



Ce document est publié sous licence
Creative Commons Attribution 4.0
<https://creativecommons.org/licenses/by/4.0/>

Quelles observations dans TXM peuvent éclairer l'hypothèse de la construction en entonnoir des plans des Actualités françaises

Bénédicte PINCEMIN (Université de Lyon, CNRS, Laboratoire IHRIM)

Document de travail de BP suite à la réunion téléphonique FM + BP du 4 novembre 2021.

Sommaire

1. MAIS COMMENT SE PRÉSENTENT LES DESCRIPTIONS PAR PLANS DANS NOTRE CORPUS ?.....	2
1.1. AF-PLANS.....	2
1.1.1. Enchaînement des plans.....	2
1.1.2. Composition des descriptions : les descriptions « uniformes ».....	3
1.2. AF-NOTICES.....	4
1.3. d'AF-PLANS à AF-NOTICES.....	4
2. RECHERCHE ET DÉCOMPTE AUTOMATIQUE DE MOTIFS ENTONNOIR.....	5
2.1. Principe général et notations.....	5
2.2. Requête pour 3 types de plan différents en entonnoir.....	5
2.3. Requête pour rechercher les débuts entonnoirs avec Zoom.....	7
2.4. Requête 3 plans en entonnoir incluant les cas de zoom.....	7
2.5. Pourquoi a-t-on si peu de résultats ?.....	8
2.5.1. Cas a : quelle est la part des descriptions très/trop courtes ? combien reste-t-il de sujets sur lesquels il est possible d'observer un entonnoir ?	8
2.5.2. Cas b : Combien a-t-on de descriptions monotones, sans variation de plans ni dans un sens (entonnoir) ni dans un autre (élargissement ou désordre) ?.....	9
2.6. Requête pour 2 types de plan différents en entonnoir (et zoom).....	9
2.7. Combien a-t-on en tout de sujets de description comportant au moins 2 plans ?.....	11
2.8. Et combien a-t-on de sujets d'au moins deux plans qui sont monotones ?.....	11
2.9. Bilan sur la répartition dans les 3 catégories de début de sujet.....	12
2.10. Épilogue 1 : Aurions-nous eu des résultats très différents si on avait considéré des entonnoirs "dilués" ?.....	12
2.11. Épilogue 2 : Une requête plus précise pour compter les entonnoirs à 2 plans.....	13

J'ai surligné en jaune ce qui me paraît le plus important :

- §1 et §1.1 : il est essentiel de voir concrètement à quoi ressemble le corpus AF-PLANS et donc les données sur lesquelles on lance TXM.
- §2.1 : la manière de s'y prendre, les catégories utilisées, les notations utilisées par la suite.
- §2.9 : ce que je propose comme résultats.

1. MAIS COMMENT SE PRÉSENTENT LES DESCRIPTIONS PAR PLANS DANS NOTRE CORPUS ?

Avant toute chose, prendre vraiment connaissance avec les textes sur lesquels on travaille, tels qu'ils se présentent pour TXM. Et prenons conscience des caractéristiques réelles de nos données, en les observant, pour éviter les fausses idées qu'on pourrait s'en faire "en théorie", à partir de l'impression générale des notices.

C'est essentiel pour ensuite concevoir les recherches qu'il peut être intéressant de formuler, et interpréter les résultats.

1.1. AF-PLANS

- on peut feuilleter et visualiser les séquences de plans de chaque notice avec la fonction EDITION (icône livre)
- pour aller plus facilement à des endroits/dates varié(e)s dans le corpus, faire une concordance sur

<div>[]

et double-cliquer sur une ligne (on voit la date à gauche) pour visualiser l'édition correspondante. Continuer de la sorte (double-clic sur différentes lignes de la concordance) pour naviguer dans l'édition en accédant directement à différentes périodes.

Rq.1 Si ce n'est déjà le cas, il est commode d'afficher la concordance à gauche et l'édition à droite (plutôt que la concordance en dessous et l'édition au dessus). On peut régler cela en glissant-déposant l'onglet de l'édition vers le bord droit de la concordance Ou bien régler la préférence pour avoir toujours l'ouverture de l'édition à droite de la concordance qui l'appelle :

Menu général Edition > Préférences > TXM > Utilisateur > Édition (après Table lexicale)

new editor position : choisir l'emplacement voulu par ex. RIGHT_OF pour un affichage à droite de la concordance.

Ce réglage ne s'applique pas aux éditions déjà ouvertes -> fermer et réouvrir une nouvelle édition.

