

Traitement de données pour une détermination de degré d'implantabilité

À partir d'une table de données obtenue après enquête auprès de deux groupes d'experts dans le domaine de la télé-détection aérospatiale, il semble possible de mettre en évidence un indice d'implantabilité en établissant des rapports entre les divers indicateurs : indicateur d'utilisation (U), indicateur d'hésitation (H), indicateur de rejet (R). Si $U > R$ le vocable paraît acceptable/accepté ; si U/R tend vers 0, un constat de refus peut être établi.

Une zone d'incertitude s'ouvre dès lors que U/R tend vers 1, le vocable n'est ni accepté ni rejeté ; si U/R tend vers $+\infty$, le vocable sera considéré en voie d'implantation à condition que H reste faible. En calculant $U/R+H$, on peut rendre compte du degré de résistance du vocable.

L'étude menée à partir des données issues d'un premier groupe d'experts doit être confrontée aux réponses d'un ou d'autres groupe(s) d'experts.

Termes-clés :

Télé-détection aérospatiale, implantation / non implantation, degré d'implantabilité, zone d'incertitude, zone de fluctuation orientée, degré de résistance du vocable.

Dans le cadre du Centre de linguistique et de dialectologie (Clid), une étude à base d'enquête a été conduite en 1992, mettant en évidence un ensemble de données (Rouges-Martinez 1992 ; Rouges-Martinez et Fossat 1994) qui révèle qu'une terminologie très structurée comme celle de la Télé-détection aérospatiale n'obéit pas à un schéma d'implantation unique et homogène mais que le processus d'implantation d'une terminologie déroule des scénarios multiples.

Le terme d'*implantation terminologique* fait référence à une vision du processus achevé d'implantation d'un vocable. Quand on ne peut pas conclure à l'implantation d'un vocable, c'est qu'il se situe dans une zone d'incertitude entre l'implantation et la non implantation terminologique. Pour aborder cette zone d'incertitude, nous avons opté pour un système de calculs à partir de la mise en rapport de certains indices ; les résultats obtenus nous permettent de mettre en avant divers degrés d'implantabilité.

1 Tables de données

Deux groupes d'informateurs ont été sollicités dans le domaine de la télé-détection aérospatiale :

- Le premier groupe est constitué de responsables de formation ou de recherche ;
- Le deuxième groupe est constitué de spécialistes de la discipline impliqués dans l'élaboration et la diffusion des produits de la télé-détection.

Nous les avons interrogés sur un ensemble de termes. Trois types de réponses étaient attendus ;

- Le vocable est déclaré connu et utilisé, il sera assorti d'un indicateur d'acceptation ;
- Le vocable est déclaré mal connu et peu ou pas utilisé, il sera assorti d'un indicateur de refus ;
- Le vocable est déclaré peu employé et le locuteur hésite quand il s'agit de préciser dans quelles conditions d'énonciation il y a recours ; il sera assorti d'un indicateur d'hésitation.

Ces trois indicateurs (U, R, H) sont reportés dans les tableaux numérotés de 1 à 4 ; les valeurs qui sont exprimées dans les colonnes correspondent au nombre d'occurrences où le terme a été déclaré utilisé, rejeté, non activé spontanément (dans le cas de l'hésitation) par les interlocuteurs. Ces valeurs sont des valeurs relatives ; en effet, potentiellement, les interlocuteurs formation-recherche (au nombre de 12) pouvaient déclarer utiliser le terme, le rejeter ou hésiter sur ce terme dans l'ensemble des situations de communication (22) soit 264 possibilités d'occurrences ; de même, du côté des interlocuteurs professionnels (8), il y avait potentiellement parlant 176 possibilités d'occurrences.

2 Corrélation des fréquences d'occurrences

Considérons trois réels non nuls x , y , z tels que x représente la fréquence d'occurrences de l'utilisation, y la fréquence d'occurrences du refus et z la fréquence d'occurrences de l'hésitation.

