

Interactions complémentaires au résumé parlé

Sommaire

7.1	Frise chronologique interactive	133
7.2	Description du prototype	134
7.2.1	Interface utilisateur	135
7.2.2	Architecture technique	137
7.3	Enquête utilisateurs	137
7.3.1	Principe	138
7.3.2	Résultats	139
7.4	Conclusion	143

Dans les chapitres précédents, une méthode de résumé automatique de parole adaptée à un contexte interactif a été présentée et évaluée de manière indirecte sur des données textuelles. Il s'avère que le gros défaut du résumé par extraction est le manque de structure des informations présentées. En effet, créer un résumé par juxtaposition de phrases retirées de leur contexte, et sans autre lien que leur représentativité thématique, a tendance à générer une réponse dénuée de cohérence. Nous allons tenter d'ébaucher quelques représentations complémentaires au résumé parlé afin de donner à l'utilisateur une idée de la structure des données représentées (section 7.1). Pour cela, le prototype implémenté est décrit en 7.2 et son potentiel est évalué par une enquête, de portée restreinte, auprès d'un panel d'utilisateurs en 7.3. Le lecteur se référera à la section 2.1.7 pour un bref historique des différentes interfaces proposées pour faciliter l'accès à des données parlées.

7.1 Frise chronologique interactive

La plupart des échecs des systèmes de recherche d'information provient d'une mauvaise perception du besoin de l'utilisateur exprimé au travers d'une requête. Pour ou-

trepasser ce genre de situation, l'utilisateur n'a d'autre choix que de reformuler son besoin jusqu'à obtenir un succès. En général, il utilise deux types de sources d'information pour cette reformulation : sa connaissance générale du domaine et la structure de la base documentaire traitée. La section 2.1.5 présente quelques méthodes d'expansion de requête pour guider l'utilisateur dans l'expression de son besoin. Ces méthodes sont orientées par les cooccurrences des mots de la requête dans les documents considérés comme pertinents. Cette seule dimension ne permet pas de discriminer les mots proposés en fonction de la structure thématique de l'«espace informatif». [Chuang et Chien \(2004\)](#) construisent par exemple une hiérarchie thématique à partir des résultats d'un moteur de recherche. La distribution temporelle de l'information est un autre axe fortement structurant dans le cas de nouvelles radio-diffusées.

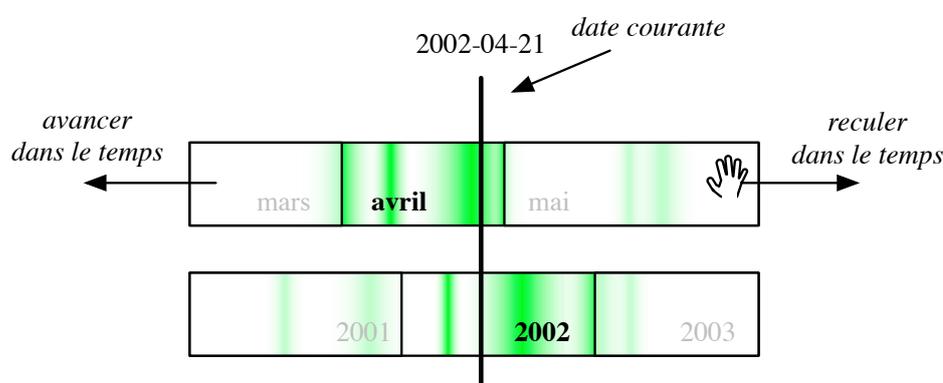


FIG. 7.1: Fonctionnement de la frise chronologique. Le curseur central représente la position de la lecture dans le temps. Chaque échelle est centrée sur ce curseur en fonction de la date courante. L'utilisateur peut faire glisser les échelles pour explorer les résultats dans le temps. La densité de résultats pertinents est représentée sur chaque échelle par un dégradé.

Nous proposons d'étudier cette distribution à travers une frise chronologique interactive. La figure 7.1 illustre son fonctionnement. L'information potentielle d'une tranche de temps est représentée comme la densité de résultats pertinents dans cette tranche. Au lieu d'avoir une résolution ajustable, les différentes granularités sont discrétisées sous la forme de plusieurs échelles temporelles synchronisées sur l'instant de lecture du flux audio. L'utilisateur peut explorer les différentes échelles en les faisant glisser vers le futur ou le passé de cet instant. L'idée derrière cette répartition en échelles est de créer une correspondance avec un système de quantification du temps naturel pour l'utilisateur : les années, les mois, les jours.... La frise chronologique est implémentée conjointement au résumé de parole dans le prototype présenté par la section suivante.

7.2 Description du prototype

L'ensemble des techniques présentées dans ces travaux est implémenté au sein d'un prototype dont l'objectif est de démontrer leur faisabilité technique et d'en effectuer

une analyse globale. La mise en place du démonstrateur complet permet la validation des concepts proposés dans une approche impliquant directement l'utilisateur.

La problématique principale étudiée dans ce document est provoquée par la quantité de données audio à écouter dans le cadre d'une recherche documentaire audio. Si le cas d'un moteur de recherche textuel est considéré, les documents retrouvés sont montrés à l'utilisateur sous forme d'une vue condensée incluant diverses informations jugées utiles pour déterminer rapidement la pertinence d'un document. Ces informations incluent souvent le titre, le thème, l'auteur, la source, ou la date de publication du document. Parfois, des extraits sont ajoutés pour contextualiser les résultats. Cette première représentation permet à l'utilisateur de décider s'il souhaite continuer l'exploration dans la direction d'un certain document. Le système lui délivre alors le document textuel que l'utilisateur commence par observer avant d'en lire une ou plusieurs parties. Très rapidement, en fonction de son expérience, l'utilisateur peut déterminer si un approfondissement s'avère nécessaire ou s'il lui faut passer au document suivant pour obtenir la réponse à son besoin en information. En général, la charte graphique, la structure thématique et une lecture rapide dirigent la décision précédente avec succès et rapidité. Dans le cas de l'audio, l'impossibilité d'avoir un aperçu global d'un document force l'utilisateur à passer beaucoup plus de temps à écouter le contenu. Cette perte de temps peut devenir fortement handicapante dans le cas de bases de données audio de grande taille.

La solution de résumer les résultats du moteur de recherche trouve sa valeur réelle dans l'exploitation de données audio. Nous proposons donc de concevoir un démonstrateur incluant un moteur de recherche interrogeable de la même façon qu'un moteur de recherche textuel, mais dont les résultats sont résumés sous la forme d'extraits audio. Afin d'améliorer la perception de la structure des résultats par l'utilisateur, deux représentations supplémentaires lui sont proposées : des mots-clés évoquant le contenu et une frise chronologique interactive pour naviguer rapidement dans les résultats.

7.2.1 Interface utilisateur

La figure 7.2 montre un écran représentatif des éléments du démonstrateur. Tout d'abord, un champ de requête permet d'entrer une requête textuelle et de la soumettre au moteur de recherche. Les résultats retrouvés sont présentés sous trois formes complémentaires : une frise chronologique interactive, une série de mots-clés et une liste d'extraits sélectionnés pour le résumé.

Les segments présentés dans la partie inférieure de la capture-écran ont été sélectionnés parmi les résultats du moteur de recherche en utilisant MMR-LSA¹, la méthode de résumé automatique présentée dans la section 5.3.1. Il est possible d'écouter les extraits, de voir leur transcription et d'explorer l'espace informatif qu'ils représentent en utilisant leur transcription comme requête. La transcription n'est pas montrée directement à l'utilisateur car il a tendance à trop lui faire confiance et à ne pas écouter l'audio

¹Les données résumées sont les 100 heures d'ESTER. Le modèle LSA a été appris sur 300 millions de mots du journal Le Monde. Les autres paramètres sont identiques à ceux appliqués sur DUC 2006.

