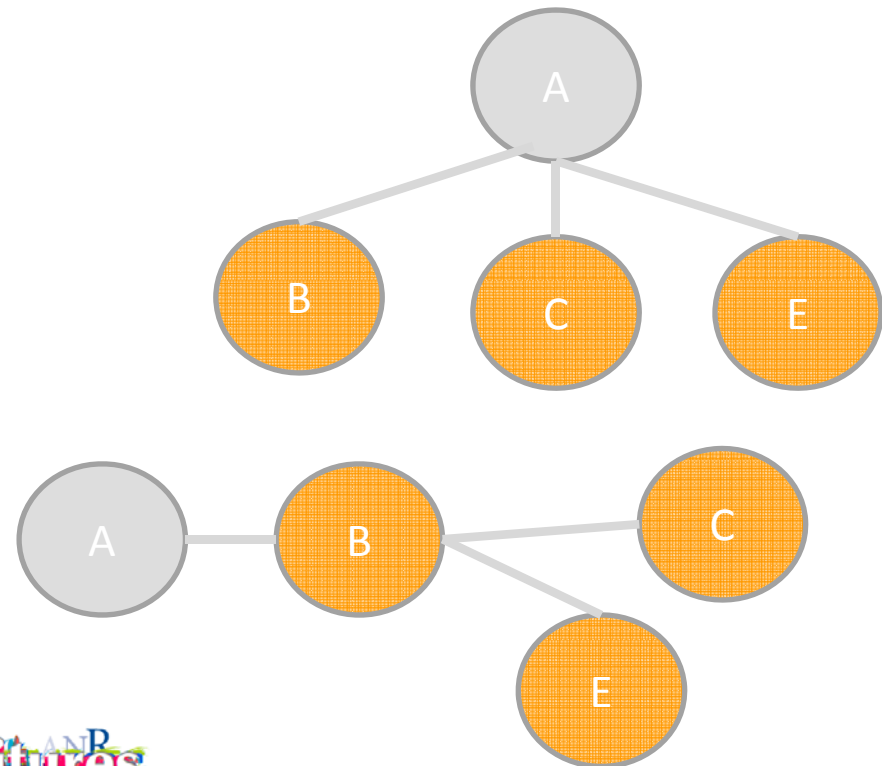
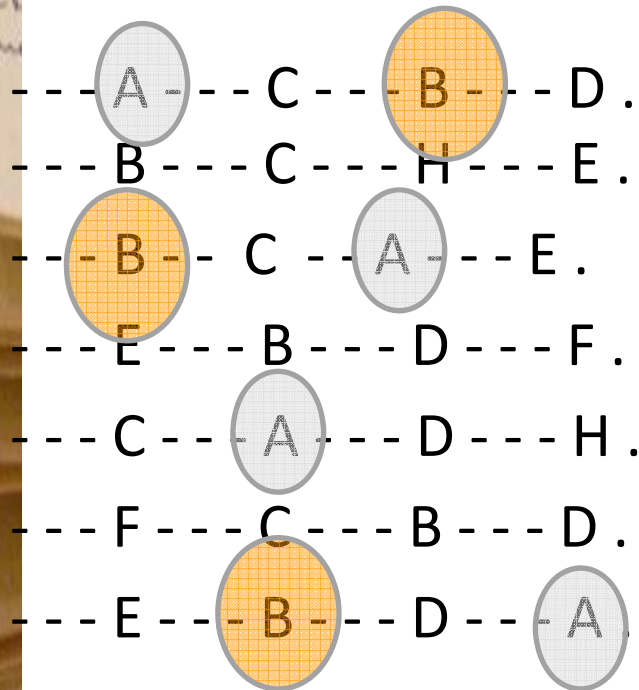
marine a commencé une orthophonie le 02 février 2008 suite à des difficultés repérées dans le cadre du collège à raison d'une séance par semaine . ce suivi s'est arrêté en septembre 2008 .

