



Analyse de la voix d'un conducteur automobile: effets acoustiques de la fatigue

Robert Ruiz, Philippe Plantin de Hugues, Claude Legros

► To cite this version:

Robert Ruiz, Philippe Plantin de Hugues, Claude Legros. Analyse de la voix d'un conducteur automobile: effets acoustiques de la fatigue. 10ème Congrès Français d'Acoustique, Apr 2010, Lyon, France. hal-00534631

HAL Id: hal-00534631

<https://hal.science/hal-00534631>

Submitted on 10 Nov 2010

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.

10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

Analyse de la Voix d'un Conducteur Automobile : Effets Acoustiques de la Fatigue

Robert Ruiz¹, Philippe Plantin de Hugues², Claude Legros³

¹LARA, Univ. Toulouse (UTM), 5 all. A. Machado, F-31058 Toulouse cedex 9, robert.ruiz@univ-tlse2.fr

²Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile, Aéroport du Bourget, F-93352 La Bourget CEDEX

³LAUTM, Univ. Toulouse (UTM), 5 all. A. Machado, F-31058 Toulouse cedex 9

Par analogie avec la charge de travail de pilotes d'avion au cours de vols quotidiens court-courrier, la voix d'un conducteur automobile a été enregistrée périodiquement au cours d'un long trajet autoroutier d'une journée. Des effets acoustiques de la fatigue sont recherchés dans les signaux de voyelles avec la mesure de plusieurs paramètres des plans dynamique, spectral et des phases. Outre l'évolution chronologique des grandeurs, une comparaison est effectuée avec des enregistrements réalisés la veille au soir et tôt le matin de l'expérimentation. Sur la base d'un corpus de près de 400 voyelles segmentées et analysées, les résultats montrent des variations probablement en lien avec l'accumulation de la fatigue au cours de la journée (shimmer et centre de gravité spectral). Enfin, les prononciations du matin se distinguent nettement des autres en termes de fréquence fondamentale et de « digital amplitude length ». Plusieurs interprétations de ce résultat sont envisagées.

1 Expérimentation

Les effets de la fatigue sont principalement mesurés par des paramètres physiologiques (rythme cardiaque ...) et peu de recherches ont été entreprises pour en estimer les manifestations vocales [1,2]. Les applications dans le domaine aéronautique et plus généralement celui des transports de passagers en font l'intérêt. Après des mesures sur des pilotes d'avion au cours de vols court-courrier quotidiens sur le territoire français [3], il apparaît qu'un conducteur automobile peut présenter une fatigue similaire à celle des pilotes.

Pour cela un trajet autoroutier d'une journée a été envisagé et le conducteur a dû se lever tôt comme les pilotes pour quitter Paris par avion pour Toulouse, lieu de départ du trajet. Les enregistrements sont les suivants :

- Enregistrement n°0 : à 19h la veille de l'expérimentation ;
- Enregistrement n°1 : à 6h50 à l'aéroport de Paris avant le vol pour Toulouse ;
- 9h départ voiture de Toulouse ;
- Enregistrement n°2 : à 11h31 après 200 km et en roulant ;
- Enregistrement n°3 : à 12h12 après 300 km, à l'arrêt avant la pause déjeuner ;
- 13h30 fin de la pause déjeuner ;
- Enregistrement n°4 : à 13h57 après 450 km et en roulant ;
- Enregistrement n°5 : à 15h après 550 km et en roulant ;
- Enregistrement n°6 : à 16h08 après 750 km et en roulant ;
- Enregistrement n°7 : à 17h15 après 900 km et en roulant
- 17h30 arrivée et arrêt du véhicule à Toulouse.

Le conducteur est équipé pour les enregistrements (enregistreur DAT SONY TCD D10 PRO) d'un microphone de proximité monté sur un serre-tête (dispositif AKG C420) maintenant une distance constante entre la

bouche et le transducteur quelle que soit la position de la tête.

Il lit les 5 phrases suivantes, choisies pour leur similitude avec le vocabulaire aéronautique et qui sont les mêmes que celles qui ont été lues par les pilotes d'avion [3], puis les 5 premières voyelles de la langue française.

