



**HAL**  
open science

## Le TAL au service de la qualité rédactionnelle

Reutenauer Coralie, Roy Varlette, Raynal Céline, Dacheux Laurence

► **To cite this version:**

Reutenauer Coralie, Roy Varlette, Raynal Céline, Dacheux Laurence. Le TAL au service de la qualité rédactionnelle. Congrès Lambda Mu 23 “ Innovations et maîtrise des risques pour un avenir durable ” - 23e Congrès de Maîtrise des Risques et de Sûreté de Fonctionnement, Institut pour la Maîtrise des Risques, Oct 2022, Paris Saclay, France. hal-03966514

**HAL Id: hal-03966514**

**<https://hal.science/hal-03966514>**

Submitted on 31 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Le TAL au service de la qualité rédactionnelle

## Using NLP to improve editorial quality

REUTENAUER Coralie  
Plateau Simplification SNCF  
10, rue Camille Moke  
93212 La Plaine Saint-Denis  
[coralie.reutenauer@sncf.fr](mailto:coralie.reutenauer@sncf.fr)

RAYNAL Céline  
Safety Data – OmniContact  
49 rue Ponthieu  
75008 Paris  
[celine.raynal@safety-data.com](mailto:celine.raynal@safety-data.com)

DACHEUX Laurence  
Safety Data – OmniContact  
49 rue Ponthieu  
75008 Paris  
[laurence.dacheux@safety-data.com](mailto:laurence.dacheux@safety-data.com)

ROY VARLETTE Patricia  
SNCF Voyageurs – Direction du  
Matériel  
4, rue André Campra  
93212 La Plaine Saint-Denis  
[patricia.roy@sncf.fr](mailto:patricia.roy@sncf.fr)

**Résumé** — En 2017, SNCF engage une refonte de son système documentaire dans le cadre d'un programme d'excellence sécurité. Un des enjeux est d'améliorer la qualité rédactionnelle pour une documentation technique plus claire. À partir de 2019, SNCF commence un travail avec Safety Data et son logiciel de vérification de la qualité rédactionnelle isChecker. La première étape consiste à tester le logiciel dans sa configuration par défaut. L'étape suivante permet de paramétrer le logiciel de telle sorte qu'il réponde plus précisément aux besoins rédactionnels de SNCF. L'analyse croisée des résultats du logiciel et des guides de rédaction SNCF incite alors les experts SNCF à engager un travail d'optimisation des guides. Cette démarche permet une double consolidation des standards de rédaction et des outils d'aide à leur application.

**Mots-clefs** — *TAL, documentation technique, qualité rédactionnelle, guide de rédaction*

**Abstract** — In 2017, the SNCF (the French national railway company) started the redesign of its information system in the frame of a safety excellence programme. One of the stakes is to improve writing quality to make technical documentation clearer. Since 2019, the SNCF has started a collaboration with Safety Data, taking advantage of its writing quality verification tool isChecker. The first step consists in testing the software with its default configuration. The following step consists in configuring the tool so as to more precisely correspond to the actual constraints of the SNCF writing rules. Cross-checking analysis results from the tool and the SNCF writing guides help SNCF experts in the optimisation of these guides. This approach leads to a twofold consolidation of writing standards and of the tools assisting in the application of these rules.

**Keywords** — *NLP, technical documentation, writing quality, writing guide*

### I. OBJECTIFS

En 2017, SNCF engage une refonte de son système documentaire dans le cadre d'un programme d'excellence sécurité. Un des enjeux est d'améliorer la qualité rédactionnelle pour une documentation opérationnelle plus claire. Dans ce but, elle mène des travaux pour standardiser la rédaction, professionnaliser des rédacteurs et les doter d'outils d'aide à la rédaction. A partir de 2019, elle engage une démarche exploratoire avec la société Safety Data (SD), spécialisée en traitement automatique des langues – notamment dans les traitements linguistiques appliqués à la documentation technique –, pour expérimenter un outil de vérification de la qualité rédactionnelle, isChecker.

Ces expérimentations ont eu pour objectif :

- d'évaluer le potentiel technologique du logiciel : sa capacité à détecter les écarts par rapport aux normes linguistiques de rédaction applicables aux textes de prescription SNCF et la pertinence de ses résultats ;

- d'évaluer le potentiel d'usage du logiciel en tant qu'outil d'aide à la rédaction et aussi d'outil de mesure de la qualité rédactionnelle, et évaluer ainsi l'utilité et l'utilisabilité des résultats produits ;
- d'exploiter l'analyse de corpus réalisée à partir des résultats fournis par le logiciel pour objectiver les habitudes rédactionnelles et renforcer la pertinence des standards de rédaction SNCF : exploiter les indicateurs pour mieux connaître les spécificités du corpus SNCF, mettre ses standards de rédaction à l'épreuve de la réalité des données.

### II. LE CONTEXTE SNCF

Le système de prescription du groupe SNCF est une base documentaire constituée de 90 000 textes de prescription. Cette documentation est produite par 5 000 rédacteurs et ouverte à toute personne du Groupe Public Ferroviaire. Les textes mis à disposition dans le système de prescription sont répartis selon 4 principales natures de textes : les principes, les

règles, les documents d'organisation & méthode et les documents opérationnels. Ils peuvent s'appliquer à un niveau national, régional ou d'établissement, et concernent des destinataires variés, couvrant les différents métiers SNCF ainsi que plusieurs profils (manager, fonctions supports, opérateurs). Le corpus est composé de plus de 50 % de textes de sécurité. La documentation est prescriptive, elle relève de la famille des textes procéduraux, destinés à dire de faire et comment faire et caractérisés par des spécificités linguistiques et discursives [1].

La documentation SNCF est complexe, avec des limites et irritants identifiés tels qu'une rédaction plus littéraire que technique, l'imbrication des informations opérationnelles et de cadrage dans des textes de nature différente, des strates multiples de documentation, une complexité liée à l'organisation et l'évolution de l'organisation, ou encore la présence de texte flou.

En 2015 est lancé un projet de transformation documentaire, réalisé dans le cadre du volet Simplification du programme PRISME d'excellence sécurité. L'ambition est de produire une documentation orientée utilisateur facilitant l'accès aux compétences et savoir-faire, avec des contenus précis et ciblés pour une prise de décision simplifiée lors des opérations d'exploitation et maintenance sur le terrain. Le projet comporte une refonte de la production documentaire, à travers trois axes participant d'une démarche globale de normalisation [6] : la mise en place de standards de rédaction, la professionnalisation des rédacteurs et la mise en œuvre de nouveaux outils pour les rédacteurs. Un des objectifs est d'introduire une nouvelle façon d'écrire, pour simplifier les contenus, limiter les risques d'interprétations différentes et dissocier règles et informations opérationnelles.