son assistante familiale était inquiète au sujet d'une prise de poids qui lui paraissait excessive de conduites alimentaires qui lui paraissait anormales (marine se relevait parfois la nuit pour manger , grignotait beaucoup)

des examens ont été effectués . ils se sont avérés négatifs et le poids de marine s'étant stabilisé dans la norme (63 kgs pour 1m68) il n'y a pas lieu

Co-occurrences / Poly-cooccurrences

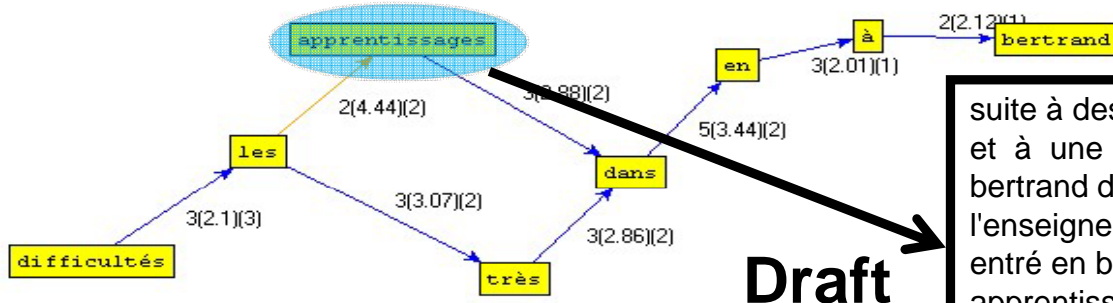
- Two words or more that appear at the same time in a predetermined span of text- lexical relationships around a pivot-form (William Martinez, 2003)
- Result: network of associative relationships...



Cooccurrences / Poly-cooccurrence

- **Frequency:** total frequency of the pivot-form
- **Co-Frequency:** the frequency with which the cooccurrence appears with the pivot-form in the defined context.
- ***Specificity:*** the probability that cooccurrence has for appearing in the defined context
- **Nombre de contextes:** the number of different context (paragraphs, sentences, etc.) that both the pivot-form and cooccurrence appear in.

ARC Couleur => bleu: (2<=Sp<=4) | vert: (4<Sp<=4) | orange: (4<Sp<=50) | rouge: (50<Sp)
 ARC Epaisseur => 1: (1<=NbContexte<=20) | 3: (20<NbContexte<=40) | 5: (40<NbContexte<=60) | 7: (60<NbContexte)
 ARC Label : nofreq(Sp) (nbContexte)