Rq.2 Pour ma part par exemple, je trouve que cela permet de prendre conscience que les descriptions sont souvent courtes, et aussi souvent étonnamment uniforme (des suites de 00DP, ou de 10PG, ou de 32PP, etc.), comme si le/la documentaliste avait tendance, une fois sa description commencée, à noter toujours le même type de plan. Est-ce un biais de description, ou a-t-on vraiment dans la vidéo une tendance naturelle ou voulue à ne pas varier les plans ?

Avec TXM on peut enquêter avec deux observations, l'une sur l'enchaînement immédiat des catégories avec les COOCCURRENCES (cf. paragraphe 1.1.1) ci-après), l'autre avec la composition d'ensemble des descriptions avec un INDEX (cf. paragraphe 1.1.2) ci-après).

1.1.1. Enchaînement des plans

On calcule quels sont statistiquement les plans sur-représentés juste après un plan donné, en tenant compte de leur fréquence.

Mode opératoire : fonction COOCCURRENCES

paramètres par défaut sauf :

- décocher "utiliser le contexte gauche"

- empan du contexte droit : de 0 à 0

(c'est-à-dire qu'on ne considère que le mot à distance 0, immédiatement voisin).

En requête, je mets par exemple les plans assimilés aux plans généraux :

"10PG|40VA|50VE|31PANO"

Le résultat présente comme successeurs privilégiés exactement le même ensemble de valeurs.

Si on se focalise sur 10PG seul, il appelle massivement un autre 10PG (indice de 127), éventuellement un 13PR (indice de 2).

Voici quelques enchaînements mis en évidence :

[word="10PG|40VA|50VE|31PANO"] -> les mêmes : 31PANO (indice 258), 10PG (127), 40VA (110), 50VE (6)

[word="11PL|51VI"] -> 11PL (45), 12PM (4) i.e. même catégorie + 1 entonnoir

[word="12PM|32PP"] -> les mêmes : 32PP (127), 12PM (127)
[word="13PR"] -> 13PR (indice 127), 32PP (indice 30, 501 occurrences de la succession 13PR 32PP, cas qui irait plutôt à l'inverse de l'entonnoir mais pas forcément en début de sujet)
[word="14GP"] -> 14GP (indice 127), 60ZOOM (6).

Bilan de ces observations d'enchaînement : répétition très marquée du même, en particulier pour les catégories les plus massivement représentées : les 10PG et assimilés (32529 occ / 91898), les 12PM et 32PP totalisant ensemble 19743 occ., et les 14GP (16234 occ.). en comparaison de

[word="13PR"] : 2888 occ.
[word="11PL|51VI"] : 322 occ.

1.1.2.Composition des descriptions : les descriptions « uniformes »

Ci-après on a mise au point une requête qui cherche des résumés ou des descriptions d'au moins deux plans, mais ne comportant que des plans d'une certaine catégorie ou des plans disons non-marqués c'est à dire hors des 6 catégories définissant l'entonnoir.

Mise au point de la requête (avec des retours à la ligne pour s'y retrouver visuellement) :

```
(<resume>
[word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,}
)</resume>)(<sequences>
[word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,}
)</sequences>
```

Requête utilisable dans TXM (en remettant tout sur une seule ligne)

```
(<resume>([word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,})</resume>)(<sequences>([word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,})</sequences>
```

Si on demande l'INDEX de cette requête, on trouve 2065 occ., et on visualise toutes les formes prises par cette description "uniforme", à commencer par les plus courantes :

00DP	00DP	128	
10PG	00DP	106	
10PG	10PG	76	
32PP	32PP	65	
10PG	10PG	10PG	37

```
70TI 10PG    34
00DP 00DP 00DP    32
10PG 10PG 00DP    30
etc.
```

Notons bien qu'il s'agit là de descriptions complètes (ici donc se réduisant à 2 ou 3 plans), et non d'extraits de descriptions.

Revenons à nos observations en corpus, on peut aussi avoir envie de retourner au texte des notices lui-même et à la présence des valeurs de plan dans ces textes :

1.2. AF-NOTICES

CONCORDANCE sur

```
[plan!="__UNDEF__|01HS"]
```

et visualisation des notices en double-cliquant sur les lignes, en essayant de varier les périodes.

L'édition s'affiche avec en surligné rose les occurrences qui ont été annotées avec une valeur de plan pour AF-PLANS.

Cela permet de voir à quoi ressemblent les notices textuellement et comment apparaissent les valeurs de plan.

Noter cependant le biais que plus une description a de mentions, plus elle occupe de lignes dans la concordance.