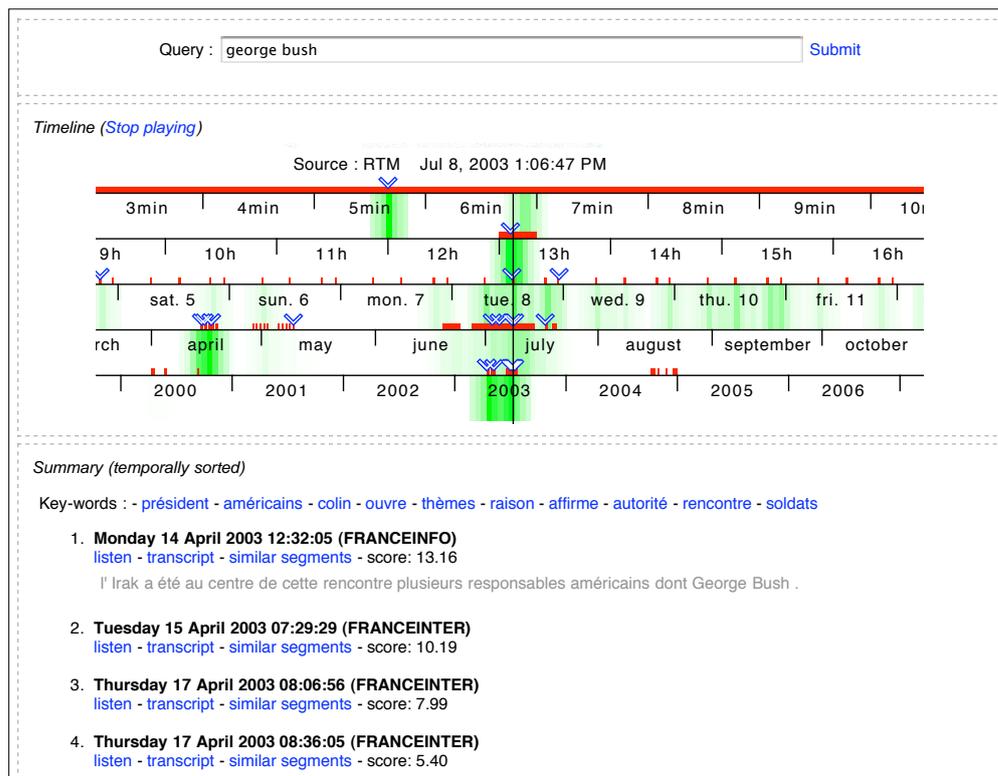


FIG. 7.2: Capture écran de l'interface utilisateur du prototype de moteur de recherche audio incluant de haut en bas : un champ de requête, une frise chronologique interactive, des mots-clés et les extraits audio sélectionnés pour le résumé. La frise chronologique permet de naviguer selon plusieurs échelles temporelles sur lesquelles sont représentés la densité de résultats (dégradé vert), les données disponibles (rectangles rouges) et les extraits sélectionnés pour le résumé (chevrons bleus).

(Hirschberg et al., 2001). Des informations supplémentaires caractérisent les extraits : leur date de diffusion, leur source et le score de pertinence assigné par le moteur de recherche. Les segments sont ordonnés temporellement pour ne pas briser la chronologie des événements.

Ce premier résumé des résultats est complété par une liste de mots-clés construite par la même méthode (MMR-LSA) mais dans l'espace de représentation des mots et non des phrases. Ces mots-clés sont généralement signifiants mais un filtrage sur les formes syntaxiques permettrait d'en améliorer l'utilité. L'utilisateur peut affiner sa requête en cliquant sur les mots-clés. Cette action les ajoute à la requête et met à jour les résultats.

Une représentation temporelle des résultats et de leur résumé est conçue dans le but d'améliorer la capacité de l'utilisateur à localiser l'information dans le temps. Les résultats du moteur de recherche font appel à deux types de chronologie : la chronologie naturelle des événements (ce qui arrive) et la chronologie de leur apparition dans les thèmes du flux audio (quand on en parle). Afin de minimiser la confusion de l'utilisateur, seul le second type de chronologie est utilisé, mais il serait très intéressant

d'intégrer ces deux chronologies dans une même visualisation bénéfique à l'utilisateur. La frise chronologique du démonstrateur permet de naviguer dans l'ensemble de la base de données audio selon plusieurs échelles temporelles (année, mois, jour, heure, minute) sur lesquelles sont représentés la densité de résultats (dégradé vert), les données disponibles (rectangles rouges) et les extraits sélectionnés pour le résumé (chevrons bleus). Le curseur central représente la date courante de lecture du flux audio. L'utilisateur peut déplacer l'une des échelles afin d'observer l'évolution de la densité de résultats sur une autre échelle, ou de positionner le curseur à une date précise. La densité de résultats est calculée par rapport au score de chaque segment, quantifié sur une durée temporelle fixe (par exemple toutes les 5 minutes pour l'échelle des heures). Les intervalles sont normalisés sur la durée affichée (pour avoir un contraste relatif plus intuitif), et interpolés selon l'équation 7.1.

$$d_k(t_1, t_2) = \sum_{j=-k}^k \frac{nb_p \{(t_2 - t_1)j < t(p) - t_1 \leq (t_2 - t_1)(j + 1)\}}{|j| + 1} \quad (7.1)$$

Dans cette équation, $d_k(t_1, t_2)$ est la densité entre un temps t_1 et un temps t_2 , prenant en compte k voisins pour l'interpolation. $nb_p(\cdot)$ est le nombre de points p dans l'intervalle $[t_1; t_2]$ décalé de $j \times (t_2 - t_1)$, $|\cdot|$ est la valeur absolue de son argument, et $t(p)$ représente la date associée à un point.

7.2.2 Architecture technique

Le schéma technique général du prototype est décrit dans la figure 7.3. La structuration automatique de la base de donnée audio et son indexation sont effectuées hors ligne. Les étapes de segmentation acoustique, et d'indexation en locuteur, de transcription, de segmentation en phrases, d'extraction d'entités nommées et d'indexation sont opérées en amont car ces traitements sont coûteux en temps et en ressources. L'architecture s'appuie sur ces données pour proposer un service de recherche d'information audio à l'utilisateur. Côté serveur, sont opérés les traitements «métier» comme la recherche d'information, l'extraction de mots-clés et la création de résumés. Côté client, les résultats sont mis en forme de façon indépendante. La diffusion de flux audio doit avoir les mêmes performances que ce soit en local ou à travers un lien réseau, tout en respectant les contraintes temps réel liées au média (délais, gigue...).

Des standards technologiques et des logiciels *open-source* sont utilisés pour faciliter l'intéropérabilité avec les systèmes existants, faciliter le déploiement du démonstrateur et préparer une éventuelle intégration produit. Les technologies utilisées sont détaillées dans la table 7.1.

7.3 Enquête utilisateurs

Cette section relate les détails de l'enquête menée auprès d'utilisateurs vis-à-vis du démonstrateur.