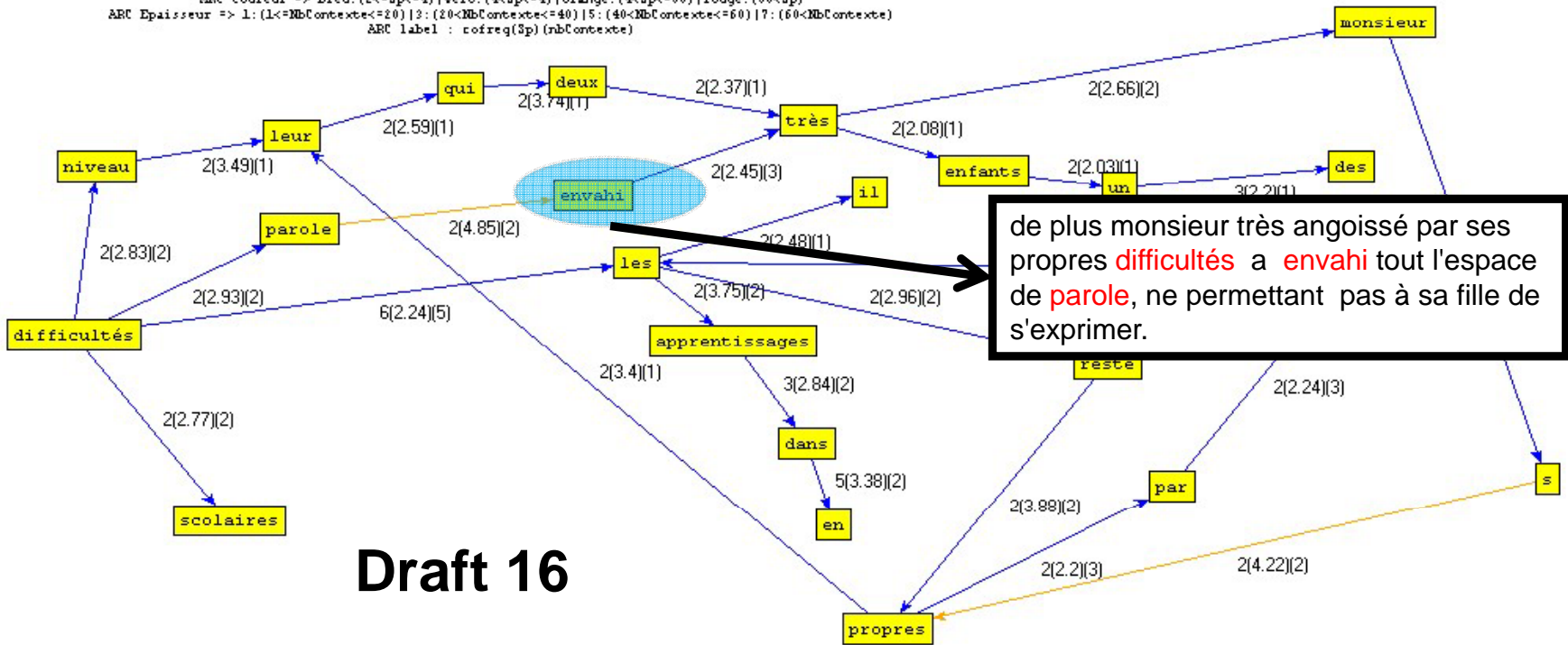


Draft 5

suite à des **difficultés** scolaires en 4ème et à une lassitude très importante de bertrand dans **les apprentissages** de l'enseignement classique, bertrand est entré en bep usinage dans le cadre d'un apprentissage à la rentrée scolaire 2008-2009.

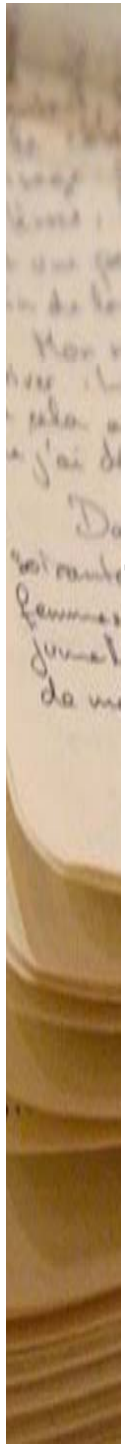
Thierry-Roux report

ARC Couleur => bleu: (2<=Sp<=4) | vert: (4<Sp<=4) | orange: (4<Sp<=50) | rouge: (50<Sp)
 ARC Epaisseur => 1: (1<=NbContexte<=20) | 3: (20<NbContexte<=40) | 5: (40<NbContexte<=60) | 7: (60<NbContexte)
 ARC Label : nofreq(Sp) (nbContexte)



Draft 16

de plus monsieur très angoissé par ses propres **difficultés** a **envahi** tout l'espace de **parole**, ne permettant pas à sa fille de s'exprimer.