Dans ce contexte, un guide SNCF décrivant les standards de rédaction est produit en 2019. Les principes s'appuient largement sur les normes du GIFAS [4], mais le guide n'est pas construit avec des utilisateurs. En parallèle a lieu le lancement d'un programme de formation des rédacteurs. De façon complémentaire, SNCF cherche à doter les rédacteurs et les responsables de la qualité documentaire de solutions logicielles pour gagner en qualité d'écriture et professionnaliser la production documentaire. Elle s'oriente ainsi vers une solution de traitement automatique du langage basée sur une implémentation des standards de rédaction technique, le logiciel isChecker de la société Safety Data, afin d'outiller l'analyse de la qualité rédactionnelle. Pour accompagner et évaluer la démarche, elle mobilise des experts métiers, représentants de différentes activités métiers du groupe SNCF, pour la plupart référents de la simplification documentaire dans le domaine d'activité, et rédacteurs de textes de prescription.

### III. LE LOGICIEL ISCHECKER

Nous présentons les principes du logiciel isChecker développé par Safety Data (groupe OmniContact) avant de décrire précisément son fonctionnement et ses cas d'usage.

#### A. Principe

Le logiciel isChecker développé par Safety Data a été créé dans le but de fournir une aide aux rédacteurs et vérificateurs de documents techniques [3], avec l'objectif de repérer les pratiques rédactionnelles pouvant entraîner un problème de compréhension à la lecture des documents, voire un risque lors de la réalisation des instructions décrites le cas échéant.

Le repérage des problèmes rédactionnels réalisé par isChecker consiste à vérifier que les règles linguistiques prescrites, par des standards internationaux ou des guides de rédaction Client par exemple (nous y revenons plus loin), sont bien appliquées. Dans le cas contraire, isChecker considère qu'il y a « écart » à la norme en vigueur et le signale en fournissant un message d'alerte à l'utilisateur spécifiant la règle enfreinte et le problème rencontré, permettant ainsi à celui-ci de corriger le texte. L'ensemble des écarts est signalé de façon à ce qu'ils puissent être facilement compris et éliminés par le rédacteur ou le vérificateur des documents.

Le logiciel se propose de répondre à un contexte de rédaction technique souvent hétérogène. En effet, bien qu'il existe souvent des guides de rédaction dans les entreprises, guides ayant pour vocation à unifier les pratiques des différents rédacteurs, il est difficile de garantir que les documents répondent à tous les critères prescrits dans la mesure où ce sont des humains qui rédigent, humains qui par définition introduisent de la variabilité et sont susceptibles d'oublier d'appliquer une règle, de l'appliquer de façon non conforme ou tout simplement de la méconnaître – voire de penser qu'elle est superflue. L'un des intérêts d'utiliser un logiciel comme isChecker tient dans la vérification exhaustive et systématique qu'il permet. En effet, tous les documents analysés par isChecker se voient appliquer les mêmes règles, quelle que soit la longueur du document par exemple, permettant ainsi de s'assurer de la cohérence rédactionnelle non seulement d'un même document, mais également de tout le fonds documentaire traité. Un autre intérêt est au niveau de la perception et de l'acceptation par un rédacteur des écarts remontés : le sentiment de critique ou de remise en question de la qualité de son travail sera moindre s'il reçoit des écarts d'application de règles de la part d'un logiciel plutôt que de la part d'un collègue voire de sa hiérarchie, surtout concernant une façon de rédiger. Le logiciel peut servir d'autocontrôle.

Notons que isChecker n'a pas vocation à corriger automatiquement les écarts détectés. En effet, non seulement il est tout simplement impossible, dans la majorité des cas, d'appliquer à un problème une solution qui soit toujours la même, mais au-delà de cet aspect purement technique, il est crucial que l'expert humain conserve l'entière maîtrise des corrections qu'il apporte et de la version finale du document qu'il rédige. Le but est donc d'accompagner le rédacteur ou le vérificateur dans sa mission, sans jamais prendre de décisions à sa place. Le logiciel devient un protagoniste des négociations qui se construisent habituellement entre les différents acteurs autour de la production d'un document : il assiste à la dynamique de construction, sans imposer son point de vue [2]. En effet, la décision finale revient toujours au rédacteur, seul à savoir si un écart doit être corrigé ou non.

#### B. Fonctionnement

isChecker se présente aujourd'hui sous la forme classique d'un environnement client-serveur et fonctionne de la manière suivante : il commence par récupérer le texte du document à analyser, puis applique les règles sélectionnées sur ce texte, *i.e.* vérifie que ces règles sont correctement utilisées. En sortie, il fournit les résultats de son analyse sous forme d'écarts. Chaque écart correspond à une déviation constatée par rapport à une règle donnée. Une même portion de texte, typiquement une phrase, peut donc comporter plusieurs écarts. C'est le cas de la phrase suivante :

*Le conducteur adapte l'utilisation ultérieure du frein en fonction des constatations qu'il aura fait au cours de ce premier serrage.*

Cette phrase contient trois écarts distincts :

- la longueur de la phrase est supérieure au maximum autorisé (22 mots au lieu des 20 prescrits) ;
- le futur antérieur est utilisé alors qu'il est interdit (« aura fait ») ;
- un marqueur de flou qui apporte de l'imprécision à la consigne est présent (« en fonction de »).

isChecker fonctionne sur n'importe quel corpus de documents techniques, et à partir de n'importe quelle norme de rédaction. Cette adaptabilité est possible grâce à l'architecture modulaire de l'outil, et notamment son module de règles. Ce module contient plusieurs ensembles de règles, et peut également intégrer toute règle spécifique à un Client donné. Il intègre par défaut trois « packs » de règles :

- pack STE (*Simplified Technical English*) : règles pour l'anglais issues de la norme *ASD-STE100* de l'ASD ;
- pack GIFAS : règles pour le français issues du *Guide du rédacteur* du GIFAS<sup>1</sup> ;
- pack SD : règles bilingues (anglais/français), conçues et implémentées par l'équipe de linguistes de Safety Data ; elles viennent en complément des règles déjà préconisées par les standards STE et GIFAS et sont issues de l'expertise de l'équipe relativement à la documentation technique.

L'intérêt de ces différents packs est de permettre de cibler l'analyse en fonction des besoins et de la demande du Client. Par exemple, si l'objectif est de rédiger en français conformément au GIFAS, le pack GIFAS peut être sélectionné à l'exclusion des autres. Notons qu'il est également possible, à la demande du Client, de réaliser une configuration spécifique et de ne sélectionner que les règles pertinentes pour le Client, au sein d'un ou plusieurs packs. Le logiciel est donc entièrement configurable du point de vue des règles vérifiées.

isChecker est d'autant plus souple et configurable qu'à ces règles existant par défaut, peuvent s'ajouter des règles spécifiques au Client. S'il existe un guide ou un standard de rédaction utilisé par le Client, il est parfaitement possible d'intégrer ses règles à isChecker, dans la mesure où celles-ci sont vérifiables automatiquement dans l'état actuel des connaissances et des ressources. Quelle que soit l'origine d'une règle, sa vérification consiste donc dans tous les cas à comparer une prescription de la norme à ce qui a été produit et qui se trouve dans le document en cours d'analyse.