Autrement dit on a plus de chance de tomber au hasard sur une notice longue avec beaucoup de plans annotés.

Pour palier ce biais, on peut utiliser d'autres requêtes pour chercher les notices formulées avec peu ou pas de valeurs de plan.

```
<div>[plan="__UNDEF__|01HS"]+</div>
```

-> 480 sujets sans aucune description avec indications techniques de plan (les séquences peuvent prendre la forme de succession d'alinéas).

```
<div>[plan="__UNDEF__|01HS"]*@[plan!="__UNDEF__|01HS"]</div>
```

748 sujets avec une seule mention de plan

```
<div>[plan="__UNDEF__|01HS"]*@[plan="00DP"]</div>
```

266 sujets sont le seul plan est un "divers plans".

Pour voir les 482 autres sujets :

```
<div>[plan="__UNDEF__|01HS"]*@[plan!="__UNDEF__|01HS|00DP"]</div>
```

1.3. d'AF-PLANS à AF-NOTICES

Quand on visualise un "texte" d'AF-PLANS, on peut avoir envie de voir comment il était à l'origine, avec ses mots, dans AF-NOTICES.

Circuit possible :

- sélectionner AF-NOTICES et demander l'ÉDITION (livre).
- placer cette édition à côté de celle d'AF-PLANS en tirant l'onglet.
- copier l'identifiant du journal qui s'affiche dans la barre de navigation en bas de l'édition AF-PLANS, sur la droite.
- coller cet identifiant à l'endroit analogue dans l'édition d'AF-NOTICES, et appuyer sur la touche entrée.

Cela positionne l'édition au début du journal concerné, reste à aller au sujet.

Dans l'édition AF-PLANS, regarder à quelle page de l'émission est le sujet (dans barre de navigation en bas de la fenêtre, mais cette fois-ci vers le centre, à gauche de l'identifiant de l'émission).

Avancer (avec la flèche simple) ou aller directement (en mettant le numéro + valider avec "entrée")

au même numéro de page dans l'édition d'AF-NOTICES.

2. RECHERCHE ET DÉCOMPTÉ AUTOMATIQUE DE MOTIFS ENTONNOIR

Rq. Finalement il ne m'a plus semblé nécessaire de recourir à la stratégie CQP "longest".
Toutes les requêtes ci-après se font avec la stratégie par défaut au lancement de TXM.

2.1. Principe général et notations

On a défini 6 grands types de plans "significatifs" du point de vue du procédé de changement d'échelle qu'on cherche à observer :

```
[plan0] aussi noté en abrégé 0 : [word="10PG|40VA|50VE|31PAN0"]  
[plan1] ou 1 : [word="11PL|51VI"]  
[plan2] ou 2 : [word="12PM|32PP"]  
[plan3] ou 3 : [word="13PR"]  
[plan4] ou 4 : [word="14GP"]  
[plan5] ou 5 : [word="15TGP"]
```

Un 7e type regroupe les plans qui ne "comptent" pas du point de vue du changement d'échelle,
et qu'on autorisera à s'intercaler, mais en quantité limitée (typiquement 2 max.) :

```
autre :  
[autre] ou X : [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]
```

Enfin un traitement à part est fait pour le zoom,
qui traduit lui-même déjà à lui seul tout ou partie d'un entonnoir :

```
[zoom] ou Z : [word="60ZOOM"]
```

En gros on va donc chercher des suites de plans :

- avec des chiffres croissants,
- avec des X qui peuvent s'intercaler au début ou entre deux chiffres
- le Z peut s'insérer partout, mais on considère que sa simple présence rend compte d'un entonnoir et on n'a pas besoin de poursuivre la recherche.

ex. de motifs à repérer *à partir du début du résumé ou des séquences* : 1134, 2XX4, X11X1XX1XX22X2, X34Z, etc.

2.2. Requête pour 3 types de plan différents en entonnoir

Mise au point du motif à chercher : on balaye systématiquement toutes les combinaisons de chiffres possibles, soit :

```
012, 013, 014, 015,  
123, 124, 125,  
234, 235,  
345.
```

Chaque chiffre peut être répété : 000112, 233355 etc.