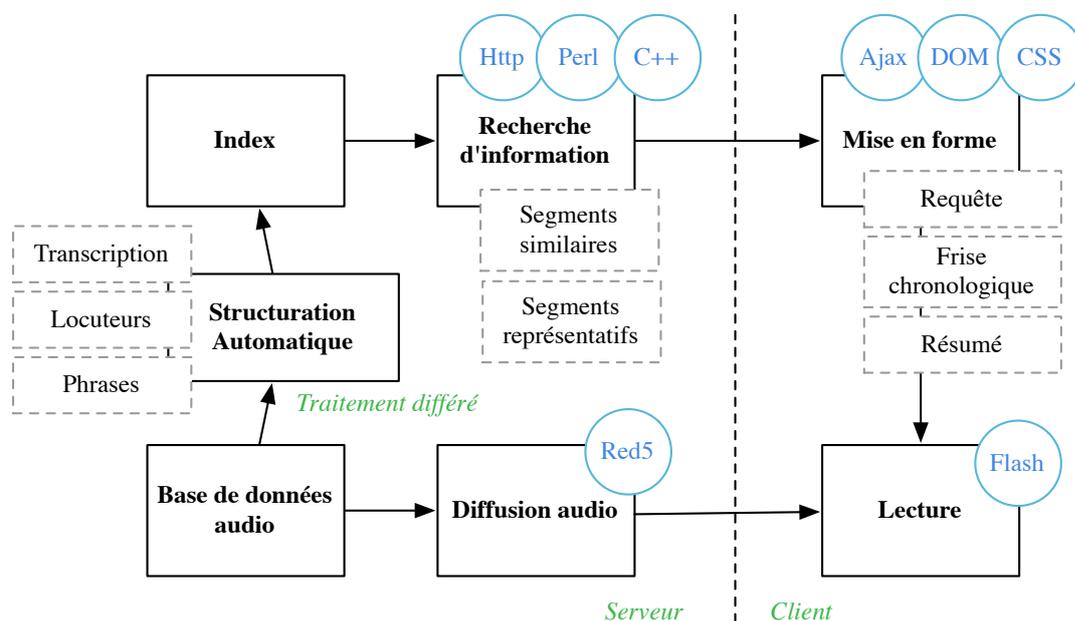


FIG. 7.3: Architecture technique pour la mise en place du prototype. La structuration est réalisée hors ligne ; la recherche d'information, le résumé et la diffusion audio sont proposés sous forme de services et exploités par un client léger fondé sur des standards technologiques.

7.3.1 Principe

En plus de l'évaluation de l'algorithme MMR-LSA sur les données d'une campagne reconnue (DUC 2006), l'interface est évaluée à l'aide lors d'une enquête auprès d'utilisateurs. Cette enquête repose sur un protocole du type « pensez tout haut » lors duquel l'utilisateur doit décrire oralement le maximum de ses réflexions, intentions et actions. Ce protocole a été étudié par [Ericsson et Simon \(1993\)](#) qui le qualifie d'« excellent moyen d'observer le cheminement mental d'un utilisateur ». Ce protocole est largement utilisé dans de nombreux domaines cognitifs ou liés à des interactions avec l'humain comme en ergonomie. [Nielsen \(1993\)](#) souligne qu'il faut choisir les sujets avec soin pour que ce protocole s'avère efficace et qu'un nombre réduit de sujets peut suffire dans le domaine des interfaces logicielles.

Dans le cadre de nos expériences, des sessions de 4 utilisateurs représentatifs ont été enregistrées et analysées. Chaque session dure environ 45 minutes pendant lesquelles le sujet s'exprime sur son expérience, guidé par un interlocuteur connaissant déjà le prototype. Les sessions se passent en plusieurs étapes durant lesquelles l'utilisateur réalise différentes tâches face au démonstrateur :

1. lecture du descriptif de l'expérience présenté en figure 7.4 ;
2. description de son profil en détaillant son expérience de l'outil informatique en général et des moteurs de recherche ;

Transcription	Speeral (Outil LIA)
Suivi de locuteurs	Alize/LIA_SpkSeg (Outil LIA)
Segmentation en phrases	CRF++
Entités nommées	AT&T fsm/grm
Serveur HTTP	serveur léger (Perl), support Ajax
Moteur de recherche	Modèle vectoriel, serveur léger (Perl/C++)
Résumé automatique	MMR-LSA, serveur léger (Perl/C++)
Serveur audio	Red5
Lecteur audio	Flash (Ming/Perl)
Frise chronologique	Applet (Java)
Mise en forme	DOM, CSS et Javascript, (navigateur web)

TAB. 7.1: Technologies utilisées dans les différents composants du démonstrateur. Red5 est un serveur de diffusion pour du contenu Flash temps réel (<http://osflash.org/red5>, visité en septembre 2006). Ming est une bibliothèque de génération d'applicatifs Flash (<http://ming.sourceforge.net>, visité en septembre 2006).

3. première confrontation visuelle avec l'interface sans l'utiliser, description et intuition sur les fonctions qu'elle peut accomplir ;
4. découverte de l'interface, l'utilisateur peut «jouer» avec l'interface en essayant ses différentes fonctions ;
5. mise en situation en remplissant un besoin en information concret avec une analyse des résultats ;
6. retour sur impressions permettant de collecter l'avis de l'utilisateur sur les points forts et les manques du logiciel.

7.3.2 Résultats

Pour chaque phase de l'enquête, les enregistrements ont été analysés. Les points suivants en résument les traits principaux.

– **Profil utilisateur :**

Tous les sujets utilisent le web régulièrement, mais montrent différents niveaux de connaissance de l'informatique théorique. Le genre est distribué uniformément entre les sujets. Ils ont de 17 à 30 ans.

– **Intuitivité a priori :**

Les utilisateurs doivent décrire l'interface et imaginer comment elle marche sans interagir avec elle. Plus précisément, ils doivent décrire l'interface comme si leur interlocuteur ne l'avait pas sous les yeux et ne la connaissait pas. Cet exercice permet de faire apparaître les points non intuitifs du logiciel.

Lors de cette phase, les sujets ont tous comparé l'interface à celle d'un moteur de recherche web comme Google en imaginant qu'elle devait fonctionner de la même

Interface d'accès à une base de données audio.

Ce logiciel permet de trouver et écouter des informations audio provenant de plusieurs radios entre 1998 et 2004. Vous allez participer à une enquête afin de tester la validité du prototype auprès d'utilisateurs. Le cheminement de l'enquête est décrit sous forme de grandes étapes guidées par des questions. Afin de mieux observer votre cheminement mental, vous devrez **penser à voix haute**. Vous serez enregistré(e) pour des analyses ultérieures mais vos informations personnelles ne seront pas diffusées; vous avez un droit de retrait conformément à la loi.

0) Votre profil.

Remplissez la feuille concernant vos informations personnelles.

1) Première approche visuelle.

Comment imaginez-vous que fonctionne ce logiciel ? Décrivez les éléments que vous voyez. A quoi pensez-vous qu'ils servent ? Qu'attendez-vous de ce logiciel ?

2) Premier contact.

Familiarisez-vous avec l'interface en jouant avec ses divers composants. Fonctionne-t-elle comme vous vous l'imaginiez ? Décrivez le fonctionnement de l'interface comme si vous l'expliquiez à quelqu'un qui ne l'a jamais utilisée.

3) Utilisation effective.

Choisissez un thème en vous aidant de la liste suivante.

- Fin de la guerre en Irak
- Arrêt de production chez Coca-cola
- Coupe d'Afrique de football
- La crise au Front National
- Le SRAS

Avez-vous des connaissances sur ce thème ? Pourquoi l'avez-vous choisi ? Retrouvez les informations importantes liées à ce thème parmi les données audio. Décrivez votre stratégie. Êtes-vous satisfait des informations trouvées ? Pensez-vous avoir retrouvé toutes les informations importantes liées au thème ?