Phrase 1 : « Autorisé décollage piste 36 gauche
XRAYFOX

Phrase 2 : XRAYFOX montez au niveau 2.5.0

Phrase 3 : Je rappelle en courte finale XRAYFOX

Phrase 4 : Autorisé atterrissage piste 36 droite
XRAYFOX

Phrase 5 : Première à gauche XRAYFOX »

Près de 400 voyelles sont segmentées (une cinquantaine par enregistrement) à l'aide d'un logiciel d'édition audio puis traitées par des programmes Matlab® spécifiquement développés pour mesurer les paramètres acoustiques suivants.

2 Paramètres acoustiques

Les paramètres du plan dynamique sont mesurés à partir d'une détection automatique des pics d'amplitude principaux de la voyelle : fréquence fondamentale moyenne (Fo), jitter (J), shimmer (S).

Les paramètres du plan spectral résultent de l'estimation de l'enveloppe spectrale de la voyelle par une analyse LPC (pré-accélération = 0.98, ordre 48, fenêtre de Hamming de 512 échantillons avec recouvrement à 50%, FFT de 1024 points sur les coefficients du modèle LPC).

Il s'agit du centre de gravité (CGS), de la fréquence d'équilibre spectral, de la fréquence d'équilibre énergétique, du coefficient d'asymétrie, du coefficient d'aplatissement et de la fréquence des 4 premiers formants.

Des calculs de distance spectrale euclidienne (DS) entre les vecteurs MFCC moyens de chaque enregistrement et

pour chaque type de voyelle sont effectués (banc de 32, puis de 3, filtres Mel sur 133-5000 Hz).

Dans le plan des phases, le plus grand exposant de Lyapunov est estimé pour chaque voyelle (dimension de prolongement = 3, délai de reconstruction égal au premier zéro de la fonction d'auto corrélation, rayon de la sphère de voisinage = 2% de la dynamique du signal, nombre d'itérations = 5).

Enfin, le « Digital Amplitude Length » (DAL) [3] représente la somme des différences (en valeur absolue) de l'amplitude A entre les échantillons i successifs d'un cycle vibratoire To de la voyelle :

$$DAL_{T_0} = \sum_{i=1}^{N-1} DAL(i) = \sum_{i=1}^{N-1} |A_i - A_{i+1}| \quad (N \text{ est le nombre d'échantillons dans la période } T_0).$$

Le rapport moyen $\langle \frac{DAL_{T_0}}{T_0} \rangle$ est ensuite calculé pour l'ensemble des périodes de la voyelle.

3 Résultats

Seuls les résultats les plus significatifs sont ici présentés à savoir les variations chronologiques de Fo, $\frac{DAL_{T_0}}{T_0}$, S, CGS et DS. Pour les autres paramètres, aucune tendance ou événement particulier n'a pu être détecté.

3.1 Fréquence fondamentale et DAL/T₀

Sur les figures 1 et 2, il est remarquable de noter une baisse sensible des valeurs, mais non statistiquement significative, pour l'enregistrement de 6h50. Le paramètre DAL/T₀ accentue le phénomène.

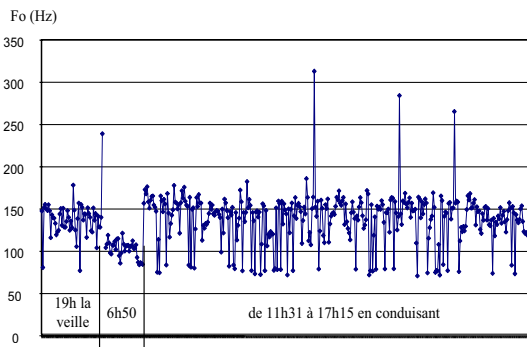


Figure 1. fréquence fondamentale moyenne par voyelle.