Cette structure permet une grande souplesse dans les traitements appliqués aux données d'une entreprise. Ainsi, dans l'hypothèse où la norme serait modifiée, ou si les documents techniques d'un nouveau service devaient être

traités, le coût de mise en place des traitements serait très réduit.

### C. Cas d'usage

Si le logiciel isChecker réalise toujours le même traitement décrit ci-dessus, à savoir analyser le contenu linguistique de documents techniques afin de s'assurer de leur conformité vis-à-vis de règles de rédaction en vigueur, l'usage qui peut en être fait peut varier. On peut au moins identifier quatre scénarios d'utilisation.

- isChecker permet de faire un audit des pratiques rédactionnelles sur un échantillon de documents, afin de mettre en évidence : les bonnes pratiques, les problématiques rédactionnelles connues, les nouveaux points de vigilance à prendre en compte, les règles existantes non pertinentes dans le contexte client. Le logiciel fournit des indicateurs pertinents sur la qualité rédactionnelle au regard des règles implémentées. Ainsi, documents et résultats à l'appui, isChecker permet d'identifier des règles de rédaction pertinentes pour créer ou optimiser un guide de rédaction existant ; il permet plus globalement de challenger une organisation pour uniformiser ses écrits.
- isChecker apporte une assistance rédactionnelle voire un auto-contrôle en alertant sur le non-respect d'une règle prescrite, en rappelant la règle le cas échéant, et en aidant le rédacteur à trouver une correction adéquate en renvoyant un message ciblé sur le problème rencontré. Dans une utilisation régulière par un rédacteur, isChecker renvoyant de façon systématique et identique les écarts, l'utilisateur se familiarise au fil du temps avec les règles et intègre des « réflexes rédactionnels ». Ainsi, après avoir vu un certain nombre d'écarts relatifs aux marqueurs de flou, les expressions ou mots déconseillés lui seront devenus familiers et il en emploiera spontanément moins ; la règle selon laquelle les expressions ou termes qui apportent de l'imprécision sont à éviter sera ainsi de mieux en mieux intégrée.
- isChecker est un support pour le relecteur dans sa tâche de vérification des documents techniques, lui permettant ainsi de cibler les améliorations rédactionnelles à faire réaliser.
- Enfin, utilisé à des fins d'analyse de haut niveau, isChecker peut aider les personnes en charge de la gestion de la documentation et de la rédaction à décider des actions à mettre en œuvre au vu des résultats isChecker : action de formation, revue des process...

Nous présentons dans ce qui suit les différentes étapes réalisées avec SNCF dans le cadre du projet d'amélioration de la qualité rédactionnelle avec isChecker, les cas d'usage pouvant se compléter, nous allons le voir.

<sup>1</sup> Ces deux standards sont proposés par les industriels de l'aéronautique et du spatial : AeroSpace and Defence Industries (ASD) pour l'anglais, et le Groupement des industries françaises aéronautiques et spatiales (GIFAS) pour le français. Néanmoins, la grande majorité des règles de

rédaction proposées valent quel que soit le domaine d'activité. Même le glossaire des mots interdits peut être pertinent, au moins en partie, dans certains secteurs tels que le transport ferroviaire

#### IV. LES ETAPES DU PROJET : METHODE

Le projet que nous présentons ici a débuté mi-2019 et est réalisé depuis son commencement dans une démarche itérative : les allers-retours sont au cœur de la démarche, non seulement entre experts métiers SNCF et SD, mais aussi et surtout entre les données et les résultats obtenus avec isChecker d'une part, et les analyses, propositions et décisions des experts d'autre part. Cette navigation constante entre données et experts a été mise en place dès la première étape du projet qui a consisté en l'expérimentation de isChecker sans aucun paramétrage spécifique sur un échantillon de documents restreint (section A) et s'est intensifiée avec les deux étapes suivantes, à savoir l'adaptation de isChecker au contexte métier SNCF (section B) et l'exploitation de l'analyse des données produite par isChecker pour optimiser le guide de rédaction SNCF (section C).

##### A. Expérimentation de isChecker sans paramétrage

La première étape du projet, qui s'est déroulée entre mi-juillet et début septembre 2019, a consisté à tester isChecker dans sa version « par défaut », sans aucun paramétrage particulier, sur un échantillon de documents. Ce corpus comprenait un ensemble de 11 extraits<sup>2</sup> de documents prescriptifs de SNCF issus de différents métiers et rédigés par des experts. Certains de ces experts ont été intégrés à un « panel d'experts métiers » composé de rédacteurs techniques, de rédacteurs métiers et de référents documentaires impliqués dans le programme de simplification documentaire, et chargés d'évaluer d'un point de vue métier la qualité des résultats du logiciel. La démarche étant centrée sur les utilisateurs, l'implication des rédacteurs était primordiale.

Le paramétrage par défaut de isChecker au moment de cette expérimentation consistait en la vérification automatique de 21 règles, 10 règles issues du standard GIFAS et 11 règles Safety Data, recouvrant la détection de problèmes de :

- complexité, par exemple : la phrase ne doit pas contenir plus d'un terme négatif (SD) ;
- formulation, par exemple : la voix passive n'est pas conforme au GIFAS ;
- ou lexicale, par exemple : certains termes sont interdits par le GIFAS.

Une fois les documents de l'échantillon traités par le logiciel, les résultats ont été mis à disposition directement dans les documents Word sous forme de commentaires, et listés exhaustivement dans un fichier Excel, ce dernier contenant également les résultats quantitatifs afin de voir rapidement la répartition des écarts par famille, de manière globale et par document.

Lors de cette première étape, les résultats obtenus, qualitatifs et quantitatifs, ont fait l'objet d'une analyse par les experts linguistes SD et ont été présentés aux experts métiers SNCF. Cette restitution des résultats et de leur analyse a été réalisée lors d'une réunion à laquelle assistait l'ensemble des experts engagés dans le projet, à savoir notamment les rédacteurs des documents analysés. Cette restitution a permis, dès cette première confrontation des experts avec les résultats produits par isChecker, de mettre en évidence des bonnes pratiques tout comme des habitudes rédactionnelles à éviter, et également la disparité entre documents, et donc parfois

entre pratiques rédactionnelles de chacun. Cela a également confirmé que la configuration de isChecker devait impérativement être paramétrée pour répondre aux besoins et particularités du contexte SNCF. Nous détaillons les résultats dans la section V.A.

##### B. Adaptation de isChecker au contexte métier

Suite à la première étape et étant donné les résultats obtenus, il a été décidé de poursuivre le projet en se focalisant sur les particularités SNCF. Ainsi la deuxième phase, qui s'est déroulée au dernier trimestre 2020, a consisté à intégrer isChecker dans le contexte SNCF selon une double démarche que nous détaillons ci-après. Soulignons dès à présent que cette phase du projet a été menée en réalisant six ateliers réunissant systématiquement experts SNCF et experts SD, dans la démarche des allers-retours présentée précédemment.