Puis on ajoute la possibilité d'intercalation de 0 à 2 X au début ou après n'importe quel chiffre :

```
ex. 00XX01X12, XX23X33XX55
```

Dès qu'on trouve le 3e chiffre différent on a trouvé ce qu'on veut donc on arrête la recherche :

```
ex. en fait on ne trouvera pas précisément XX23X33XX55 parce qu'on se sera arrêté à XX23X33XX5 (qui inclut le cas plus précis  
XX23X33XX55).
```

on écrit d'abord la requête en allant à la ligne pour y voir clair, on peut simplifier le dernier terme (on s'aperçoit qu'il n'était pas régulier dans la requête précédente : 5 traité différemment des autres) :

```
(<resume>|<sequences>) [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}(
```

```

(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="12PM|32PP"] ) |
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="13PR"] ) |
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="14GP"] ) |
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="15TGP"] ) |
(( [word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="13PR"] ) |
(( [word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="14GP"] ) |
(( [word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="15TGP"] ) |
(( [word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="14GP"] ) |
(( [word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="15TGP"] ) |
(( [word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
([word="14GP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="15TGP"] )
) within div

```

En recollant tout sur une seule ligne, cela donne la requête qu'on peut mettre dans TXM :

```

(<resume>|<sequences>) [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} (( ([word="10PG|40VA|50VE|31PANO"] [word="00DP|
20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="12PM|32PP"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="11PL|
51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="13PR"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|
20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="14GP"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="11PL|51VI"]
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="15TGP"] ) | (( [word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|
70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="13PR"] ) |
(( [word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|
70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="14GP"] ) | (( [word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="15TGP"] ) | (( [word="12PM|32PP"] [word="00DP|
20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+
[word="14GP"] ) | (( [word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ ([word="13PR"] [word="00DP|20PLON|
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="15TGP"] ) | (( [word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]
{0,2} )+ ([word="14GP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"] {0,2} )+ [word="15TGP"] )) within div

```

==> 40 occurrences

Déception : très peu sujets (40 sur 10 000) présentant le motif entonnoir avec 3 valeurs.
Mais on n'a pas encore pris en compte les zooms :

2.3. Requête pour rechercher les débuts entonnoirs avec Zoom

Mise au point du motif à chercher :

- le zoom peut éventuellement précédé de chiffres, tant qu'ils sont croissants (ils peuvent aussi être répétés)
- on autorise toujours l'intercalation de X, mais jamais plus de 2 d'affilée

Donc en gros cela donne :

peut-être un ou deux X, puis

peut-être un ou plusieurs 1 (avec max. 2 X intercalés entre chaque), puis

peut-être un ou plusieurs 2 (avec max. 2 X intercalés entre chaque), puis

peut-être un ou plusieurs 3 (avec max. 2 X intercalés entre chaque), puis

peut-être un ou plusieurs 4 (avec max. 2 X intercalés entre chaque), puis

peut-être un ou plusieurs 5 (avec max. 2 X intercalés entre chaque), puis

le Z cherché.

Écriture avec des retours à la ligne pour s'y retrouver :

```
(<resume>|<sequences>) [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}
([word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*
([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*
([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*
([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*
[word="60ZOOM"]
```

En recollant tout sur une seule ligne, cela donne la requête qu'on peut mettre dans TXM :

```
(<resume>|<sequences>) [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}([word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*[word="60ZOOM"] within div
```

==> 55 occurrences

Les zooms (n')ajoutent (que) 55 sujets.

Ce n'est pas énorme, vue qu'on est moins exigeant,
on n'a pas besoin de beaucoup de valeurs de plan différentes,
même une seule mention de zoom suffit, si elle est proche du début du sujet.

2.4. Requête 3 plans en entonnoir incluant les cas de zoom

D'où requête unique qui couvre à la fois le cas de 3 plans en entonnoir, ou un zoom précédé éventuellement de plans dans le bon ordre

```
(<resume>|<sequences>) [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}((( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="12PM|32PP"]) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="13PR"]) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="14GP"]) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="11PL|51VI"]
```

```
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] | (([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"]) | (([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"]) | (([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"]) | (([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"]) | (([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"]) | (([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+([word="10PG|40VA|50VE|31PAN0"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}))*([word="11PL|51VI"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}))*([word="12PM|32PP"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}))*([word="13PR"] [word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2}))*[word="60ZOOM"] ) ) within div
```

====> 95 occ

(C'est ce qu'on attendait : 40 + 55).

C'est intéressant de voir le résultat de cette requête en INDEX :

cela montre les cas concrets de réalisation de ce motif "entonnoir".

On peut notamment trier l'INDEX alphabétiquement, sur la première colonne, comme cela on voit des familles de motifs :

tous ceux qui commencent par des divers plans, ou un plan général, etc.

Qualitativement, c'est intéressant, mais quantitativement, on semble loin de chiffres à l'appui de la thèse de l'introduction en entonnoir.