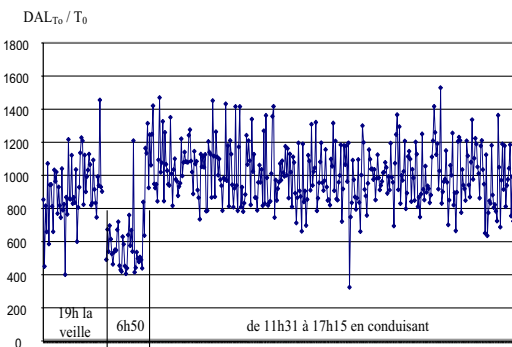


Figure 2. DAL/T₀ moyen par voyelle.

La valeur moyenne de Fo est de 137,5 Hz ($\sigma = 22,5$ Hz) à 19h la veille, de 138,3 Hz ($\sigma = 29,9$ Hz) pendant toute la journée de conduite et de 101,8 Hz ($\sigma = 10,1$ Hz) à 6h50.

Les mesures de jitter n'ont pas permis d'observer une quelconque tendance croissante ou décroissante au cours de la journée ni d'isoler une période particulière de la journée.

3.2 Shimmer (S)

$$S = \frac{1}{n-1} \sum_{i=1}^{n-1} \left| 20 \cdot \log \left(\frac{A_i}{A_{i+1}} \right) \right| \quad (\text{en dB}), \text{ où } A_i \text{ désigne}$$

ici l'amplitude du ième pic principal d'amplitude et n est le nombre de ces pics pour la voyelle.

Sur l'évolution chronologique du shimmer (Fig.3), une baisse est observée à partir de 17h15 à la fin du parcours autoroutier.

Le shimmer moyen vaut 0,76 dB ($\sigma = 0,53$ dB) à 17h15 et vaut en moyenne 1,12 dB ($\sigma = 0,93$ dB) pour toutes les voyelles de la journée de conduite. La variation n'est pas statistiquement significative.

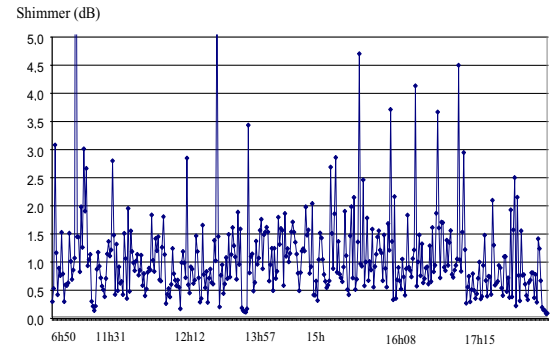


Figure 3. shimmer par voyelle.

3.3 Centre de gravité spectral (CGS)

$$C.G.S = \frac{\sum_{i=1}^N f_i \times 10^{L_i/10}}{\sum_{i=1}^N 10^{L_i/10}} \quad \text{où } N \text{ désigne le}$$

nombre de fréquences discrètes dans le spectre. Chacune d'entre elles possède un niveau L_i (dB).

La figure 4 montre l'évolution du CGS pour les 85 [i] segmentés depuis 6h50 jusqu'à 17h15.

Le calcul d'une régression linéaire montre un accroissement de 26%. Pour tous les autres types de voyelle un accroissement est également mesuré et vaut en moyenne 6,2%.

Aucun autre paramètre spectral mesuré, comme la fréquence des 4 premiers formants, n'a présenté de variation sensible et a fortiori pour chacun des 5 types de voyelles analysées.

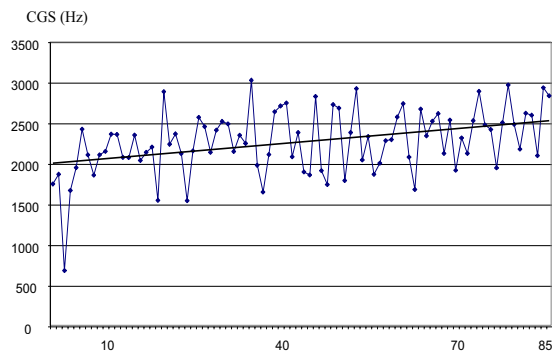


Figure 4. centre de gravité spectral des [i] par voyelle.