###### 1) Types de document et règles applicables spécifiques à SNCF

L'objectif étant de préparer une configuration de isChecker la plus adaptée possible au contexte SNCF, la démarche a d'abord consisté à analyser ce contexte. En effet, il était crucial de comprendre l'organisation de la documentation au sein de l'entreprise, les différents types de documents produits et utilisés, leurs spécificités respectives, leurs finalités, leurs publics, etc. Cet état des lieux du fonds documentaire SNCF avait pour objectif de clarifier la typologie des documents afin notamment de pouvoir leur appliquer des règles de rédaction précises, potentiellement distinctes, et ainsi de prévoir une configuration de isChecker adaptée à chacun.

Pour ce faire, un ensemble représentatif de documents a été sélectionné par les experts SNCF pour permettre encore une fois une démarche s'appuyant sur les données réelles. Le corpus était constitué de 27 documents, initialement répartis selon les quatre natures de document déjà utilisées comme métadonnée descriptive du système de prescription SNCF, avec au moins 5 documents par classe. Par ailleurs, un soin particulier a été porté au choix de ces documents afin qu'ils témoignent d'une variété des métiers représentés dans l'entreprise avec un souci d'équilibre entre SA Réseau et SA Voyageurs par exemple, ou encore entre les profils des utilisateurs entre des agents de la ligne managériale *versus* des opérateurs de terrain. L'enjeu dans la constitution du corpus était d'avoir une bonne représentativité de la variété des documents tout en conservant un volume restreint afin de permettre une analyse qualitative fine, mais suffisante pour obtenir des résultats quantitatifs significatifs.

À l'issue de la phase d'analyse des documents du point de vue organisationnel, les experts SNCF ont validé une structuration des documents en deux types de document selon lesquels adapter le paramétrage de l'outil : les documents opérationnels et les documents de cadrage (nous y revenons dans la section V.B dédiée aux résultats).

Une fois cette étape réalisée, le travail a consisté à analyser les guides de rédaction internes SNCF afin d'avoir une idée précise de leur contenu relativement à isChecker. Pour chaque règle ou recommandation explicitée dans ces guides, les experts SD ont précisé si elle était déjà vérifiée par une ou plusieurs règles isChecker existantes, et le cas échéant laquelle ou lesquelles. On observe plusieurs cas de figure :

<sup>2</sup> Soit un total d'environ 200 pages de fichiers Word.

- à une règle SNCF correspond une règle isChecker : « Le nombre de phrases par paragraphe ne doit pas être supérieur à 6 » ;
- à une règle SNCF correspondent plusieurs règles isChecker : la règle SNCF « Une phrase doit être : précise, correcte, courte, concise », de par son périmètre très large, est couverte par plusieurs règles disponibles dans isChecker aussi bien SD que GIFAS, et notamment les règles GIFAS interdisant les instructions coordonnées, la présence du participe présent (qui introduit généralement une demande d'action de manière déguisée), la voix passive (qui complexifie la structure de la phrase) ;
- plusieurs règles SNCF sont couvertes par une règle isChecker : l'une des règles isChecker permet de repérer les structures syntaxiques complexes, or cette règle va permettre à la fois de répondre – en partie au moins – à la règle SNCF citée au point précédent, mais également à celle demandant d'« Ordonner la phrase selon l'ordre logique des événements décrits » par exemple.

Dans certains cas, c'est seulement une partie de la règle SNCF qui trouve sa correspondance dans une règle isChecker. Ce cas de figure peut recouvrir plusieurs situations :

- Soit la partie de la règle SNCF non vérifiée automatiquement jusqu'alors peut l'être et il s'agira de l'implémenter,
- soit elle ne peut pas être vérifiée et la question se pose de la règle initiale ; en effet, ce cas de figure met parfois en évidence une règle complexe voire floue qu'il est bon de clarifier et parfois de subdiviser pour faciliter sa compréhension et son application.

Dans le cas où une règle d'un guide de rédaction SNCF n'était pas présente dans isChecker, il convenait d'acter de la possibilité ou de l'impossibilité de vérifier automatiquement la règle en question. Ne pas pouvoir vérifier automatiquement une règle préconisée par un guide tient soit au fait qu'elle n'est pas linguistique (on pense par exemple à des préconisations techniques), soit au fait qu'elle n'est pas modélisable, *i.e.* qu'il n'est pas possible de créer un algorithme utilisant les informations linguistiques disponibles pour la vérifier. C'est par exemple le cas des recommandations discursives, tenant à l'organisation du discours, de l'information ; par exemple :

#### Organisation du paragraphe :

- + Phrase clé : c'est la phrase la plus importante. Elle indique le sujet traité ;
- + Phrases de développement : ces phrases développent le sujet. Elles fournissent les informations, détails, et aspects particuliers.

Fig. 1. Règle discursive SNCF

À l'issue de cet état des lieux, les règles présentes dans les guides SNCF et absentes de la configuration par défaut de isChecker, mais parfaitement modélisables, ont été identifiées en une liste de « règles SNCF à créer » le moment venu.

Dans une démarche complémentaire, « à rebours », les règles isChecker ont toutes été confrontées aux guides de rédaction SNCF afin de pointer celles ne figurant pas dans les référentiels de l'entreprise et de décider de leur pertinence

avec les experts métiers pour les conserver dans la configuration finale – et les ajouter au guide de rédaction, nous y revenons dans la section suivante – ou non.

Enfin, notons que lors du passage en revue des règles SNCF et isChecker, la question de leur adéquation aux différents types de document établis s'est posée. En effet, pour chaque règle à faire vérifier à isChecker, il s'est agi de décider si elle était pertinente pour tous les documents et si oui de la même façon ou non, ou si elle n'était valable que pour l'un des types de document définis, document opérationnel ou document de cadrage.

#### 2) Définition des règles de rédaction à conserver dans isChecker

Une fois le corpus constitué, les documents ont d'abord été analysés avec isChecker dans sa configuration par défaut. En effet, il s'agissait dans un premier temps de confirmer les résultats obtenus et les tendances observées lors de la première étape du projet (sur seulement 200 pages)<sup>3</sup>.

En s'appuyant sur ces résultats et sur le travail fait à partir des guides de rédaction internes SNCF, les ateliers regroupant les experts métiers et les experts linguistes ont permis de préciser le sort à faire à chacune des règles présentes dans la configuration par défaut :

- inhibition de la règle isChecker en raison de sa non-pertinence pour le contexte SNCF : c'est le cas de la règle GIFAS qui interdit l'utilisation du participe présent ;
- ajustement de la règle isChecker pour s'adapter au contexte SNCF : c'est le cas de la règle sur les temps et modes interdits, à la liste de ceux interdits par le GIFAS, on a ajouté le passé composé ;
- conservation de la règle isChecker : on pense à la règle limitant le nombre de phrases par paragraphe, à celles interdisant la voix passive ou la coordination d'instructions.