Si $x > y$, le vocable paraît acceptable mais il reste à savoir dans quelle mesure il est accepté.

Si le rapport $\frac{x}{y}$ tend vers 1, le vocable ne sera ni refusé, ni accepté. Deux hypothèses sont alors envisageables : soit le vocable est en phase d'implantation parce qu'il est récent ou parce qu'il n'a pas d'équivalent réel, soit il est fortement concurrencé par un substitut et nous le considérerons en phase de rejet.

Par ailleurs, dans un tel cas, il sera intéressant de mesurer le degré d'implantation probable du substitut dominant.

Si le rapport $\frac{x}{y}$ tend vers $+\infty$, le vocable sera dit implanté si la fréquence d'occurrences de l'hésitation reste faible. En effet, si z est trop petit pour que le rapport $\frac{x}{z}$ tende vers 1, il est suffisamment grand pour indiquer un seuil fréquentiel significatif.

Lorsque z devient significatif, un ou plusieurs substituts apparaissent. Le vocable est en concurrence avec un ou plusieurs substituts possibles. La première hypothèse voudrait que le vocable «résiste» au(x) substitut(s). En réalité, parmi un ensemble de substituts, il n'existe pas un substitut vraiment dominant. L'un d'entre eux est en pseudo-phase d'implantation. Le vocable de base résiste. Il tendra vers l'abandon dès que le substitut dominant sera reconnu et par conséquent en phase d'implantation.

Tableau n°1 : Relevé des données concernant les interlocuteurs professionnels

Vocables	Utilisation	Refus	Hésitation	$\frac{U}{R}$	$\frac{U}{(R+H)}$
<i>bande spectrale</i>	159	4	12	39,75	9,938
<i>canal spectral</i>	63	13	7	4,846	3,15
<i>capteur</i>	165	1	6	165	23,571
<i>chatoiement</i>	21	51	43	0,412	0,223
<i>cible</i>	129	1	30	129	4,161
<i>classification dirigée</i>	29	25	9	1,16	0,853
<i>classification non dirigée</i>	17	25	10	0,68	0,486
<i>composition colorée</i>	142	21	9	6,762	4,733
<i>détecteur</i>	168	1	7	168	21
<i>déversement radar</i>	24	50	73	0,48	0,195
<i>donnée brute</i>	168	1	7	168	21
<i>donnée d'apprentissage</i>	107	51	7	2,098	1,845
<i>donnée de terrain</i>	162	2	8	81	16,2
<i>donnée géocodée</i>	158	4	14	39,5	8,778
<i>donnée image</i>	115	24	8	4,792	3,594
<i>équidensité colorée</i>	109	27	9	4,037	3,028
<i>fausse couleur</i>	138	2	7	69	15,333
<i>filtrage numérique</i>	125	14	19	8,929	3,788
<i>fouillis d'échos</i>	14	39	66	0,359	0,133
<i>hyperfréquence</i>	152	2	9	76	13,818
<i>imagerie</i>	165	1	10	165	15
<i>imagette</i>	129	1	16	129	7,588
<i>image chronoséquentielle</i>	0	33	50	0	0
<i>image d'écran</i>	103	24	13	4,292	2,784
<i>image diachronique</i>	49	58	43	0,845	0,485
<i>image filtrée</i>	154	13	9	11,846	7
<i>image prétraitée</i>	155	11	10	14,091	7,381
<i>image radar</i>	170	1	5	170	28,333
<i>image satellite</i>	164	1	3	164	41
<i>indice de végétation</i>	128	8	7	16	8,533
<i>infrarouge couleur</i>	143	3	9	47,667	11,917
<i>ligne de balayage</i>	142	5	8	28,4	10,923
<i>lobe d'antenne</i>	113	13	40	8,692	2,132
<i>mosaïque</i>	166	4	6	41,5	16,6
<i>numérisation</i>	167	2	7	83,5	18,556
<i>orthophotoplan</i>	83	33	7	2,515	2,075
<i>pas d'échantillonnage</i>	162	4	10	40,5	11,571
<i>pixel</i>	162	5	9	32,4	11,571