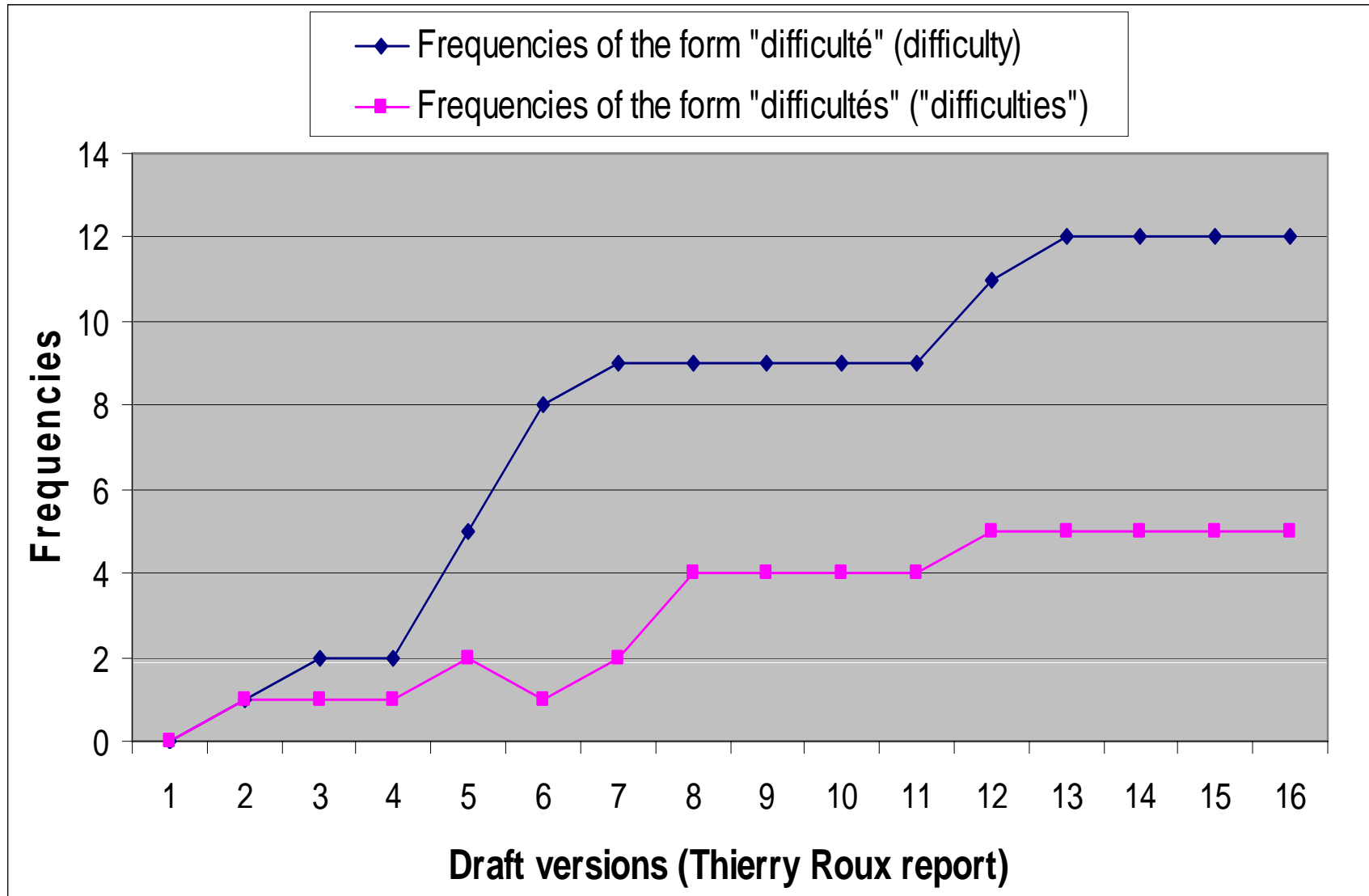
Discursive Routines, local changes and common strategies of the social writing/writers

The “**écrits au travail**” or “writings at work” (Boutet, Franckael, Delcambre 1992) involve some discursive routines. This property can be observed in the social reports through detection of typical usages and sequences. The genetic and textometric analysis of social reports, by comparing different draft versions for each report, reveals the routine aspect of this discourse from its writing process.