4) Sentiment général.

Qu'est-ce qui vous a plu, déplu ? Que faudrait-il ajouter à ce logiciel ?

FIG. 7.4: Feuille de questions remis aux utilisateurs.

façon mais sur des données audio. Malgré l'intuitivité apparente du moteur de recherche, ils ont trouvé que la frise chronologique paraissait obscure et difficile à utiliser. En effet, cette dernière ne correspond pas à une représentation classique et contient des éléments qui ne « s'expliquent pas d'eux même ». De plus, certains éléments liés aux résultats du moteur de recherche ne sont pas affichés avant la première interaction, ce qui ne facilite pas l'interprétation.

– **Découverte du système :**

Les utilisateurs peuvent interagir avec le logiciel et sont encouragés à en décou-

vrir les fonctionnalités en les essayant une par une. Ils n'ont toutefois aucune explication sur le fonctionnement du système.

Durant cette étape, les sujets ont dû inventer des requêtes portant sur des émissions radiophoniques de 1998 à 2004 sans avoir d'autre *a priori* sur les données. Les approches utilisées pour inventer ces requêtes sont très intéressantes et font apparaître une problématique souvent ignorée dans le domaine de la recherche d'information : comment représenter une *vue* informative et concise du contenu d'une base de données spécialisée qui soit utile pour explorer de l'information nouvelle ? Les stratégies mises en œuvre pour inférer le contenu de la base sont, d'une part, d'utiliser des requêtes très génériques et d'analyser les réactions du système et, d'autre part, d'utiliser des requêtes précises sur des événements bien connus appartenant à la période étudiée. La seconde approche est beaucoup plus frustrante car la base de données ne couvre qu'une centaine d'heures et des dates très peu uniformes, donc un petit nombre d'événements.

Au niveau de la frise chronologique, la plupart des utilisateurs ont aimé la façon d'explorer le temps en faisant glisser l'une des échelles temporelles, mais pour des raisons ergonomiques, certains utilisateurs ont pensé que la frise était figée et n'ont même pas essayé d'interagir avec elle. Il est intéressant de noter que la transcription des extraits sélectionnés pour le résumé est cachée à l'utilisateur et qu'il doit cliquer sur un lien pour la faire apparaître. Le but de cette approche était de forcer l'utilisateur à écouter plutôt que de s'appuyer sur sa vision, car [Hirschberg et al. \(2001\)](#) ont observé que les utilisateurs ont tendance à trop avoir confiance dans la transcription et n'écoutent plus l'audio. Ainsi, les utilisateurs ne regardent la transcription que pour valider le comportement du système, dans les cas où il n'est plus cohérent (lorsque les erreurs de transcription ont un impact sur les résultats).

– Utilisation en situation réelle :

Après des explications sur le fonctionnement de l'interface, les utilisateurs doivent répondre à des besoins réels en information, sélectionnés dans une liste. Les sujets détaillent dans un premier temps leurs connaissances du thème choisi, les raisons de leur choix et la stratégie qu'ils comptent mettre en œuvre.

Dans cette mise en situation, les utilisateurs ont développé des stratégies variées tirant parti de la frise chronologique, de la liste d'extraits ou une utilisation conjointe de ces deux éléments. Il est intéressant de noter qu'un utilisateur s'est concentré sur la liste de mots-clés et sur les transcriptions, sans écouter l'audio avant d'être sûr d'avoir trouvé l'information recherchée.

Les différentes requêtes proposées aux utilisateurs ont été spécialement conçues pour couvrir un maximum de situations : pour certaines, le moteur de recherche retrouve plus ou moins d'extraits, voire aucun extrait ; d'autres nécessitent une reformulation ; certaines expriment des arguments à l'opposé de ce qui sera retrouvé dans les données ; d'autres sont délibérément à propos d'événements ambigus. Les utilisateurs, confrontés à ces requêtes, ont bien compris les particularités des requêtes proposées et ont été satisfaits par les résultats obtenus.

– Retour sur expérience et suggestions :

À la fin de l'expérience, il est demandé aux utilisateurs de commenter leur session, de donner leur opinion sur le système et éventuellement de suggérer des nouveautés à intégrer au démonstrateur.

La majorité des suggestions concernent l'ergonomie de l'interface qui n'est pas suffisamment intuitive et nécessite beaucoup de mouvements à la souris. Par exemple, la lecture audio peut être arrêtée soit en cliquant sur le lien en haut de la page, soit en faisant glisser le curseur de la frise temporelle sur une zone non couverte par des données. Un raccourci clavier ou le remplacement du lien « lecture » des extraits issus du résumé faciliteraient l'ergonomie du lecteur audio. Il a aussi été suggéré que des mouvements fluides de la frise chronologique pourraient améliorer la compréhension du processus (lorsqu'un extrait du résumé est sélectionné sur la frise ou sur l'interface). Afin d'améliorer l'intuitivité de la frise chronologique, il faudrait afficher moins d'échelles temporelles différentes (par exemple 2) avec la possibilité d'en changer la précision. Le problème des multiples échelles est qu'elles représentent plusieurs fois la même information. Plus généralement, les utilisateurs auraient apprécié des icônes représentatives, des pointeurs de souris et des aides contextuelles leur montrant les possibilités de l'interface.

La fonctionnalité jugée comme étant la plus intéressante est l'idée de supprimer complètement la liste des extraits du résumé et de représenter leur contenu au sein même de la frise chronologique. Dans cette optique, des mots-clés sont utilisés pour annoter les zones denses en information pertinente. Il faudrait également développer un moyen de représenter la transcription, sachant que l'échelle de représentation du texte par rapport au temps n'est pas la même que celle de l'audio. À propos du sentiment général des utilisateurs sur le prototype, ils ont eu des réponses variées et pas toujours positives. Par exemple, une utilisatrice a remarqué qu'elle ne voyait « pas l'intérêt d'une base de données audio constituée de journaux radio-diffusés alors qu'elle peut trouver les mêmes informations au format textuel sur internet, le texte étant tellement plus rapide à lire ». Un autre utilisateur a précisé qu'il avait « peur des moteurs de recherche car ils indexent de plus en plus d'informations liées à notre vie privée qui pourraient être utilisées à mauvais escient par la personne qui les contrôle ».

Conclusions de l'enquête

Nous avons voulu expérimenter le résumé automatique de parole dans les conditions réelles et une utilisation conjointe avec d'autres moyens de localiser l'information. Un prototype a été implémenté pour tester l'ensemble des composants dans des conditions réelles (chaîne de structuration, résumé automatique, interface). Ces composants permettent de construire un résumé de parole à partir de données audio entièrement automatiquement. Le prototype a été évalué au moyen d'une enquête préliminaire auprès d'utilisateurs. Cette enquête a révélé qu'une frise chronologique interactive était

un bon complément à la couverture thématique du résumé. Toutefois, il serait intéressant de parfaire cette approche par l'introduction de la chronologie des événements, reliée à celle des données. Une telle perspective nécessite tout de même la résolution des références temporelles pour dater les informations au niveau de la phrase. Un autre point attirant serait une représentation de la granularité thématique sous une forme similaire à celle de la frise chronologique. Couper une arborescence thématique à plusieurs hauteurs de l'arbre peut être une première idée, mais y introduire une interactivité ne sera sans doute pas trivial.