Tableau n°2 : Relevé des données concernant les interlocuteurs professionnels

Vocables	Utilisation	Refus	Hésitation	$\frac{U}{R}$	$\frac{U}{(R+H)}$
<i>rabattement radar</i>	17	48	77	0,354	0,136
<i>radargraphie</i>	40	53	33	0,755	0,465
<i>radiomètre</i>	164	4	8	41	13,667
<i>rayonnement électromagnétique</i>	165	0	11	-	15
<i>réalité de terrain</i>	41	23	8	1,783	1,323
<i>réflectance</i>	161	7	8	23	10,733
<i>pouvoir de résolution</i>	130	9	5	14,444	9,286
<i>retrodiffusion</i>	137	31	8	4,419	3,513
<i>scanner</i>	96	1	5	96	16
<i>scène</i>	169	2	5	84,5	24,143
<i>segment de données</i>	104	24	18	4,333	2,476
<i>signature spectrale</i>	158	9	9	17,556	8,778
<i>spatiocarte</i>	142	19	15	7,474	4,176
<i>tachè</i>	2	23	48	0,087	0,028
<i>thermographie</i>	140	11	25	12,727	3,889
<i>trace de satellite</i>	168	0	8	-	21
<i>traitement d'image numérique</i>	118	0	8	-	14,75
<i>vérité-terrain</i>	127	23	4	5,522	4,704
<i>zone d'apprentissage</i>	102	27	17	3,778	2,318
<i>zone d'essai</i>	41	53	19	0,774	0,569
<i>zone d'entraînement</i>	47	66	6	0,712	0,653

Lorsque $x < y$ le rapport $\frac{x}{y}$ tend vers 0. Le refus est catégorique. Le vocable n'est pas assimilé par les informateurs; l'étude de l'implantation se dirigera vers les substituts existants.

À l'issue de ce dégrossissement, les phases d'implantation, de résistance et/ou de rejet sont en partie définies. La seconde étape du travail consiste, à présent, à calculer un indice d'association du vocable à la phase d'implantation. Cet indice d'association déterminerait le degré d'implantabilité du vocable.

Lorsque nous calculons $\frac{x}{(y+z)}$ nous pouvons rendre compte du degré d'implantation, voire du degré de résistance du vocable. Cette opération sera l'indice d'association du vocable à la phase d'implantation ou de non implantation du vocable.

Il doit être le paramètre décisif dans la mesure du degré d'implantabilité.

Dans la détermination du degré d'implantabilité, il importe d'observer les indices $\frac{U}{R}$ et $\frac{U}{(R+H)}$.

Pour le vocable *classification dirigée*, le quotient $\frac{U}{R} = 1,16$ (tableau n°1) montre un vocable implanté. Or, l'indice $\frac{U}{(R+H)} = 0,85$ est contradictoire : le vocable ne serait pas implanté, mais en phase d'implantation et il pourrait être en concurrence avec un ou plusieurs substituts.

Dans le cas des vocables *zone d'essai* et *zone d'entraînement*, le quotient $\frac{U}{(R+H)}$ précise le degré d'implantabilité. La phase d'implantation de ces vocables paraît délicate; nous serions tentés de porter un diagnostic de résistance à l'égard des vocables.

Jusqu'à présent, nous avons discuté la notion d'implantabilité d'un vocable à l'intérieur d'un groupe d'individus bien défini. Maintenant, la question est de savoir s'il est possible de déterminer le degré d'implantabilité d'un même vocable en conjugant les données issues de plusieurs groupes?

Considérons quelques vocables dans nos deux groupes. (voir tableau 5 page 20)

A priori, il existe une corrélation entre les deux groupes. Cependant, nous constatons, d'une manière générale, que les vocables « bien implantés » dans le groupe interlocuteurs professionnels ont un degré d'implantation moindre dans le groupe formation/recherche.