Draft reports: common sections

Synthesis report (documents de synthèse)	Final reports (rapports d'échéance)
<i>Evolution (evolution), prise en charge (taking charge of the children), environnement (environment), santé (health), scolarité (education).</i>	<i>Soins (care), scolarité (education), Enfant (NP) (child + name), évolution (evolution), conclusion (conclusion)</i>
Comments : All these sections account for half the sections that are used in Annie-Pauty's final synthesis, and for 5 on 6 sections in Anthony-Viti's final synthesis	Comments : All these sections account for one third (1/3) of the sections in Thierry-Roux's final report and half the sections in Marine-Dumont's final report

Densification of nouns indicating “what is wrong”: a “**phraseological effect**”





« [...] elle capte très **difficilement** l'attention de son père très envahit par ses propres **difficultés**. »

*Tr.: [...] she catch « with difficulties » her father's attention who is very overcome by his own **difficulties***

(Thierry Roux report, draft 5/16)

« Thierry a beaucoup grandi durant cette année, il n'est pas toujours très à l'aise avec sa taille et se tient **difficilement** droit. Son allure générale laissait penser à des **difficultés** au niveau du dos ».

Tr.: Thierry has grown a lot this year, he does not feel very comfortable with his height and stand up straight « with difficulty ».

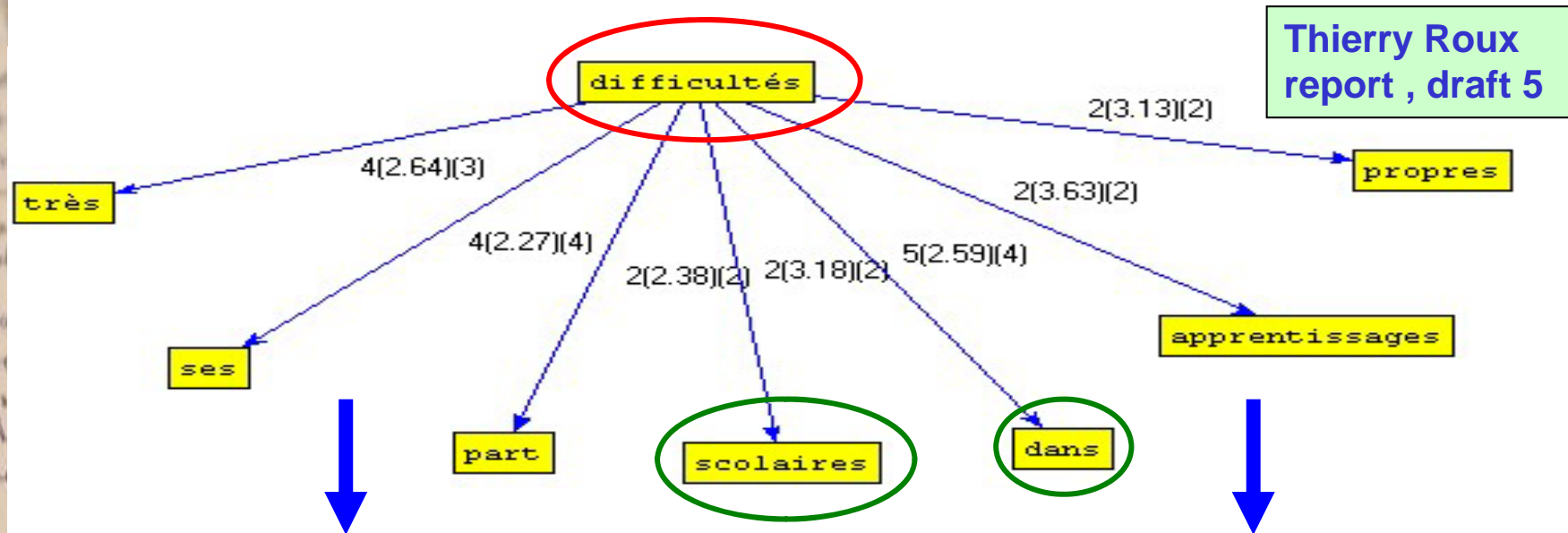
(Thierry Roux report, draft 5/16)