7.4 Conclusion

Dans ce chapitre, un prototype de résumé automatique de parole complet a été implémenté pour tester les réactions d'utilisateurs dans des conditions réelles. Ce prototype sert avant tout de support pour tester des représentations conjointes au résumé dans le but de dépasser les limites de l'approche par extraction. Ces derniers travaux représentent une ouverture vers la recherche de méthodes pour l'accès à l'information parlée pour aller plus loin qu'une simple amélioration des méthodes de structuration et de sélection de l'information.

Chapitre 8

Conclusion

Depuis près d'un siècle, l'humanité construit une société de l'information, centrée sur le partage des connaissances. Grâce à de nombreuses avancées techniques dans le domaine des communications, chaque individu peut potentiellement accéder au savoir de l'ensemble de la société. Internet, un des principaux vecteurs d'informations et de connaissances, semble jouer le rôle de système nerveux de notre société. Savoir accéder à l'information semble devoir devenir plus important que le Savoir. Toutefois, bien que la communication permette un transport des différents médias (texte, voix, image, vidéo), seule la connaissance écrite est réellement pérenne. En effet, l'accès à l'information n'est possible que par l'intermédiaire d'une recherche dans les descripteurs conceptuels du contenu des données. Alors que pour le texte, ces descripteurs sont directement extraits à partir des mots, les autres médias nécessitent une annotation textuelle (manuelle ou automatique) de leur contenu.

Dans ce document, nous nous sommes focalisés sur la parole, le moyen le plus utilisé pour communiquer. La parole a la caractéristique d'être éphémère : elle disparaît avec l'onde sonore qui la transporte. Les technologies actuelles permettent de transporter et conserver cette onde, et la rendent ainsi exploitable à plus long terme. Dès lors qu'il est possible de transcrire son contenu, elle acquiert les mêmes possibilités d'indexation que le texte. Ce processus a facilité l'apparition de moteurs de recherche sur le contenu parlé de flux audio. Ces derniers génèrent une liste d'extraits pour satisfaire le besoin de l'utilisateur. Cependant, une écoute exhaustive des extraits est essentiel pour déterminer leur pertinence. Cet effet est comparable à la nécessité d'explorer plusieurs pages de résultats dans un moteur de recherche Internet si l'information n'est pas trouvée dans la première page. Ce défaut, déjà présent à l'écrit, est exacerbé par le temps passé à l'écoute d'informations non pertinentes.

Nous avons proposé dans cette étude d'améliorer l'efficacité de l'accès à des bases de données parlées à l'aide d'une approche reposant sur le résumé automatique de parole. Ce concept, proche du résumé textuel, correspond à générer un résumé parlé à partir des données audio répondant au besoin d'un utilisateur. Une telle approche fait apparaître de nombreuses problématiques sur la portabilité des méthodes dédiées au texte, sur le traitement des données parlées et sur la manière de satisfaire l'utilisateur.

Le principal objectif de ce travail était de mettre en place une chaîne de traitement complète, en réutilisant les briques disponibles et en développant les éléments manquants, pour pouvoir étudier l'impact de chaque sous-système sur l'application visée.

Cette conclusion expose d'abord les résultats obtenus, puis détaille les perspectives de recherche majeures découlant de ce travail.

8.1 Résultats obtenus

Le chapitre 2 a tout d'abord présenté la recherche d'information à travers ses origines textuelles et ses adaptations à un média oral. Cette description s'est appuyée sur le domaine de la recherche documentaire (section 2.1) et celui du résumé automatique (section 2.2). Il est ressorti de cette étude qu'une bonne structuration était nécessaire pour passer de l'acoustique à une sémantique. Les éléments de structuration suivants ont été abordés dans le chapitre 3 : la segmentation en classes acoustiques, la segmentation en tours de parole, l'indexation en locuteurs et la transcription de la parole. Une implémentation cohérente de ces tâches sous la forme de la chaîne de transcription enrichie du LIA a ensuite été décrite et évaluée sur les données radiophoniques francophones de la campagne ESTER.

La chaîne de structuration a été complétée au chapitre 4 par deux descripteurs sémantiques et structurels majeurs : la segmentation en phrases et l'extraction d'entités nommées. La première est importante pour le résumé par extraction car la qualité des frontières de phrases est un paramètre influençant fortement la qualité de la forme du résumé. La seconde a pour but d'améliorer la projection du contenu dans un espace sémantique en faisant émerger des entités du domaine (dans notre cas : des noms de personnes, d'organisation, de lieux...) peu ambiguës hors contexte.

La solution proposée dans ce document pour la segmentation en phrases (section 4.1) repose sur une modélisation des frontières de phrase par des paramètres acoustiques et prosodiques à l'aide de *Conditional Random Fields* (CRF). Notre approche a obtenu des résultats au niveau de l'état de l'art sur les frontières de phrases ESTER avec une F_1 -mesure de 0.68. Par rapport aux techniques de segmentation en phrases de la littérature, notre approche constitue une première voie pour modéliser conjointement des événements linguistiques et prosodiques. Cependant, cette approche ne profite pas des valeurs de confiance des processus sous-jacents (transcription, segmentation en locuteurs) et ne permet pas un apprentissage des différentes modalités sur des jeux de données disjoints.

Notre solution pour l'extraction d'entités nommées dans le flux de parole (section 4.2) intègre l'espace d'hypothèses de la transcription pour tenter d'alléger les erreurs qui pourrait être commises dans sa meilleure hypothèse. La méthode proposée repose sur une annotation du graphe d'hypothèses de transcription par des grammaires locales couplées à un modèle probabiliste génératif. Elle a obtenu les meilleurs résultats sur la tâche expérimentale d'annotation en entités nommées de la campagne ESTER. Ce résultat doit cependant être tempéré : Les écarts de performances relevés entre les différentes

approches et systèmes proposés durant ESTER restent cependant peu significatifs.

La seconde partie de ce document (chapitres 5 et suivants) a été consacrée au résumé automatique de parole dans notre cadre applicatif. Le chapitre 5 a présenté les problèmes liés à l'interactivité et à un contenu parlé. Tout d'abord, les contraintes d'interactivité nous ont poussés à modifier le modèle général de résumé automatique, en séparant les paramètres liés au besoin de l'utilisateur de ceux considérés comme indépendants de ce besoin. En effet, l'estimation de la contribution d'une phrase pour le résumé peut être plus gourmande en ressources et calculée en temps différé pour les paramètres indépendants du besoin de l'utilisateur (ou adaptés à un utilisateur moyen). Par contre, les paramètres liés à une demande spécifique requièrent une inférence rapide. Le modèle proposé repose sur l'approche *Maximal Marginal Relevance* (MMR), vue comme une solution tendant à séparer les influences du fond et de la forme (les contraintes liées à la forme n'étant pas capturées par l'expression du besoin). Toutefois, le manque de données dédiées au résumé automatique de parole dans le cadre d'ESTER (cette campagne ne proposait pas de tâche directement liée au résumé automatique), n'a pas permis une évaluation exhaustive de la méthode proposée.