Il est important de noter que les ateliers ont permis aux experts métiers et aux experts linguistes de discuter des résultats obtenus avec isChecker, et de passer du temps à débattre du bien-fondé (ou pas) de conserver telle ou telle règle. En effet, il n'est pas toujours évident de trouver tout de suite un consensus sur l'application d'une règle. C'est pourquoi, les ateliers, préparés et construits en s'appuyant sur les données analysées et les résultats obtenus, ont été très précieux : du temps était laissé aux échanges, permettant ainsi de prendre en compte les différents points de vue, usages et contextes des experts présents avant de prendre une décision. L'enjeu était de définir les règles adaptées aux pratiques et usages SNCF, sans tout remettre en question mais en suivant une ligne claire d'amélioration de la rédaction de la documentation technique dans la dynamique de professionnalisation de l'activité. Les arbitrages et choix de paramétrage ont été faits pour chaque règle pour chaque type de document (document opérationnel et document de cadrage).

À l'issue de cette étape de précision du périmètre isChecker, des ajustements ont été réalisés dans la

<sup>3</sup> Entre ces deux étapes, une règle Safety Data a été ajoutée à la configuration par défaut de isChecker, portant à 22 le nombre total de règles vérifiées.

configuration du logiciel (sur les 22 règles de la configuration utilisée par défaut, 7 ont été inhibées et 6 ajustées) et les 27 documents réanalysés afin de vérifier la pertinence des choix faits au cours des ateliers. Ces nouveaux résultats ont été comparés par les experts SD avec les résultats obtenus initialement (avec isChecker dans sa configuration par défaut), puis présentés aux experts SNCF afin de procéder à la validation finale de la configuration isChecker. Quelques ajustements supplémentaires ont ainsi pu être décidés. Les résultats sont décrits précisément dans la section V.B.

### C. Exploitation de l'analyse de données produite par isChecker pour optimiser le guide de rédaction SNCF

L'analyse outillée et l'évaluation des résultats par des experts métiers et des linguistes ont fait apparaître la nécessité d'optimiser le guide de rédaction. L'objectif était d'augmenter la précision et la couverture des règles de rédaction et plus précisément :

- d'enrichir le guide existant de nouvelles règles en intégrant les règles discutées et validées lors des étapes précédentes ; par exemple la règle SNCF « Une phrase doit être : précise, correcte, courte, concise » devait être scindée en plusieurs règles et énoncée de façon plus précise, afin d'éviter une interprétation subjective et donc créer de la variabilité entre rédacteurs ;
- d'organiser les règles selon chaque type de document : document opérationnel/document de cadrage ; le guide devait respecter les types de document SNCF définis en amont ;
- d'illustrer les règles avec des exemples et des reformulations issus des analyses en corpus ;
- de définir un modèle de description commun à toutes les règles ;
- de définir une mise en forme qui garantisse l'utilisabilité du guide.

D'un point de vue méthodologique, le guide est issu d'un travail collaboratif entre un expert linguiste SD et trois experts métiers SNCF. La construction du guide s'est déroulée sur une période de 3 mois, de septembre à novembre 2021. Il nous a semblé important de ne pas trop espacer dans le temps les ateliers pour maintenir une dynamique entre les différentes phases du projet que nous détaillons à présent : conception, rédaction, relecture, concertation et validation.

La phase de conception a consisté à :

- proposer un plan : comment organiser les règles et la navigation dans le document ;
- définir les éléments constitutifs de chaque règle : énoncé, description, objectif, applicabilité, exemples et reformulations associées ;
- définir le modèle de présentation de chaque règle.

La rédaction du guide s'est inscrite dans les étapes suivantes :

- la rédaction de l'introduction pour définir le périmètre du guide, présenter les deux types de document (opérationnel/cadrage) et donner des précisions terminologiques ;

- La rédaction de l'énoncé et de la description de chaque règle ;
- l'extraction d'exemples pertinents à partir des analyses de corpus faites par isChecker. Les exemples issus de corpus nécessitent des ajustements du texte de manière à cibler la règle à illustrer. Donner un exemple de temps verbal non approuvé dans une instruction de plus de 20 mots vient contredire la règle sur le nombre de mots maximum par phrase, et est donc contre-productif.

La phase de relecture du guide prépare les échanges et la concertation. La rédaction du guide a donné lieu à 5 versions. Chaque version était relue par les experts métiers SNCF et faisait l'objet de commentaires, suggestions et ajustements à discuter et valider au cours de l'atelier suivant. Cette phase de travail est cruciale pour optimiser le temps de concertation dans les ateliers.

La concertation, étape essentielle de mise en commun et d'échanges, s'est déroulée de la façon suivante :

- envoi de la version  $n$  du guide en amont de l'atelier ;
- retours des experts métiers, toujours en amont de l'atelier ;
- confrontation de ces retours et validation en atelier ;
- à l'issue de l'atelier, préparation de la version  $n+1$  incrémentée de nouvelles règles et envoi de cette version (nouveau cycle).

Chaque atelier était espacé de 15 jours environ. Le premier atelier a permis de valider le plan et l'introduction du guide et de stabiliser un patron commun de description des règles. Les ateliers suivants ont consisté à mettre en commun des commentaires issus des relectures et de confronter les points de vue linguistiques et métier.

La validation s'est déroulée de façon progressive jusqu'à l'étape de clôture. Au cours des cinq ateliers, l'expert linguiste et les trois experts métiers ont validé l'intégralité du guide. Une réunion de clôture a permis de faire le point sur le travail réalisé.

Enfin, la valorisation s'est effectuée en trois étapes. En interne, les responsables de la conduite du changement ont repris les travaux réalisés par les experts SNCF et l'expert linguiste. Ils ont ainsi mis le guide produit en conformité avec la charte graphique SNCF. Ils ont également intégré le guide linguistique au guide complet comprenant les parties sur les illustrations, etc.

## V. RESULTATS

La section IV a permis de présenter en détail la façon dont le travail a été mené au cours du projet d'amélioration de la qualité rédactionnelle avec isChecker ; nous détaillons à présent les résultats obtenus à chacune des trois étapes du projet.

### A. Test de isChecker sans paramétrage

Pour mémoire, cette première étape a consisté à analyser environ 200 pages (11 extraits de documents) avec isChecker dans sa configuration par défaut regroupant 21 règles traitant de problèmes de complexité, de formulation et de lexiche.

À l'issue du traitement, on observe en moyenne 26 écarts par page, tous les documents confondus. Il est intéressant de noter que des variations existent d'un document à l'autre ; ainsi, deux documents ont une moyenne de 18 écarts par page, alors que quatre autres ont une moyenne de plus de 35 écarts par page, le maximum observé étant 38 écarts par page en moyenne pour un document. Au-delà de son aspect purement comptable, cette variation met en exergue la variabilité entre rédacteurs. En effet, comme évoqué précédemment, bien que des guides de rédaction soient présents dans l'entreprise, ils n'excluent pas les différences de pratiques entre rédacteurs.

Notons que le grand nombre d'écarts est parfois dû au fait qu'une même phrase concentre plusieurs écarts (parfois beaucoup). Considérons ainsi la phrase suivante :

*Dans le cas exceptionnel où le remplacement de l'engin avarié ne peut être réalisé, et sauf si l'exploitant ferroviaire ou le présent document prescrivent des mesures plus restrictives, sont autorisés :*

L'analyse automatique faite par isChecker produit 7 écarts relatifs à la présence d'une modalité, de termes déconseillés (par le GIFAS ou par SD qu'ils soient flous ou complexes), de coordination verbale, d'un nombre de mots trop important.