Ce constat marque la dominance d'un groupe dans la détermination de l'implantation d'un vocable. En admettant que nous ayons un groupe dit *majeur* dans le niveau d'implantation d'un vocable et un groupe dit *mineur* dans le niveau d'implantation de ce même vocable, pouvons-nous affirmer que le groupe majeur influence le groupe mineur?

Cette hypothèse supposerait une étroite relation entre les deux groupes. Elle impliquerait des interactions langagières véhiculant une fluctuation de détermination du degré d'implantation.

Si à l'inverse, il n'existait pas d'interaction langagière entre les deux groupes, la différence de degré d'implantation d'un même vocable entre les deux groupes mettrait en évidence des caractéristiques intrinsèques à chaque catégorie socio-professionnelle.

Dans ce cas, l'étude des groupes devrait comporter une description des concepts de chaque vocable afin de comprendre et de cerner les préférences de chaque groupe.

Tableau n°3 : Relevé des données concernant les interlocuteurs formation / recherche

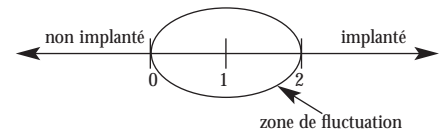
Vocables	Utilisation	Refus	Hésitation	$\frac{U}{R}$	$\frac{U}{(R+H)}$
<i>bande spectrale</i>	227	228	24	0,996	0,901
<i>canal spectral</i>	41	70	47	0,586	0,35
<i>capteur</i>	227	232	23	0,978	0,89
<i>chatolement</i>	5	84	87	0,06	0,029
<i>cible</i>	135	161	64	0,839	0,6
<i>classification dirigée</i>	63	70	25	0,9	0,663
<i>classification non dirigée</i>	216	224	24	0,964	0,871
<i>composition colorée</i>	235	235	24	1	0,907
<i>détecteur</i>	216	226	25	0,956	0,861
<i>déversement radar</i>	12	47	151	0,255	0,061
<i>donnée brute</i>	206	221	30	0,932	0,821
<i>donnée d'apprentissage</i>	109	135	48	0,807	0,596
<i>donnée de terrain</i>	197	220	25	0,895	0,804
<i>donnée géocodée</i>	145	184	51	0,788	0,617
<i>donnée image</i>	113	134	35	0,843	0,669
<i>équidensité colorée</i>	135	175	53	0,771	0,592
<i>fausse couleur</i>	219	230	28	0,952	0,849
<i>filtrage numérique</i>	158	19	49	8,316	2,324
<i>fouillis d'échos</i>	10	86	128	0,116	0,047
<i>hyperfréquence</i>	156	31	35	5,032	2,364
<i>imagerie</i>	184	15	29	12,267	4,182
<i>imagette</i>	165	27	34	6,111	2,705
<i>image chronoséquentielle</i>	24	74	57	0,324	0,183
<i>image d'écran</i>	80	87	26	0,92	0,708
<i>image diachronique</i>	84	127	25	0,661	0,553
<i>image filtrée</i>	203	35	25	5,8	3,383
<i>image prétraitée</i>	224	11	25	20,364	6,222
<i>image radar</i>	225	8	31	28,125	5,769
<i>image satellite</i>	219	2	22	109,5	9,125
<i>indice de végétation</i>	223	11	23	20,273	6,559
<i>infrarouge couleur</i>	234	3	22	78	9,36
<i>ligne de balayage</i>	130	34	68	3,824	1,275
<i>lobe d'antenne</i>	91	37	124	2,459	0,565
<i>mosaïque</i>	212	23	24	9,217	4,511
<i>numérisation</i>	207	6	47	34,5	3,906
<i>orthophotoplan</i>	84	53	77	1,585	0,646
<i>pas d'échantillonnage</i>	177	18	68	9,833	2,058

L'implantation terminologique nécessiterait une étude sur deux plans: l'un consisterait à déterminer un degré d'implantabilité dans chaque groupe, le second établirait les interrelations entre les groupes.