Pour pallier ce problème, le chapitre 6 a proposé une évaluation indirecte du modèle proposé sur des données textuelles, lors de la campagne *Document Understanding Conference* (DUC) 2006. Elle a donné lieu à une collaboration LIA-Thales pour une soumission fusionnant 5 systèmes et obtenant des résultats état de l'art (environ 6^e/34 soumissions, dont deux significativement meilleures). L'approche développée dans ce document a obtenu des performances juste en dessous de la fusion et meilleures que les 4 autres systèmes selon l'évaluation automatique Rouge-2. Obtenir des performances satisfaisantes sur des données textuelles n'offre pas la garantie d'un même niveau de performances sur la parole, essentiellement à cause des erreurs et des imprécisions ajoutées par la chaîne de structuration (découpage en phrases, erreurs de transcription...). En conséquence, nous avons évalué l'impact de tels problèmes sur la qualité des résumés proposés par notre approche en «bruitant» les données textuelles, pour simuler des données provenant d'une chaîne de structuration automatique. Les résultats obtenus ont démontré la robustesse de notre approche à de telles dégradations. Ce point reste cependant à confirmer, à cause de la mesure de performance utilisée (Rouge) et de la nature artificielle des données employées.

Dans le but de prouver la faisabilité d'un système de recherche d'information reposant sur les approches proposées dans ce document, nous avons décrit dans le chapitre 7 une implémentation complète sous la forme d'un démonstrateur. Réaliser un démonstrateur nous permettait également de prendre en compte les facteurs liés à l'interaction utilisateur-système. Cette implémentation a prouvé la faisabilité de la méthode proposée en terme de complexité de développement, de calcul et d'efficacité. Elle a facilité l'étude des différentes composantes en application réelle (données radiophoniques ESTER). Le facteur «interactivité» a été intégré sous la forme d'une présentation originale des résultats correspondant au besoin exprimé par l'utilisateur. Chaque segment audio extrait par le procédé de résumé est placé dans une frise chronologique interactive. L'utilisateur peut aisément naviguer dans les résultats et sélectionner une partie ou l'autre des segments, affinant ainsi l'expression de son besoin. Bien que l'impact de

cet élément soit difficile à évaluer, nous en avons proposé une première étude à l'aide d'une enquête utilisateur sur des aspects ergonomiques.

8.2 Perspectives

Nous allons explorer les perspectives de ce travail au regard des challenges listés par [Zechner \(2003\)](#) pour le résumé automatique de parole :

1. « améliorer la transcription automatique du contenu parlé ;
2. intégrer les informations prosodiques dans les méthodes textuelles de résumé automatique ;
3. faire converger les méthodes de résumé issues de la recherche d'information, robustes au domaine, et celles issues de l'intelligence artificielle, dépendantes du domaine ;
4. rapprocher le résumé automatique et la problématique questions/réponses ;
5. évaluer le résumé de parole grâce à des critères adaptés. »

Le résumé automatique de parole nécessite une amélioration des performances de la transcription automatique (1), et plus généralement de l'ensemble de la chaîne de structuration. L'étude proposée dans ce document est restreinte à un type de contenu radiodiffusé sur lequel la plupart des travaux en transcription ont été réalisés. En changeant de domaine et de conditions (moins bonnes conditions acoustiques, plus de parole spontanée, langue ou thème différents), la qualité de la structuration sera certainement détériorée à cause d'une plus faible quantité de données d'apprentissage et d'une plus grande variabilité des observations. Cette variabilité peut être compensée par la transformation des observations (débruitage), par augmentation de la diversité des situations observées en apprentissage, et par une amélioration des capacités de généralisation des algorithmes de reconnaissance. D'un autre côté, la variabilité des observations peut être amortie par l'utilisation de mesures de confiance entre les tâches de structuration. Ce procédé correspond à une extension des espaces de recherche de la meilleure hypothèse d'une étape à la suivante qui pourrait être généralisée à une fusion de tous les espaces de recherche. Une telle fusion n'est pas triviale car elle demande d'accorder des hypothèses provenant de données d'apprentissage et de cadres théoriques différents (espaces acoustiques ou linguistiques, discrets ou continus, cadre probabiliste...). L'apprentissage supervisé reste très coûteux car il requiert une annotation complète des données : il convient peut-être de se tourner vers des approches non-supervisées ou semi-supervisées comme celle de [Haghighi et Klein \(2006\)](#). Cette approche se fonde sur des exemples peu nombreux dont l'étiquetage est étendu à un corpus par des statistiques globales. Les performances restent bien en dessous des approches sur des corpus totalement étiquetés, mais un annotateur peut observer les résultats de l'étiquetage et réitérer l'apprentissage après ajout de nouveaux exemples représentatifs.

Plus spécifiquement, parmi les éléments ajoutés à la chaîne de traitement, la méthode de détection de frontières de phrases doit être étendue à une méthode de segmentation, au sens propre du terme. En effet, bien que les paramètres de *Conditional*

Random Fields dépendent d'une séquence dans son ensemble, le modèle n'est pas capable de prendre en compte la cohérence globale de la phrase, voire l'enchaînement des phrases. Cette limitation est due à l'hypothèse de Markov (localité des dépendances) qui facilite la résolution du problème. Les CRF Semi-Markoviens de [Sarawagi et Cohen \(2005\)](#) remettent en cause cette hypothèse avec pour conséquence une augmentation de la complexité de l'apprentissage et du décodage. Ce genre de modèle sera certainement bénéfique à l'annotation en entités nommées. Il faudra cependant instancier le modèle dans le cadre des automates à états finis ou trouver un autre moyen de prendre en compte le treillis d'hypothèses de transcription.

Nous n'avons pas abordé la détection de l'emphase par analyse prosodique (2) dans ce document car les méthodes d'extraction de prosodie sont relativement peu robustes dans un environnement acoustique contenant un fond sonore. Notre approche repose surtout sur les statistiques de cooccurrences des mots. Toutefois, la formulation présentée pour *Maximal Marginal Relevance* permet d'introduire des informations d'emphase au niveau de la granularité de la phrase. Il serait très intéressant de pouvoir employer cette caractéristique au niveau des mots et surpondérer les mots à forte intonation dans le modèle vectoriel. Néanmoins, ces considérations nécessitent de trouver un bon équilibre entre la contribution des informations prosodiques, des informations de fréquence et la contribution des autres mots de la phrase dans la normalisation du vecteur la représentant. De plus il faut différencier les nombreux mots informatifs non accentués des mots accentués non informatifs. De façon similaire, la méthode proposée dans ces travaux facilite l'intégration d'informations provenant du domaine (3) comme des indices lexicaux ou structurels. Pour compléter ce type d'intégration, il faut se poser la question de l'interaction entre les mesures de confiance issues de la structuration et les patrons linguistiques découverts.

La campagne d'évaluation DUC nous a offert la possibilité de construire un résumé par extraction grâce à la fusion des sorties de systèmes spécialisés dans le résumé automatique et de systèmes orientés questions/réponses (4). Ce rapprochement des deux domaines a été bénéfique (la fusion est meilleure que les systèmes séparés), mais la fusion proposée n'utilise que le pouvoir informatif estimé des phrases et ne prend pas en compte la redondance du résumé. Pour remédier à cette limitation, le problème de sélection des phrases doit être formulé à l'aide d'une fonction objective globale. Les équations 8.1 et 8.2 donnent une éventuelle formulation de ce problème prenant en compte la représentativité d'une phrase par rapport au besoin de l'utilisateur et aux documents d'origine, et l'inférence du contenu d'une phrase du résumé par une autre phrase du résumé. Dans ces équations, S est une sélection de phrases, s_i est une phrase, b est le besoin, D est l'ensemble des documents source.

$$\hat{S} = \operatorname{argmax}_S \text{information}(S) - \text{redondance}(S) \quad (8.1)$$

$$= \operatorname{argmax}_S \sum_{s_i \in S} \text{représentativité}(s_i, b, D) - \sum_{s_i \in S} \sum_{s_j \in S} \text{inférence}(s_i, s_j) \quad (8.2)$$

Cette formulation peut être assimilée à un problème de sac-à-dos quadratique en ajoutant la contrainte de longueur du résumé. Une solution exacte à ce problème étant

inaccessible lorsque le nombre de phrases augmente, il conviendra de mettre en place des heuristiques et des approximations pour traiter le problème dans un temps raisonnable.