2.5. Pourquoi a-t-on si peu de résultats ?

On a de premières idées après être allé feuilleter le corpus et voir à quoi ressemblaient les descriptions

(cf. première partie au début de ce document) :

a- peut-être que beaucoup de descriptions sont très courtes : une description de 1 ou 2 plans ne peut pas réaliser un entonnoir de 3 catégories de plan (sauf cas de zoom).

b- beaucoup de descriptions pourraient être "monotones", c'est-à-dire répètent toujours le même type de plan

2.5.1. Cas a : quelle est la part des descriptions très/trop courtes ? combien reste-t-il de sujets sur lesquels il est possible d'observer un entonnoir ?

Recherche des sujets avec une description d'au moins 3 plans :

```
(<resume>[] {3,}</resume>)|(<sequences>[] {3,}</sequences>)
```

On met cette requête en INDEX, comme cela on a un aperçu du genre de séquences qu'on trouve.

On peut être frappé par le nombre de répétitions du même plan (cf. Cas b-)

====> 8 611 occurrences

On a tout de même de l'ordre de 8 000 sujets avec une description ayant au moins 3 mention de plan.

La requête peut sur-évaluer légèrement le nombre de sujets, car un même sujet pourrait avoir été compté deux fois, si on a au moins 3 mentions de plan dans son résumé et 3 mentions de plan dans ses séquences.

On peut d'ailleurs le mesurer précisément :

```
<resume>[] {3,}</resume><sequences>[] {3,}</sequences> within div
```

---> seuls 15 sujets ont des descriptions d'au moins 3 plans à la fois dans le résumé et les séquences.

Soit donc exactement 8596 sujets d'au moins 3 plans.

2.5.2.Cas b : Combien a-t-on de descriptions monotones, sans variation de plans ni dans un sens (entonnoir) ni dans un autre (élargissement ou désordre) ?

Motif à rechercher : un mélange de X et d'un seul type de chiffre
(il peut n'y avoir que des X ou que des chiffres, mais quand il y a un chiffre c'est toujours le même).
ex. XXXXX, 00000, XX3X3XXXXX3333

Recherche des constants d'au moins 3 plans :

Écriture avec des retours à la ligne pour s'y retrouver :

```
(<resume>(  
[word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,}  
)</resume>)|(<sequences>(  
[word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} |  
[word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,}  
)</sequences>)  
  
(<resume>([word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="11PL|51VI|00DP|20PLON|21CPL|  
30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="13PR|00DP|20PLON|  
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="15TGP|00DP|  
20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,})</resume>)|(<sequences>([word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|  
71BT|72GR|73FLOU"]{3,} | [word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="12PM|32PP|00DP|20PLON|21CPL|  
30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="14GP|00DP|20PLON|21CPL|  
30TRAV|70TI|71BT|72GR|73FLOU"]{3,} | [word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{3,})</sequences>)
```

====> 1352 occ.

Donc sur 8596 sujets d'au moins 3 plans, 1352 ne présentent pas de variation de catégorie.

Vue l'allure de nos données, et la maigre récolte en cherchant des entonnoirs un peu développés à 3 catégories différentes, il semble intéressant déjà de s'intéresser au mouvement en début de sujet, sur 2 catégories de plan :
au début du sujet, la première variation va-t-elle dans le sens de l'élargissement, ou du focus ?
(le 3e cas étant quand on n'a pas de variation du tout).

2.6. Requête pour 2 types de plan différents en entonnoir (et zoom)

Mise au point du motif à chercher : on balaye systématiquement toutes les combinaisons de chiffres possibles, soit :
01, 02, 03, 04, 05,
12, 13, 14, 15,

23, 24, 25,
34, 35,
45.

Chaque chiffre peut être répété : 00011, 2333 etc.

Puis on ajoute la possibilité d'intercalation de 0 à 2 X au début ou après n'importe quel chiffre :

ex. 00XX01X1, XX23X33

Dès qu'on trouve le 2e chiffre différent du 1er on a trouvé ce qu'on veut donc on arrête la recherche :

ex. en fait on ne trouvera pas précisément 00XX01X1 parce qu'on se sera arrêté à 00XX01 (qui inclut le cas plus précis 00XX01X1).

Idem pour XX23X33, on s'arrête à XX23.