« La scolarité est toujours aussi **difficile**. [...] la rencontre avec le directeur et la surveillante principale a montré que les enseignants prenaient actes des **difficultés** d'Annie [...]

*Tr.: She still has **many difficulties** at school. The meeting with the headmaster and the chief supervisor showed that the teachers noted Annie's **difficulties** [...]*

(Rapport Annie Pauty, draft 9/12)

The pivot-form *difficultés* (difficulties) and prepositions : co-occurrences



En troisième , Evelyne a rencontré des **difficultés** croissantes, **dans les apprentissages** [...]

Tr. *In 9th grade, Evelyne has met increasing **difficulties**, **with learning** [...]*

Suite à des **difficultés scolaires** en 4ème et à une lassitude très importante de Bertrand **dans les apprentissages** de l'enseignement classique [...]

Tr. *Following **academic difficulties** when he was in the 8th grade and to a very important weariness **of learning** standard subjects [math, geography,...] [...]*

Constructions and lexical uses more or less fixed

X_{l'enfant} a des **difficultés** dans + SN

Tr.: X_{the child} has **difficulties** in or at + N

difficultés scolaires, maternelles

Tr.: *difficulties at school, maternal difficulties (=difficulties of the mother)*

l'enfant (n)est (pas)en (grande) **difficulté** pour FAIRE qqch

Tr.: X_{the child} is (not) in difficulties to DO sth

X_{l'enfant/les parents} a/ont des difficultés/difficultés de X_{l'enfant/les parents}
à/pour + FAIRE qqch

Tr.: X_{the child/parents} has/have difficulties to do/not to do sth

*Language dictionaries: avoir une/quelque **difficulté** à/pour faire qqch*

Tr.: to have (a) **difficulty** to do sth

1. « elle (Evelyne) capte très difficilement l'attention de **son père très envahi par ses propres difficultés** »

(Draft 7, description of relationships between Evelyne and her father)

2. « De plus **Mr très** angoissé **par ses propres difficultés** a **envahi** tout l'espace de parole, ne permettant pas à sa fille de s'exprimer. »

(Draft 7, description of relationships between Odile and her father)

3. « En effet, **Monsieur** est **très envahi par ses difficultés propres** »