Les difficultés d'évaluation du résumé automatique sont déjà nombreuses à l'écrit à cause de l'impossibilité de définir un résumé modèle de référence (5). Il n'existe pas à l'heure actuelle d'évaluation du résumé de parole à grande échelle, bien que NIST compte orienter prochainement *Document Understanding Conference* vers la parole. Nous avons proposé quelques critères spécifiques à ce média (Q6-Q11, section 2.2.1, page 47) pour une évaluation manuelle de la forme d'un résumé parlé. Les critères sur le fond sont-ils identiques à ceux utiles au résumé textuel ? Les mesures automatiques d'évaluation comme Rouge sont-elles adaptées à un média parlé ? Nous avons étudié la différenciation entre un résumé « lu » et un résumé « écouté ». Les deux types de résumé ont des applications très différentes et mèneront certainement à des approches différentes. Nous avons simulé l'impact d'un contenu parlé sur des données textuelles pour évaluer ces nouvelles conditions, mais une évaluation sur des données réelles sera nécessaire pour confirmer les résultats obtenus. Rouge, la mesure d'évaluation automatique dédiée à un développement rapide des systèmes de résumé, atteint ses limites, car elle n'offre aucune contrainte sur la forme du résumé et ignore sa redondance. La recherche d'une meilleure mesure d'évaluation pourrait être l'objet d'une prochaine campagne DUC, les participants proposant des systèmes capables de prédire au mieux le comportement des juges.

Un dernier point n'est pas évoqué par Zechner (2003) : les interactions avec l'utilisateur pour construire un résumé. En plus du besoin exprimé sous forme textuel, DUC introduit en 2007 les connaissances présumées de l'utilisateur au travers d'une tâche de mise à jour d'un résumé. Toutefois, ce type d'évaluation n'a pas la dynamique nécessaire pour étudier les interfaces complémentaires au résumé comme la frise chronologique présentée dans nos travaux. L'enquête préliminaire associée à cet élément devra être étendue en focalisant l'évaluation sur la perception de la structure temporelle des informations présentées par les résumés. Cette évaluation peut par exemple prendre la forme d'un questionnaire présenté à des utilisateurs ayant écouté un résumé et d'autres utilisateurs ayant écouté le résumé et ayant eu la possibilité d'interagir avec la frise chronologique. En plus de l'information temporelle, cette frise pourra être améliorée en ajoutant des indicateurs thématiques. Pour cela, la structure thématique devra être représentée conjointement à la structure temporelle tout en offrant les mêmes possibilités de navigation que la frise actuelle.

Finalement, il semble que les travaux sur le résumé automatique de parole doivent s'orienter vers un *modèle complet* ne faisant aucun *a priori* d'indépendance entre chacun des éléments du triplet acoustique, utilisateur et résumé.

Annexe A

Résultats DUC détaillés pour le *topic* D0641 (réchauffement climatique)

Topic	D0641E
Titre	global warming
Détails	Describe theories concerning the causes and effects of global warming and arguments against these theories.

A.1 Résumés générés par les systèmes S_1 à S_5 et leur fusion F_2

Résumé S_1 (MMR-LSA)

To them, the observed surface warming of about 1 degree over the with an especially sharp rise in the last quarter century is mostly or wholly natural, and there is no significant human influence on global climate. While the Global Climate Coalition questions whether global warming is happening, it advocates that companies voluntarily explore and employ new technology to reduce emissions that contribute to global warming. The changes are caused by a temperature increase, which is one of the effects of global warming. For that reason, slightly higher rates of warming and sea level rise related to warming are expected, according to a climate study written by Wigley and released June 29 by the Pew Center on Global Climate Change in Washington. The findings suggest that global warming blamed on human industry is manifesting itself not just by causing a global rise in temperatures but by favoring certain established weather patterns.

Rouge-2	0.12444
Rouge-SU4	0.17684

Résumé S_2 (CORTEX)

The changes are caused by a temperature increase, which is one of the effects of global warming. Greenhouse gas emissions blamed for global warming may cause the collapse of the West Antarctic Ice Sheet and raise the average global sea level by four to six metres, beginning as as the, a new scientific study predicted recently. While the Global Climate Coalition questions whether global warming is happening, it advocates that companies voluntarily explore and employ new technology to reduce emissions that contribute to global warming. The study believed that because carbon dioxide is one of the primary greenhouse gases contributing to the warming of the planet, the decrease of this greenhouse gas may help slow the global warming. The researchers, at the Max Planck Institute for Meteorology in Hamburg, Germany, used a computerized model of the climate system's general circulation to investigate the effects of global warming on El Nino's frequency.

Rouge-2	0.07072
Rouge-SU4	0.12892

Résumé S_3 (alignement)

Agriculture is estimated to be responsible for eight percent of the total greenhouse gas emissions blamed for global warming. South African Environment Minister Valli Moosa said on Friday that global warming resulted in the rise in incidence of malaria. While the Global Climate Coalition questions whether global warming is happening, it advocates that companies voluntarily explore and employ new technology to reduce emissions that contribute to global warming. Scientists looking for signs of global warming should spend more time scrutinizing Earth's weather circulation patterns, a new study suggests. Chinese scientists believe that global warming, particularly higher winter temperatures, will continue to the first half of the 21 st century. Scientists who argue against global warming often cite nearly 20 years of satellite readings showing slight cooling of the atmosphere two miles up. These effects open up a new way of tracking the progress of global warming without the uncertainties in simple temperature measurements. An American biologist has obtained the clearest evidence so far that global warming is forcing living species to move toward the poles and to higher altitudes. New study suggests that El Nino may help slow the global warming, though it has been blamed for many climate disorders such as floods and droughts. Under way in Buenos Aires is a giant international conference on global warming, a follow up to last December's United Nations sponsored confab in Kyoto, Japan.

Rouge-2	0.06572
Rouge-SU4	0.10855

Résumé S_4 (Q/A densité)

Methane and nitrous oxide are more powerful in retention of heat and potential causes for global warming. One ton of methane will have the same warming effect as 56 tons of carbon dioxide over a period of 20 years. North said the effect of global warming in Texas and the projected corresponding 1 to 2 foot rise in the sea level could add additional stress on already tight water supplies. The increase rate of gases that cause the greenhouse effect will accelerate, so much so that the earth's warming trend will continue in the coming dozens of years. The predicted warming is expected to cause widespread climatic, ecological and economic dislocation. McKelvey said the team's ultimate aim was to show the likely effect of future global warming. It is an enormous and expensive task to study a continental ice sheet that is twice as big as Australia. Heat waves triggered off by global warming have caused thousands of additional deaths each year in major cities worldwide, and would lead tens of millions of people to the risk of malaria, the report warns. When Wentz and co author Matthias Schabel took the falling satellite effect into account, they found the lower troposphere was warming 0.13 degrees Fahrenheit per decade. It can disrupt climate around the world, producing extra rain in the southeastern United States and in Peru during the winter, while causing drought in the western Pacific.