3.4 Distance spectrale (DS)

Que ce soit avec 32 ou 3 filtres MEL et pour chaque type de voyelle, les distances spectrales ont été calculées entre chaque enregistrement et respectivement : l'enregistrement à 19h la veille (E0), celui à 6h50 (E1), et celui à 11h31 (E2).

Avec d_{ij} qui désigne la distance entre le vecteur MFCC moyen de l'enregistrement i et celui de l'enregistrement j pour le type de voyelle considérée, les tableaux suivants présentent les résultats obtenus.

	d_{01}	d_{02}	d_{03}	d_{04}	d_{05}	d_{06}	d_{07}
[a]	7,56	4,58	4,23	3,34	4,07	3,14	2,54
[Ø]	7,64	5,29	4,43	3,79	3,54	3,75	2,90
[i]	5,54	2,96	2,50	2,68	2,77	1,92	3,03
[o]	4,49	5,23	4,47	3,60	2,97	4,43	3,75
[e]	5,51	4,49	3,48	3,36	3,68	3,47	3,28

Table 1. distance spectrale entre l'enregistrement à 19h (E0) et les suivants pour 32 coefficients MFCC.

	d_{12}	d_{13}	d_{14}	d_{15}	d_{16}	d_{17}
[a]	9,06	9,99	8,73	8,64	8,34	6,69
[Ø]	8,47	7,55	7,27	6,57	7,24	7,10
[i]	6,33	6,61	6,02	5,88	6,04	5,65
[o]	8,58	7,82	6,93	6,34	7,76	6,76
[e]	9,06	8,18	7,34	7,88	7,89	2,33

Table 2. distance spectrale entre l'enregistrement à 6h50 (E1) et les suivants pour 32 coefficients MFCC.

	d_{23}	d_{24}	d_{25}	d_{26}	d_{27}
[a]	2,77	2,16	1,15	2,4	3,48
[Ø]	2,17	2,34	2,95	2,05	4,26
[i]	1,3	1,43	1,01	1,41	3,09
[o]	2,05	2,04	2,84	1,45	1,65
[e]	1,55	2,5	1,59	1,55	3,02

Table 3. distance spectrale entre l'enregistrement à 11h31 (E2) et les suivants pour 32 coefficients MFCC.

L'examen des tableaux 1, 2 conduit aux remarques suivantes.

- que la référence du calcul soit à 19h la veille ou à 6h50 le matin, la distance avec le dernier enregistrement est toujours plus petite qu'avec le premier ;
- la tendance générale est plutôt à une décroissance des distances même si des exceptions ont lieu pour certains types de voyelles et à certaines heures de la journée.

L'examen du tableau 3 pour lequel la référence de calcul des distances est à 11h31 c'est-à-dire pour le premier enregistrement en roulant montre que c'est la distance avec le dernier enregistrement (d_{27}) qui est plus grande que la première (d_{23}) (sauf pour les [o]). Le sens de variation est donc contraire à celui observé pour les deux autres références (Tableaux 1 et 2).

Il est difficile de conclure à une tendance croissante ou décroissante dans ce cas car les sens de variation changent d'une distance à l'autre et d'une voyelle à l'autre.

Les calculs de distance avec des vecteurs MFCC à 3 coefficients conduisent aux mêmes observations qu'avec les vecteurs à 32 coefficients MFCC.

3.5 Exposant de Lyapunov (λ)

Sur la figure 5 sont observées les variations du plus grand exposant de Lyapunov au cours de la journée de conduite.

Outre la grande dispersion des valeurs, la droite de régression ne montre pas de tendance particulière. Aucune période de la journée ne peut être distinguée des autres.

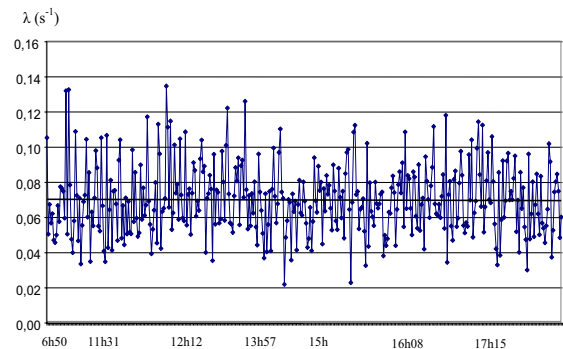


Figure 5. Exposant de Lyapunov par voyelle.