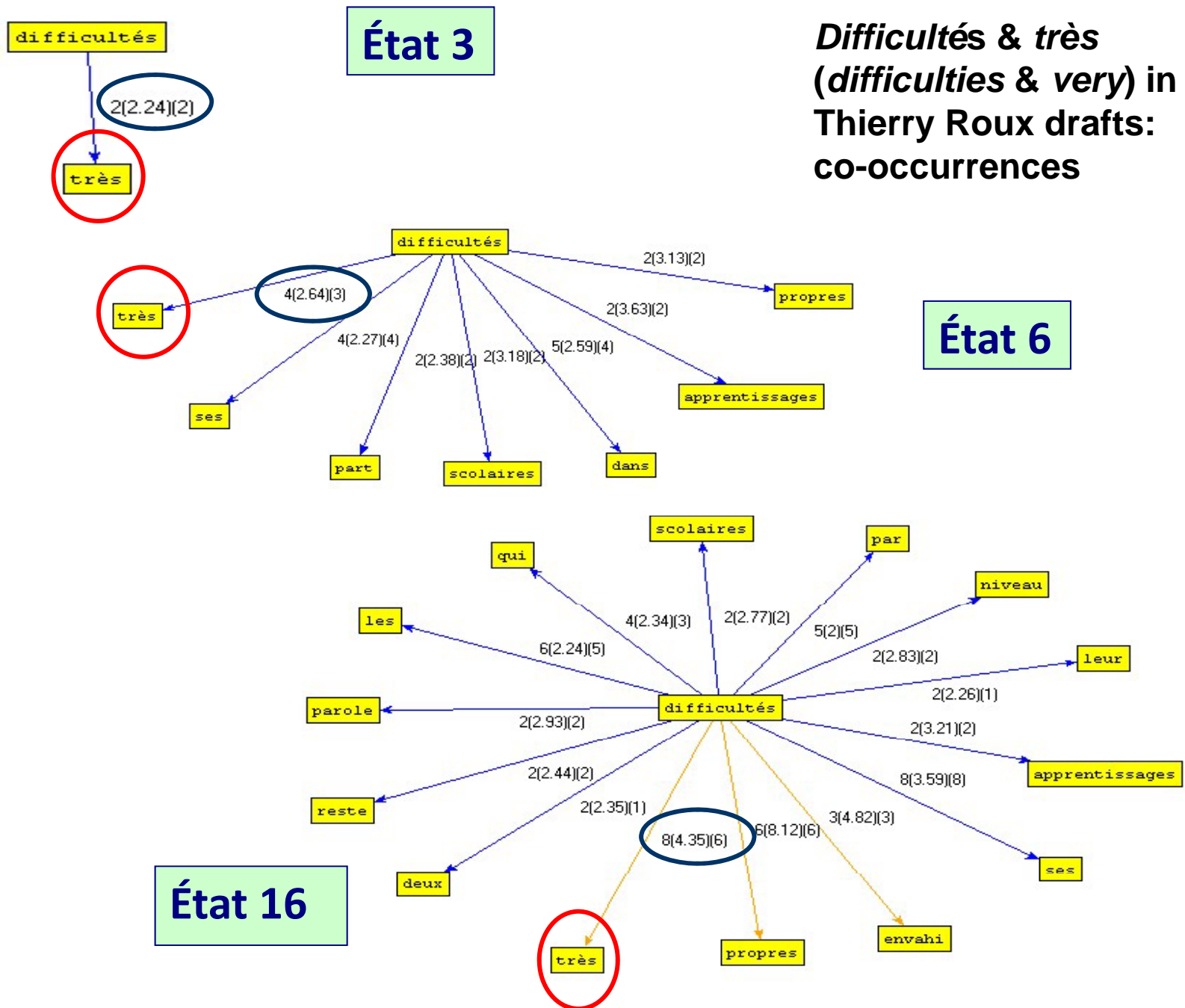
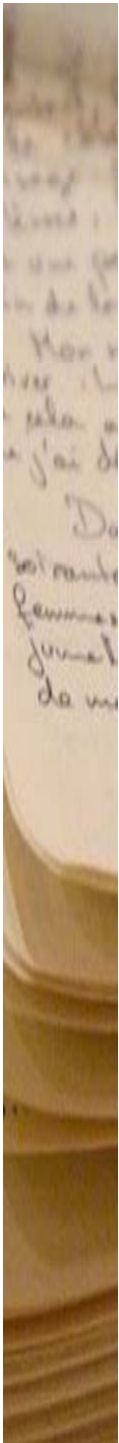
(draft 14, relationships between the four children and their parents)

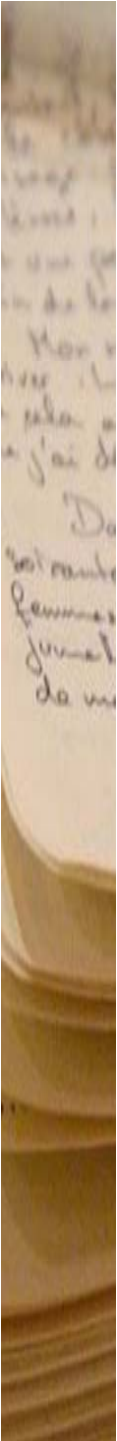
4. « A l'inverse elle [la mère] évoque **ses difficultés propres** (financières, de conflit, de voisinage) ».