Écriture avec des retours à la ligne pour s'y retrouver :

```
(<resume>|<sequences>)[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2} ((  
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="11PL|51VI"] ) |  
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="12PM|32PP"] ) |  
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"] ) |  
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"] ) |  
(( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) |  
(( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="12PM|32PP"] ) |  
(( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"] ) |  
(( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"] ) |  
(( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) |  
(( [word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"] ) |  
(( [word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"] ) |  
(( [word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) |  
(( [word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"] ) |  
(( [word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) |  
(( [word="14GP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] )  
 ) | (  
[word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*  
[word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*  
[word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*  
[word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*  
[word="60ZOOM"] ) ) within div
```

```
(<resume>|<sequences>)[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2} ((( ([word="10PG|40VA|50VE|31PANO"] [word="00DP|  
20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="11PL|51VI"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|  
30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="12PM|32PP"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|  
72GR|73FLOU"]{0,2})+[word="13PR"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+  
[word="14GP"] ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) |  
(( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="12PM|32PP"] ) | (( [word="11PL|51VI"]  
[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"] ) | (( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|  
70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"] ) | (( [word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+  
[word="15TGP"] ) | (( [word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"] ) | (( [word="12PM|  
32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"] ) | (( [word="12PM|32PP"] [word="00DP|20PLOW|21CPL|  
30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) | (( [word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+  
[word="14GP"] ) | (( [word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) | (( [word="14GP"]  
[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"] ) ) | (( [word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLOW|  
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})* ([word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]
```

```
{0,2})*([word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*([word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})*[word="60ZOOM"]) within div
```

====> 5599 occ.

Rien à voir avec les résultats pour 3 plans (5599 >> 95), ici c'est beaucoup plus massif !

Mais en même temps finalement c'est assez normal, car il n'y a que 3 cas possibles :

- soit une configuration monotone (0 ou 1 seule catégorie de plan, donc pas de variation),
- soit la 2e catégorie est plus resserrée que la première (ce qu'on vient de mesurer ici),
- soit la 2e catégorie est plus large que la première.

Pour pouvoir vraiment comparer l'équilibre de ces trois catégories, il faudrait se limiter à l'ensemble total des sujets sur lesquels elles peuvent être observées : il faut que le sujet ait au moins 2 plans.

2.7. Combien a-t-on en tout de sujets de description comportant au moins 2 plans ?

```
(<resume>[]{2,}</resume>)|(<sequences>[]{2,}</sequences>)  
9 580 occurrences.
```

Voyons combien de sujets sont comptés deux fois

(car avec à fois un résumé et une séquence d'au moins deux plans) :

```
<resume>[]{2,}</resume><sequences>[]{2,}</sequences> within div  
29 occurrences
```

====> On a 9551 sujets ayant une description d'au moins 2 plans.

Donc en gros pour 5600 sujets / 9600 décrits par au moins 2 plans, on a un entonnoir (chiffres précis : 5599 / 9 580)

Le cas des zooms fait qu'en théorie on pourrait aussi avoir des résultats d'entonnoir pour une description d'un seul plan qui serait 60ZOOM, cependant en pratique le cas ne se rencontre pas, les deux requêtes suivantes n'ont aucun résultat :

```
<resume>[word="60ZOOM"]</resume>  
<sequences>[word="60ZOOM"]</sequences>
```

Donc tous les résultats de la requête entonnoir font partie des sujets à au moins deux plans.

2.8. Et combien a-t-on de sujets d'au moins deux plans qui sont monotones ?

```
(<resume>( [word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |  
[word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |  
[word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |  
[word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |  
[word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |  
[word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,}  
)</resume>)|(<sequences>
```

```
[word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} |
[word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,}
)/</sequences>
```

```
(<resume>([word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="11PL|51VI|00DP|20PLON|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="12PM|32PP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="13PR|00DP|20PLON|
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="14GP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="15TGP|00DP|
20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,})</resume>)|(<sequences>([word="10PG|40VA|50VE|31PANO|00DP|20PLON|21CPL|30TRAV|70TI|
71BT|72GR|73FLOU"]{2,} | [word="11PL|51VI|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="12PM|32PP|00DP|20PLON|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="13PR|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="14GP|00DP|20PLON|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"]{2,} | [word="15TGP|00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{2,})</sequences>)
```

====> 2065 occ.

Donc en gros pour 2000 sujets / 9600 décrits par au moins 2 plans, on n'a pas de variation de plan indiquée.