D'un point de vue qualitatif, on remarque que les trois familles d'écarts (complexité, formulation, lexique) sont presque également représentées avec respectivement 37 %, 28 % et 35 % de l'ensemble des écarts. On observe toutefois que cette répartition varie d'un document à l'autre, certains d'entre eux montrant majoritairement des écarts relatifs à la formulation (documents 1, 3, 9).

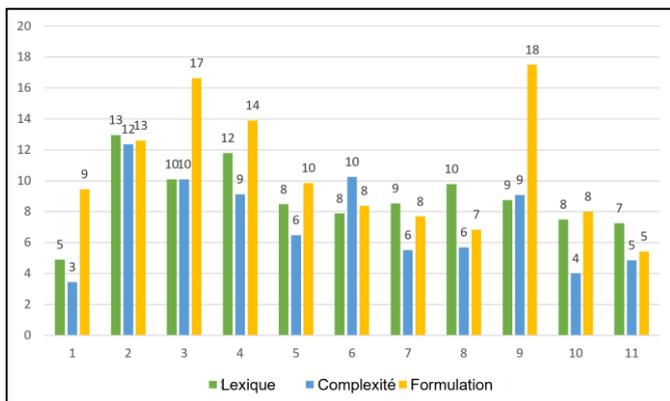


Fig. 2. Répartition des familles d'écarts par document

En analysant les résultats plus précisément, au niveau des règles elles-mêmes, on observe que certaines génèrent plus d'écarts que d'autres, à savoir celles relatives :

- aux termes interdits par le GIFAS ;
- aux phrases trop longues ;
- aux formulations avec un temps ou un mode inadapté ;
- aux marqueurs d'ambiguïté ;
- à la trop forte présence de verbes dans la phrase.

Ces écarts sont illustrés par la phrase suivante qui contient des termes interdits par le GIFAS (« lorsque »), trop de mots (44 au lieu des 20 prescrits), un marqueur d'ambiguïté (« dans certains cas ») et un nombre de verbes élevé (« peut conduire », « allonger », « résulte ») :

*La recherche d'une signalisation claire ou d'une amélioration de la visibilité des signaux peut donc conduire, dans certains cas, à allonger la zone de ralentissement ou d'annonce lorsqu'il n'en résulte pas de dégradation significative du débit et des performances.*

À l'inverse, on ne constate aucun écart sur trois des règles vérifiées par isChecker. Pour deux d'entre elles, on peut sans aucun doute interpréter cela comme une intégration de ces règles dans les bonnes pratiques des rédacteurs, à savoir : ne pas faire des paragraphes comprenant plus de 6 phrases (règle SNCF en adéquation avec une règle GIFAS), et ne pas donner d'instruction entre parenthèses (règle SD). En revanche pour la troisième, plusieurs interprétations sont possibles. En effet, cette règle issue du standard GIFAS précise que les notes ne doivent pas contenir d'instruction. Or, si cette notion de « note » est très claire et définie dans les documents relatifs à l'aéronautique et au spatial<sup>4</sup>, cela ne se retrouve pas tel quel dans l'échantillon de documents SNCF analysé. Par conséquent, l'absence d'écarts signale ici qu'il est nécessaire d'investiguer davantage le sujet plutôt que de considérer une bonne application de la règle en vigueur sur les notes.

Enfin, remarquons que du point de vue du temps de traitement, la version de isChecker utilisée à l'époque permettait de traiter une moyenne de 9 pages par minute, soit environ 20 minutes de traitement pour l'échantillon traité à cette étape. L'architecture ayant été revue et améliorée depuis 2019, l'analyse par isChecker de ce même échantillon d'environ 200 pages ne requiert plus aujourd'hui que 55 secondes.

La présentation aux experts métiers SNCF de ces résultats obtenus avec isChecker sans aucun paramétrage a permis de mettre différentes conclusions en évidence. La première concerne l'efficacité de l'outil tant dans le contenu des écarts, cohérents avec la formation interne SNCF aux standards de rédaction, que dans les formats des sorties produites, le tableau Excel récapitulatif et les documents Word commentés, pertinents, clairs et adaptés aux différents niveaux d'analyse faite par les experts. Ce qui amène au deuxième point relevé par les experts : l'usage multiple qui peut être fait du logiciel. En effet, isChecker peut aussi bien servir au diagnostic pour avoir une vision d'ensemble de la qualité rédactionnelle d'un fonds documentaire, qu'à fournir une aide à la rédaction en signalant les écarts sans pour autant les corriger, ce qui laisse le rédacteur libre de trouver la reformulation qui correspond le mieux au contexte. Et enfin, il apparaît clairement que isChecker est une aide à la formation : formation par la pratique en complémentarité de la formation aux guides existants plus standard. Toutefois, et c'est la quatrième et dernière remarque à ce stade, il est indispensable de paramétrer isChecker car la configuration par défaut n'est pas toujours adaptée au contexte en jeu et le logiciel risque par conséquent, en signalant des écarts qui n'en sont pas au regard des guides de rédaction SNCF, d'être perçu comme non

<sup>4</sup> La « Note » ou le « Nota » contient des informations complémentaires non essentielles à la procédure.

pertinent et de décrédibiliser le logiciel voire même de démotiver les rédacteurs pour l'utiliser.

### B. Test avec paramétrage de isChecker adapté au contexte SNCF

Pour rappel, la deuxième étape du projet a été organisée en trois phases. Dans la première phase, à partir d'un échantillon de 27 documents représentatifs des documents existant dans l'entreprise, il s'est agi de définir des types de document pertinents. C'est ainsi qu'au vu de leur usage (à quoi servent-ils ?), de leur public (à qui sont-ils adressés ?), de leur contenu (que contiennent-ils ?) et de leur gestion (font-ils l'objet de déclinaison ?), les documents ont pu être regroupés en deux types :

- les documents opérationnels destinés à des opérateurs devant « lire pour faire », qui contiennent majoritairement des instructions mais peuvent également donner des informations non prescriptives (descriptions, définitions, etc.) ;
- les documents de cadrage destinés au management ou à la main d'œuvre opérationnelle, qui contiennent des exigences pouvant être complétées par d'autres informations plus descriptives et peuvent faire l'objet de déclinaisons.

Ainsi, chacun des deux types de document peut contenir deux types d'information différents. Par conséquent, il est nécessaire de se poser la question de l'application de règles de rédaction pour chacun de ces quatre types d'information : les instructions et les informations/descriptions dans les documents opérationnels, et les exigences et informations/descriptions dans les documents de cadrage.