(draft 14, behavior of the mother when visiting her children)

.....
Tr. 1. she catches « with difficulties » **her father's** attention **who is very overcome by his own difficulties**

Tr. 2. : Furthermore, **Mister who is very** worried with **his own difficulties has invaded** the entire talking area, so much so that her daughter can't express herself





(1) En troisième, **Evelyne** a rencontré **des difficultés croissantes**, [right co-text =>] dans les apprentissages, notamment en mathématiques et en anglais.

Tr.: In 9th grade, **Evelyne [the child]** has met **increasing difficulties**, **with learning mathematics and English**.

(2) **Madame** est en **grande difficulté** pour se mobiliser autour des besoins de ses enfants et de l'exercice de son autorité parentale.

Tr.: **Madam** has **many difficulties** to mobilize herself to her children's needs and to exercise parental rights

(3) En effet, **Monsieur** est **très** envahi par ses **difficultés** propres et s'il revendique les droits de visite, il les investit **surtout** comme un lieu de parole pour lui.

Tr.: In fact, **Mister is very** overcome by his own **difficulties** and while he's claiming visiting rights, he « invests » it as a talking place for himself **above all**



Conclusion of Thierry Roux report

Les relations des deux parents avec leurs enfants sont **très** anxiogènes pour eux, chacun à leur niveau **en fonction de ses difficultés propres** il ressort des deux années écoulées une évolution des quatre enfants, qui reste **extrêmement** fragile, empreinte d'un fonctionnement familial qui remet en cause régulièrement leur développement psychique et nécessite un étayage éducatif et psychologique **très** important. Il est **fondamental** de maintenir le dispositif actuel de placement qui garantit la protection des enfants.

Tr.: The relationship of parents with their children creates **alot of** anxiety for them, each at their level and **in accordance with their own difficulties** over the course of the past two years there appears to be an evolution of the four children who remain **extremely** fragile, marked by family behavior that regularly questions their psychological development and necessitates both a **very** important psychological and educative follow-up. It is **fundamental** to maintain the current placement that guarentees the protection of the children.

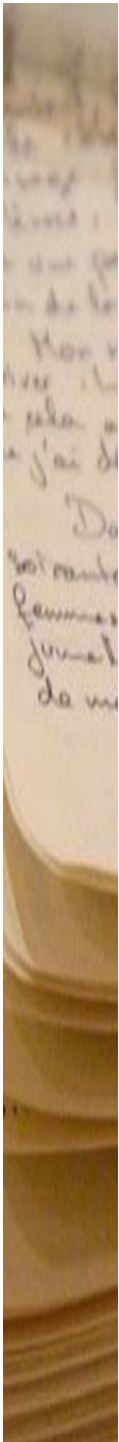
To conclude...

- **Methodology**

- Textometry sheds light on the inherent cycle-like behavior of discourse revealing persuasive strategies to convince a judge of the best possible future for the child.
- Statistics make visible these discursive patterns that could have otherwise gone undetected if we were simply to read the reports. This is the case with the “insistence” phenomena seen in the presentation

- **From a pragmatic point of view**

- The social workers leave discursive “clues” while trying to code them in such a way that they would not be hurtful to the families or child they concern - the use of the vague repeated word “difficult”.
- These reports are also part of the routine of any social worker’s daily tasks.
- The use of the such vague fixed expression, beyond being part of the professional jargon, could be a way for the social workers to emotionally distance themselves from the situation they are writing about.





Thank you for your attention

SYLED

SYstèmes Linguistiques, Enonciation et Discours (EA2290)