Rouge-2	0.07263
Rouge-SU4	0.12170

Résumé S_5 (Q/A compacité)

The changes are caused by a temperature increase, which is one of the effects of global warming. Hurricane forecasters doubt extreme weather events such as Mitch show the influence of global warming. I think it's agreed that the globe is about 1 degree Fahrenheit warmer, but as its effect on tropical cyclones, we haven't been able to notice any difference, said forecaster Brian Jarvinen at the U.S. National Hurricane Center in Miami. The theory is that at certain critical altitudes, roughly from 6 to 12 miles, increasing carbon dioxide has the effect of warming the tropics but cooling the polar regions. Methane and nitrous oxide are more powerful in retention of heat and potential causes for global warming. Sea levels would rise as the Antarctic ice sheet melted because of higher temperatures caused by global warming. The theory is that a warmer atmosphere holds more water vapor, increasing the warming more. This warming would touch off widespread disruptions in climate and weather and cause the global sea level to rise and flood many places. North said the effect of global warming in Texas and the projected corresponding 1 to 2 foot rise in the sea level could add additional stress on already tight water supplies. Computer models of the effects of warming point to the appearance of jets of fast moving wind between the troposphere and stratosphere, roughly 12 kilometers above the Earth's surface.

Rouge-2	0.08556
Rouge-SU4	0.13721

Résumé F_2 (Fusion des 5 systèmes)

Scientists looking for signs of global warming should spend more time scrutinizing Earth's weather circulation patterns, a new study suggests. The findings suggest that global warming blamed on human industry is manifesting itself not just by causing a global rise in temperatures but by favoring certain established weather patterns. While the study does not prove that human induced global warming is what caused the change in frequency of circulation patterns, he said, the change is consistent with it. The theory is that at certain critical altitudes, roughly from 6 to 12 miles, increasing carbon dioxide has the effect of warming the tropics but cooling the polar regions. The data suggest that global warming has caused a slowdown of the Earth at a rate of 0.56 milliseconds a century..

This warming would touch off widespread disruptions in climate and weather and cause the global sea level to rise and flood many places.

In Beijing, chinese scientists believe that global warming, particularly higher winter temperatures, will continue to the first half of the 21 st century.

In San Antonio, while the Global Climate Coalition questions whether global warming is happening, it advocates that companies voluntarily explore and employ new technology to reduce emissions that contribute to global warming.

In Washington, greenhouse gas emissions blamed for global warming may cause the collapse of the West Antarctic Ice Sheet and raise the average global sea level by four to six metres, beginning as as the, a new scientific study predicted recently.

Rouge-2	0.08564
Rouge-SU4	0.14662

A.2 Résumés de référence

Référence D0641E — C

Global warming is thought to be at least partly caused by emissions of waste industrial gases like carbon dioxide, produced by burning fossil fuels like coal, oil and natural gas. These emissions trap solar radiation and produce a greenhouse effect. Methane and nitrous oxide emissions from agriculture (ruminants and manure) make up 8% of greenhouse gases. Controls on sulfur dioxide emissions reduce a balancing cooling effect.

Global warming already causes more frequent El Nino appearances, receding shorelines, longer warm seasons, and a slower earth spin. It affects habitats and threatens marine life. If emissions are not reduced, average surface temperature will rise 2-6 degrees over the next century, bringing widespread climatic, ecological and economic dislocation. Floods and droughts will increase in frequency and intensity. Melting polar

ice will cause rising sea levels and coastal flooding. Malaria will increase. Rates of habitat loss and species extinction will increase. Communities will need to adapt to new conditions.

Skeptics argue that human activities have little influence on climate. Most observed warming is due to natural causes like changes in solar radiation or the circulation of heat-bearing ocean waters. Measurements taken by satellites have found little temperature rise in the upper atmosphere. Computer models are unreliable. Any warming over the next century would be most pronounced in the winter, at night, and in sub-Arctic regions, doing little harm and creating benefits like longer growing seasons and faster plant growth. Industry argues that reducing the use of fossil fuels would cause economic harm to consumers.

Référence D0641E — E

As early as 1968 scientists suggested that global warming might cause disintegration of the West Antarctic Ice Sheet. Greenhouse gas emissions created by burning of coal, gas and oil were believed by most atmospheric scientists to cause warming of the Earth's surface which could result in increased frequency and intensity of storms, floods, heat waves, droughts, increase in malaria zones, rise in sea levels, northward movement of some species and extinction of others.

Some scientists, however, argued that there was no real evidence of global warming and others accepted it as a fact but attributed it to natural causes rather than human activity. In 1998 a petition signed by 17,000 U.S. scientists concluded that there is no basis for believing (1) that atmospheric CO₂ is causing a dangerous climb in global temperatures, (2) that greater concentrations of CO₂ would be harmful, or (3) that human activity leads to global warming in the first place.

By 1999 an intermediate position emerged attributing global warming to a shift in atmospheric circulation patterns that could be caused by either natural influences such as solar radiation or human activity such as CO₂ emissions.

By 2000 opponents of programs to cut back greenhouse emissions admitted that there was evidence of global warming but questioned its cause and dire consequences. Proponents of plans to control emissions to a large extent admitted that the size of the human contribution to global warming is not yet known.

Référence D0641E — F

By 2000 most climate scientists agreed that the Earth's atmosphere warmed about 1 degree Fahrenheit in the 1990s alone. Earlier debates about conflicting evidence for atmospheric warming were resolved when weather satellite data were recalibrated. Also agreed is that atmospheric concentrations of greenhouse gases are rising – 360 parts per million, up from 315 ppm in the late 1950s ; that carbon dioxide (CO₂) is nearly 30% higher than before the Industrial Revolution and the highest in the last 420,000 years.

Agriculturally produced methane and nitrous oxide make up 8% of greenhouse gases, small compared to CO₂. Atmospheric angular momentum correlates with atmospheric temperature rises.

All agree that we don't yet know definitively what is causing global warming. The debate is not sharply polarized, but a broad spectrum of nuanced scientific views, always open to new evidence.

The dominant scientific view, expressed by a U.N. science panel, concluded that increased man-made greenhouse gases, especially CO₂ from fossil fuel burning, are significant causes of rapid global warming, which if not reduced, would raise global temperature by 2.3 to 7.2 degrees over the next 100 years. In 2000, the U.S. EPA concluded that fossil fuel burning contributes to global warming.

A significant minority – over 170,000 scientists – have questioned the accuracy of climate models predicting dangerous heat-raising effects of man-made greenhouse gases. They said naturally-produced water vapor is far more significant, and that doubling atmospheric CO₂ would add only 1 degree Fahrenheit by 2100 – near the low end of opponents' estimates.

Référence D0641E — G

There is wide agreement by scientists that the average temperature of the earth's surface has risen some 1.2 degrees over the past century. There is a wide range of opinion, however, as to what is actually causing this global warming and what its direct effects are.

Global warming is the change in climate generally thought to occur mostly from the release of methane and nitrous oxide and especially carbon dioxide gas from agriculture. The result is retention of heat in the atmosphere. However, there is a dominant view that surface warming is at least partly due to human activities; namely, the emissions of heat-trapping waste industrial gases like carbon dioxide from burning fossil fuels like coal, oil, and natural gas. A UN scientific panel has predicted that unless these greenhouse gas emissions are reduced, the earth's average surface temperature will rise by some 2 to 6 degrees F over the next century. The panel says that the warming would touch off widespread disruptions in climate and weather causing the melting of polar ice packs. This, in turn, would cause widespread flooding and droughts threatening ecosystems that support marine life and numerous animal species. Studies also show that wind changes due to global warming are responsible for a one-third slowing down of the Earth's spin.

Skeptics say that global warming is wholly natural, that there is no significant human influence on global climate, and any future warming may be inconsequential. They do agree, however, that fossil fuel burning should be reduced.