C'est en cela qu'a consisté la deuxième phase de cette étape du projet : les 51 règles identifiées dans les guides SNCF ont été globalement considérées sous l'angle d'une vérification automatique et en particulier mises en vis-à-vis des 22 règles disponibles par défaut dans isChecker. Pour chacune d'entre elles, les experts se sont posé la question de leur pertinence pour chacun des quatre types d'information présents dans les documents SNCF, et acté leur validité et leur périmètre, des adaptations étant parfois requises selon le type de document ou d'information considéré.

Sur les 51 règles SNCF identifiées, 28 règles ont été précisément revues et une décision a été prise :

- 10 règles correspondent à des règles isChecker : 6 règles sont des règles par défaut, 2 règles ont été ajustées au cours du projet et 2 règles nécessitaient un ajustement complémentaire ;
- 9 règles sont modélisables et à créer pour être intégrées dans les vérifications de isChecker ;
- 9 règles ont été abandonnées et ne requièrent donc pas de création dans isChecker.

Les 23 autres règles nécessitent des précisions.

- 3 de ces règles peuvent être couvertes par des règles déjà présentes dans isChecker *modulo* quelques ajustements, mais ces ajustements nécessitent une investigation complémentaire plus fine du corpus. C'est le cas de la règle SNCF qui préconise de ne pas utiliser d'adverbes : en effet, cela semble radical et

mérite une analyse des données pour s'assurer de sa validité ou assouplir un peu la recommandation.

- 9 règles n'étant pas vérifiables automatiquement (ex : « Ne pas substituer le logigramme au texte »), elles n'ont pas été intégralement revues et il reste nécessaire de valider leur pertinence pour chaque type de document et d'information (au-delà de leur vérification par isChecker).
- Et enfin, 11 règles sont modélisables et peuvent être créées et intégrées dans isChecker ; leur pertinence et leur périmètre d'application devront alors être confirmés.

Ainsi, au total, ce sont 33 règles des guides SNCF, sur les 51 identifiées, qui sont déjà ou peuvent être modélisées et vérifiées automatiquement par isChecker.

En changeant de perspective et en considérant les 22 règles de la configuration par défaut de isChecker, la configuration SNCF est la suivante :

- 16 règles sont conservées, dont 7 dans leur configuration par défaut. Pour les 9 autres, soit des ajustements ont déjà été faits au cours du projet pour répondre au contexte SNCF, soit une analyse plus fine est requise pour statuer sur les ajustements à réaliser précisément.
- 4 règles restent en attente d'une étude complémentaire car aucun consensus n'a permis de prendre de décision définitive sur leur pertinence et leur périmètre d'application, et seule une analyse fine des phénomènes linguistiques concernés peut aider les experts SNCF à trancher. C'est par exemple le cas des participes présents dont la présence est autorisée par certains lorsqu'elle est déconseillée par d'autres. Vérifier l'usage de ces participes présents en corpus aidera à prendre une décision éclairée (par exemple préciser les contextes où il faut en interdire l'usage et ceux où il est autorisé).
- 2 règles sont exclues, non pertinentes pour le contexte SNCF : celle du lexique autorisé/interdit par le GIFAS et la longueur des listes à puce (SD).

Rappelons que toutes ces règles ont été discutées et leurs application et périmètre validés en fonction du type de document et du type d'information concernés. Ainsi, certaines règles ne sont pas toujours appliquées de la même façon ; c'est par exemple le cas de la règle spécifiant le nombre maximal de mots autorisé par phrase. En effet, celui-ci varie selon le type d'information : les instructions des documents opérationnels ne doivent pas dépasser 20 mots, les descriptions de ces mêmes documents ne doivent pas dépasser 25 mots, tandis que les phrases contenues dans les documents de cadrage, quelles qu'elles soient, peuvent contenir jusqu'à 30 mots.

À l'issue de cette deuxième étape du projet, un bilan a été fait par les experts SNCF sur le potentiel d'usage de isChecker après paramétrage. Si l'efficacité du logiciel et son utilité ne font aucun doute, il est crucial d'accompagner les utilisateurs dans son intégration dans le processus rédactionnel. Il semble d'ailleurs plus pertinent qu'un tel outil soit mis au service des rédacteurs professionnels ou aux qualitatifs de la documentation plutôt qu'à des rédacteurs occasionnels. isChecker est à considérer comme un appui, un support pour

les rédacteurs et en ce sens, il est fondamental de bien définir son périmètre d'action. En effet, cette étape de configuration adaptée au contexte SNCF confirme l'idée qu'une utilisation sans paramétrage n'est pas envisageable étant donné le « bruit » qu'elle est susceptible de produire. Enfin, ces ajustements doivent également s'accompagner d'une réflexion sur la cible que l'on cherche à atteindre ; en effet, il est nécessaire de bien délimiter le périmètre de vérification, quitte à l'élargir petit à petit, pour que isChecker soit adopté par les rédacteurs et éviter le possible découragement qu'un trop grand nombre de règles à respecter, et potentiellement d'écarts, pourrait entraîner.

Parallèlement, cette phase d'allers-retours entre les règles SNCF, les règles isChecker et les données analysées ont également mis en évidence le besoin de clarifier certaines règles des guides, voire pour les experts de se mettre d'accord sur une pratique unique alors même que différents usages avaient jusqu'alors cours. En effet, rappelons qu'au sein de SNCF, cohabitent des contextes et des pratiques métiers différents, et que les usages peuvent par conséquent varier d'un métier à l'autre, l'arbitrage pour arriver à un modèle commun se révélant alors parfois délicat à réaliser. Le recours aux documents eux-mêmes et à la réalité de leur contenu rédactionnel, éclairé à la lumière de l'analyse isChecker, a permis d'éclaircir les débats qui pouvaient avoir lieu entre des experts métiers de différentes spécialités n'ayant pas tous la même façon d'appréhender les choses.

L'exercice mené a ainsi mis en évidence le besoin d'avoir des règles très explicites, exemples à l'appui, afin de rendre les préconisations plus concrètes et permettre une meilleure compréhension et une assimilation facilitée des règles. C'est ainsi que la troisième étape a vu le jour : l'optimisation du guide de rédaction SNCF.

### C. Mise en qualité du guide de rédaction

À l'issue des ateliers de construction et de rédaction du nouveau guide de rédaction SNCF, trois documents ont été produits :

- un guide complet avec 28 règles linguistiques ;
- un guide mémo pour les documents opérationnels ;
- un guide mémo pour les documents de cadrage.

Chaque règle linguistique du guide complet respecte le même patron qui contient le numéro et l'énoncé de la règle, sa description, son applicabilité, ses objectifs, les risques associés en cas de non-application et deux exemples : une phrase ne respectant pas la règle et une phrase conforme à ce qui est demandé par la règle. La figure ci-dessous illustre la règle relative au nombre de mots autorisés dans une instruction (document opérationnel).