4 Analyse et interprétation des résultats

Une hypothèse pourrait consister à poser que l'accroissement attendu de la fatigue provoquerait une décroissance (pas nécessairement linéaire) des valeurs numériques d'un paramètre acoustique.

De ce point de vue, seules les distances spectrales semblent vérifier le postulat. A condition toutefois qu'une référence de « repos » soit prise en dehors de la conduite automobile.

Dans une certaine mesure, le shimmer peut satisfaire à l'hypothèse car bien qu'il ne décroisse pas dans la journée, il présente une baisse nette à l'arrivée du parcours.

Le centre de gravité spectral semble quant à lui augmenter au cours de la journée. Il traduit une perte du poids des basses fréquences dans le spectre (ou une augmentation de celui des hautes fréquences). La forme des spectres change sans doute provoquant également les phénomènes observés pour les distances spectrales.

La fatigue attendue ne semble pas affecter l'aspect mélodique des voyelles. Ni la fréquence fondamentale, ni le jitter ne sont modifiés pendant la journée de conduite. En revanche, une chute manifeste existe à 6h50. Elle pourrait être imputable à un phénomène de somnolence compte tenu de l'heure matinale accompagné de prononciations avec une voix encore peu utilisée.

La représentation des vecteurs MFCC moyens (à 3 coefficients) dans un espace à 3 dimensions isole nettement ces prononciations de toutes les autres pour tous les types de voyelles. Le phénomène est donc également confirmé dans le plan spectral.

L'enregistrement à 6h50 ne constituerait donc pas une référence de l'état de « repos » du locuteur et il serait préférable de considérer celui-ci à 19h la veille de l'expérimentation.

Enfin, il a été procédé à l'analyse de voyelles prononcées dans des mots hors du contexte de vocabulaire contrôlé imposé par la lecture des 5 phrases du protocole. Le conducteur indiquait en effet le kilométrage parcouru et le lieu de l'enregistrement.

Il s'avère que la faible valeur moyenne de la fréquence fondamentale à 6h50 est confirmée, de même que les faibles valeurs du shimmer en fin de journée.

Le choix du vocabulaire définit pour l'expérience ne semble pas affecter les conclusions générales.

5 Conclusion

Pour cette étude exploratoire, un seul locuteur a été testé ce qui n'autorise pas de généralisation des résultats.

Par ailleurs, il se prépare à s'exprimer lors des enregistrements et peut se « ressaisir » masquant ainsi des effets vocaux de la fatigue qui pourrait peut-être se manifester si l'appareillage et le protocole étaient moins invasifs.

Une quantification plus précise de l'état de « repos » s'impose en effectuant sans doute un plus grand nombre de mesures à plusieurs jours d'intervalle et à plusieurs heures de la journée.

Bien que les résultats conduisent à cette perspective, il faut considérer que pour des applications futures les échantillons de parole « au repos » peuvent faire défaut comme c'est le cas pour les pilotes d'avion.

En dépit de ces remarques critiques, les résultats obtenus montrent que des modifications vocales de la fatigue provoquée par une longue période de conduite automobile existent.

Remerciements

Les auteurs remercient le Bureau d'Enquêtes et d'Analyses qui supporte la recherche et qui a aimablement prêté le véhicule pour l'expérimentation. Un remerciement particulier est adressé au conducteur.

Références

- [1] Shiomi K., "Fatigue and Drowsiness Predictor for Pilots and Air Traffic Controllers", Proc. of 45th Annual ATCA Conference, (Oct. 2000).
- [2] Shiomi K., Sato K., Sawa M., Suzuki A., "Experimental Results of Measuring Human Fatigue by Utilizing Uttered Voice Processing", Proc. of IEEE-SMC P557, Singapore (2008).
- [3] Ruiz R., Plantin de Hugues P., Legros C., "Advanced Voice Analysis of Pilots to Detect Fatigue and Sleep Inertia" *Acta Acustica united with Acustica* (accepté pour publication) (2010).