2.9. Bilan sur la répartition dans les 3 catégories de début de sujet

Nombre total de sujets ayant une description à au moins deux plans
(donc sur lesquels l'observation de variation ou monotonie est possible) : **9551**
Parmi ceux-ci,

- Nombre de sujets présentant *en début de sujet* une évolution en entonnoir,
d'un plan donné à un plan moins large : **5582**, soit 58 %
(plus précis que 5599, cf. épilogue 2 ci-après)

- Nombre de sujets présentant une configuration *entièrement monotone*
(du début à la fin) : **2065**, soit 22 %

- Sujets restants (variation de plan qui s'élargit en début de sujet,
ou variation plus tardive en entonnoir ou non) : **1904**, soit 20 %

2.10. Épilogue 1 : Aurions-nous eu des résultats très différents si on avait considéré des entonnoirs "dilués" ?

On ne limite plus le nombre de X qui peuvent s'intercaler, que ce soit en début de sujet ou entre deux mentions de plans.
L'intérêt, c'est que le résultat exclut toujours un pattern marqué,
soit de variation désordonnée soit de type élargissement.
On n'a que des entonnoirs.

```
(<resume>|<sequences>)[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]* ((([word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="11PL|51VI"])| (([word="10PG|40VA|50VE|31PANO"] [word="00DP|20PLON|21CPL|30TRAV|70TI|
```

```

71BT|72GR|73FLOU"]*)+[word="12PM|32PP"]) | ((([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+
[word="13PR"]) | (([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="14GP"]) |
((([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="15TGP"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="12PM|32PP"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="13PR"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="14GP"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="15TGP"]) | (([word="12PM|32PP"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="13PR"]) | (([word="12PM|32PP"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="14GP"]) | (([word="12PM|32PP"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="15TGP"]) |
((([word="13PR"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="14GP"]) | (([word="13PR"])[word="00DP|20PLON|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"]*)+[word="15TGP"]) | (([word="14GP"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]*)+
[word="15TGP"]))) within div

```

====> 5686 occ

(en gros 100 occ de plus avec la dilution)

Donc pas de différence énorme finalement,
la contrainte du nombre de X qui s'intercalent ne semble pas être très sensible,
le choix de mettre 2 n'a pas une incidence très forte sur les proportions trouvées.

2.11.Épilogue 2 : Une requête plus précise pour compter les entonnoirs à 2 plans

On pourrait nous objecter qu'en l'état,
la requête peut commencer dans un résumé et finir dans les séquences ;
ou encore, elle pourrait se réaliser deux fois dans un même sujet (dans son résumé et ses séquences).

Cela semblait peu probable vu ce que nous connaissons des notices.
En effet, les descriptions plan à plan ne sont que dans un seul des deux champs :
normalement les séquences, et parfois il n'y a pas de séquences et
la description plan à plan est faite directement dans le résumé.
Par ailleurs, les valeurs de plan qui comptent ici ont des formes assez typées (ex : "PG", "gros plan")
qui ont peu de chances de se trouver dans une rédaction qui ne consiste pas en une description plan à plan.

Mais bon, soyons précis.

La recherche pour s'assurer de ne compter que des réalisations au sein d'un des deux champs (résumé ou séquences), et pas à cheval sur
les deux :

```

Construction de la requête avec un retour à la ligne pour bien manier les deux parties :
(<resume>[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2} (((([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="11PL|51VI"]) | (([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|
72GR|73FLOU"]{0,2})+[word="12PM|32PP"]) | (([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]
{0,2})+[word="13PR"]) | (([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"]) |
(([word="10PG|40VA|50VE|31PANO"])[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="12PM|32PP"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"]) | (([word="11PL|51VI"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"]) |
(([word="12PM|32PP"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="13PR"]) | (([word="12PM|32PP"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="14GP"]) | (([word="12PM|32PP"|
[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"]{0,2})+[word="15TGP"])))

```



```

73FLOU"}{0,2})+[word="15TGP"] | ((([word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR"] |
(( [word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"] | (([word="12PM|32PP"] [word="00DP|
20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"] | (([word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|
73FLOU"}{0,2})+[word="14GP"] | (([word="13PR"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"] |
(( [word="14GP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"]))) | (([word="10PG|40VA|50VE|31PANO"
[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) * ([word="11PL|51VI"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|
73FLOU"}{0,2}) * ([word="12PM|32PP"] [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) * ([word="13PR"] [word="00DP|20PLOW|
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) * [word="60ZOOM"]))) ] * </sequences>
5593 occ.

```

Cherchons les descriptions à cheval pour vérifier qu'on retombe bien sur le même ordre de grandeur.
Elles commencent au début du résumé et le dernier mot est dans les séquences.