R03	Utiliser 20 mots maximum par phrase.	
Description	Chaque mot de la phrase compte, y compris les mots entre parenthèses.	
Applicabilité	Document opérationnel : <input checked="" type="checkbox"/> Instruction <input type="checkbox"/> Description	Document de cadrage : <input type="checkbox"/> Exigence <input type="checkbox"/> Description
Objectifs	Optimiser la compréhension de la phrase. Limiter le nombre d'informations dans une phrase. Éviter l'omission d'instructions.	
Risques	Complexité, omission	
Exemples	<p>✗ Le remplacement à l'identique d'un fusible est effectué après avoir vérifié l'absence de tension de chaque côté du fusible et suivant un mode opératoire défini au document technique. (30 mots)</p>	<p>✓ Vérifier l'absence de tension de chaque côté du fusible. (11 mots) Remplacer le fusible selon le mode opératoire du document technique. (10 mots)</p>

Fig. 3. Exemple de règle du Guide de rédaction

Il est apparu nécessaire d'avoir un guide spécifique pour chaque type de document, document opérationnel et document de cadrage, appelé « Guide mémo ». Son format est restreint à l'énoncé de la règle, son applicabilité et ses exemples associés (Fig. 4).

R08	Utiliser la voix active.	
Applicabilité	Document de cadrage : <input checked="" type="checkbox"/> Exigence <input checked="" type="checkbox"/> Description	
Exemples	<p>✗ Les temps de stationnement sont attribués par les horaristes en tenant compte des contraintes de l'Exploitation.</p>	<p>✓ Les horaristes attribuent les temps de stationnement en tenant compte des contraintes de l'Exploitation.</p>
	<p>✗ La traçabilité des missions et tâches correspondantes doit être assurée.</p>	<p>✓ La ligne managériale doit assurer la traçabilité des missions et des tâches correspondantes.</p>
	<p>✗ Cette étude est complétée par une enquête terrain organisée par le mainteneur en collaboration avec l'expert traction.</p>	<p>✓ Le mainteneur organise une enquête terrain en collaboration avec l'expert traction pour compléter l'étude.</p>

Fig. 4. Exemple de règle du Guide Mémo pour les documents de cadrage

Le travail de validation en ateliers a donné lieu à des débats constructifs faisant émerger des points de vue métiers différents et des convergences.

- La validation du plan, de l'introduction et d'un patron commun des règles en amont de la rédaction du guide a posé une base aux échanges et débats sur les règles elles-mêmes.
- Au cours des ateliers, chaque règle rédigée en amont a fait l'objet d'adaptations tant de son périmètre que dans sa formulation, de manière à garantir son applicabilité par les utilisateurs finaux, les rédacteurs.
- Le format du tableau pour présenter toutes les informations relatives à une règle a fait rapidement consensus. Les détails de format du tableau se sont affinés au fil des échanges.

Le profil des experts et leur implication dans le projet devaient répondre à l'enjeu majeur de l'utilisabilité du guide.

- Les experts étaient des référents documentaires et représentaient chacun une société du groupe SNCF : SNCF Voyageurs, SNCF Réseau, SNCF SA. Ce profil offrait la garantie d'une représentativité des utilisateurs du guide, deux des trois experts étant également eux-mêmes rédacteurs.
- L'implication des experts métiers dès le début du projet (participation à la constitution du corpus et à la restitution des résultats) est un prérequis au bon déroulement des ateliers et à la qualité des échanges. Les experts avaient ainsi une bonne connaissance du sujet et de la finalité du guide.
- Réaliser les ateliers avec un nombre de trois experts métiers a permis de représenter les différents métiers sans compromettre la fluidité des échanges.

Une organisation précise des ateliers a défini en amont le calendrier, le rôle de chacun et les délais à observer pour l'envoi des nouvelles versions du guide, des retours et des mises à jour. L'important était de maintenir une dynamique dans un travail collaboratif où la participation de chacun était cruciale. L'espacement de 15 jours en moyenne entre les ateliers s'est avéré un bon compromis pour assurer la continuité du projet et rester compatible avec la charge de travail des experts.

Il faut noter ici que le choix primordial d'extraire des exemples du corpus s'est avéré chronophage. Les exemples extraits de corpus demandent à être remaniés pour illustrer au plus près la règle et faire sens auprès des utilisateurs finaux.

## VI. CONCLUSION

À l'initialisation du projet, l'enjeu était de s'assurer que isChecker était suffisamment pertinent pour être mis en place dans le processus de standardisation de la rédaction des documents techniques. L'expérimentation telle qu'elle a été réalisée a permis de tester progressivement le logiciel et de confirmer son potentiel, mais également d'impliquer les utilisateurs finaux tout au long du projet, renforçant ainsi l'intérêt et l'adhésion à l'utilisation d'un tel outil. Soulignons également le cercle vertueux de la démarche d'adaptation progressive de isChecker au contexte SNCF : les vérifications automatiques des documents techniques ont permis de confronter les règles de rédaction SNCF à la réalité des textes, et ainsi d'opérer des ajustements sur les règles elles-mêmes. Un travail d'optimisation du guide de rédaction SNCF a ainsi été réalisé ; la diffusion de cette nouvelle version est prévue pour 2022, tout comme la finalisation du mode expérimental à travers une mise en service en mode pilote du logiciel. isChecker pourra ainsi accompagner le déploiement du guide de rédaction standard de SNCF, en évaluant la conformité des futurs documents produits.

La démarche a contribué à la dynamique de professionnalisation et normalisation de la rédaction de SNCF, dans le cadre de sa transformation documentaire. Ancrée dans des règles, méthodes et outils à l'origine

développés dans le milieu aéronautique, cette démarche a confirmé la volonté de renforcer la dynamique de normalisation dans le ferroviaire.

## REMERCIEMENTS

Nous remercions le panel d'experts métiers qui a largement contribué aux résultats présentés dans cet article.

## RÉFÉRENCES

- [1] J.-M. Adam, « Types de textes ou genre de discours ? Comment classer les textes qui disent de et comment faire ? », *Langages*, n° 141, 2001, p. 10-27.
- [2] S. Dalbin et B. Guyot, « Documents en action dans une organisation : des négociations à plusieurs niveaux », dans *Études de communication*, 2007. Consulté le 20 décembre 2021 : <https://journals.openedition.org/edc/467>.
- [3] F. Ganier, « Écrire pour transmettre des connaissances procédurales : le cas des documents techniques procéduraux », *Pratiques*, 2016, p. 171-172, mis en ligne le 07 février 2017 : <http://journals.openedition.org/pratiques/3168>.
- [4] GIFAS, *Guide du français rationalisé*, 2<sup>e</sup> édition, 1999.
- [5] G. Olivier, B. Berthelet (ArianeGroup), V. Andréani et C. Raynal (Safety Data), « Le Traitement Automatique des Langues : facilitateur de la gestion documentaire », *Congrès Lambda Mu 22 « Les risques au cœur des transitions »*, 22<sup>e</sup> Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement, Institut pour la Maîtrise des Risques, Le Havre, e-congrès, France, 2020.
- [6] J.-P. Sintès, N. Kerzreho, et J.-J. Thomasson, « 2. Le défi des normes », *Documentaliste-Sciences de l'Information*, vol. 51, n° 1, 2014, p. 46-55 : <https://www.cairn.info/revue-documentaliste-sciences-de-l-information-2014-1-page-46.htm>.