3 Conclusion

La détermination du degré d'implantabilité d'un vocable nécessite le calcul de plusieurs indices d'association du vocable au(x) groupe(s) d'informateurs.

Ces indices échelonnent l'implantabilité d'un vocable dans la zone d'incertitude. De ce fait, la zone d'incertitude devient une zone de fluctuation orientée:



Certains termes révèlent, à l'aide de cette approche, un profil qui comporte déjà quelques composantes d'un profil d'implantation (zone comprise entre 1 et 2); d'autres termes, au contraire, manifestent un profil qui d'emblée laisse présager de grandes difficultés à l'implantation (zone comprise entre 1 et 0).

Des études portant sur ces deux groupes de termes, avec pour objectif de mettre en évidence les traits spécifiques des unités de l'un et l'autre, apporteraient des données linguistiques et extralinguistiques qui pourraient être sollicitées dans les opérations de primo désignation.

*Ch. Langard, J. Rouges-Martinez,
Centre de linguistique et de
dialectologie,
Université de Toulouse-le-Mirail,
France*

Bibliographie

Rouges-Martinez (J.), 1992: *Rapport final du programme de recherches sur l'implantation terminologique dans le domaine de la télédétection aérospatiale*, Paris, DGLF.

Rouges-Martinez (J.) et Fossat (J.-L.), 1994: « Implantation terminologique en télédétection aérospatiale: méthodologie et mesures », dans *Terminologies nouvelles*, n° 12, p. 123-135.

Tableau n° 4 : Relevé des données concernant les interlocuteurs formation / recherche

Vocables	Utilisation	Refus	Hésitation	$\frac{U}{R}$	$\frac{U}{(R+H)}$
<i>pixel</i>	239	1	24	239	9,56
<i>rabattement radar</i>	22	83	158	0,265	0,09
<i>radargraphie</i>	25	76	131	0,329	0,12
<i>radiomètre</i>	229	14	21	16,357	6,54
<i>rayonnement électromagnétique</i>	214	28	20	7,643	4,46
<i>réalité de terrain</i>	126	53	25	2,377	1,62
<i>réflectance</i>	232	4	24	58	8,29
<i>pouvoir de résolution</i>	222	3	28	74	7,16
<i>retrodiffusion</i>	194	27	37	7,185	3,03
<i>scanner</i>	121	30	28	4,033	2,09
<i>scène</i>	218	22	23	9,909	4,84
<i>segment de données</i>	82	130	49	0,631	0,46
<i>signature spectrale</i>	237	3	24	79	8,78
<i>spatiocarte</i>	193	41	30	4,707	2,72
<i>tachèle</i>	14	46	70	0,304	0,12
<i>signature spectrale</i>	237	3	24	79	8,78
<i>spatiocarte</i>	193	41	30	4,707	2,72
<i>tachèle</i>	14	46	70	0,304	0,12
<i>thermographie</i>	199	17	42	11,706	3,37
<i>trace de satellite</i>	191	29	28	6,586	3,35
<i>traitement d'image numérique</i>	216	9	23	24	6,75
<i>vérité-terrain</i>	135	34	25	3,971	2,29

Tableau n° 5

Vocables	Interlocuteurs Professionnels		Formation Recherche	
	$\frac{U}{R}$	$\frac{U}{(R+H)}$	$\frac{U}{R}$	$\frac{U}{(R+H)}$
<i>classification dirigée</i>	1,16	0,853	0,9	0,663
<i>image diachronique</i>	0,845	0,485	0,661	0,553
<i>détecteur</i>	168	21	0,956	0,861
<i>déversement radar</i>	0,48	0,195	0,252	0,061
<i>donnée brute</i>	168	21	0,932	0,821
<i>donnée d'apprentissage</i>	2,098	1,845	0,807	0,596
<i>donnée de terrain</i>	81	16,2	0,895	0,804