Écriture avec des retours à la ligne pour s'y retrouver :

```

<resume>[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2} ((
(( [word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="11PL|51VI" & sequences]) |
(( [word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="12PM|32PP" & sequences]) |
(( [word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR" & sequences]) |
(( [word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) |
(( [word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) |
(( [word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="12PM|32PP" & sequences]) |
(( [word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR" & sequences]) |
(( [word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) |
(( [word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) |
(( [word="12PM|32PP" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR" & sequences]) |
(( [word="12PM|32PP" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) |
(( [word="12PM|32PP" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) |
(( [word="13PR" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) |
(( [word="13PR" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) |
(( [word="14GP" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences])
) | (
([word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) *
([word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) *
([word="12PM|32PP" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) *
([word="13PR" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2}) *
[word="60ZOOM" & sequences])) within div

```

```

<resume>[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2} (((([word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="11PL|51VI" & sequences]) | (([word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="12PM|32PP" & sequences]) | (([word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|
30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR" & sequences]) | (([word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|
70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) | (([word="10PG|40VA|50VE|31PANO" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|
72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) | (([word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+
[word="12PM|32PP" & sequences]) | (([word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR" &
sequences]) | (([word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) |
(( [word="11PL|51VI" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) | (([word="12PM|32PP"
[word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR" & sequences]) | (([word="12PM|32PP" [word="00DP|20PLOW|
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP" & sequences]) | (([word="12PM|32PP" [word="00DP|20PLOW|21CPL|30TRAV|70TI|71BT|

```

72GR|73FLOU"}{0,2})+[word="15TGP" & sequences]) | ((([word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+
[word="14GP" & sequences]) | ((([word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences])
| ((([word="14GP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP" & sequences])) | ((([word="10PG|40VA|50VE|
31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*([word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|
72GR|73FLOU"}{0,2})*([word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*([word="13PR"][word="00DP|20PLON|
21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*[word="60ZOOM" & sequences])) within div
72 occ.

Il y en a plus que ce qu'on pensait, mais en fait dans ces entonnoirs à cheval, ceux qui commencent par un ou deux X dans le résumé et se poursuivent ensuite dans les séquences sont en fait des entonnoirs valides et qui, avec la requête plus précise, restent comptabilisés, cette fois-ci au niveau des seules séquences.

Requête pour dénombrer les sujets qu'on aurait compté deux fois :

```
<resume>[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2} (((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|  

30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="11PL|51VI"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|  

72GR|73FLOU"}{0,2})+[word="12PM|32PP"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}  

{0,2})+[word="13PR"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"])|  

(( [word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])| ((([word="11PL|51VI"  

[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="12PM|32PP"])| ((([word="11PL|51VI"][word="00DP|20PLON|21CPL|  

30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR"])| ((([word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}  

{0,2})+[word="14GP"])| ((([word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])|  

(( [word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR"])| ((([word="12PM|32PP"][word="00DP|  

20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"])| ((([word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|  

73FLOU"}{0,2})+[word="15TGP"])| ((([word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"])|  

(( [word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])| ((([word="14GP"][word="00DP|20PLON|  

21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|  

71BT|72GR|73FLOU"}{0,2})*( [word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*( [word="12PM|32PP"  

[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*( [word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}  

{0,2})*( [word="60ZOOM"])))]</resume><sequences>[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2} (((([word="10PG|40VA|50VE|  

31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="11PL|51VI"])| ((([word="10PG|40VA|50VE|31PANO"]  

[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="12PM|32PP"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|  

20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|  

70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"])| ((([word="10PG|40VA|50VE|31PANO"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}  

{0,2})+[word="15TGP"])| ((([word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="12PM|32PP"])|  

(( [word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR"])| ((([word="11PL|51VI"][word="00DP|  

20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"])| ((([word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|  

73FLOU"}{0,2})+[word="15TGP"])| ((([word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="13PR"])|  

(( [word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="14GP"])| ((([word="12PM|32PP"][word="00DP|  

20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])| ((([word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|  

73FLOU"}{0,2})+[word="14GP"])| ((([word="13PR"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])|  

(( [word="14GP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})+[word="15TGP"])| ((([word="10PG|40VA|50VE|31PANO"]  

[word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*( [word="11PL|51VI"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|  

73FLOU"}{0,2})*( [word="12PM|32PP"][word="00DP|20PLON|21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*( [word="13PR"][word="00DP|20PLON|  

21CPL|30TRAV|70TI|71BT|72GR|73FLOU"}{0,2})*[word="60ZOOM"])))]</sequences> within div  

11 occ.
```

Donc finalement le nombre précis de sujets débutant en entonnoir (sur deux catégories de plan) est : 5593-11 = 5582

Ça sera déjà pas mal pour aujourd'hui.

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!! THE END ;-) !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!