

CFA/VISHNO 2016

Nouveau paradigme pour l'étude du temps de traitement de sources sonores par le système auditif : RASP, Rapid Audio Sequential Presentation

V. Isnard^a, I. Viaud Delmon^a et C. Suied^b

^aIRCAM, 1, Place Stravinsky, 75004 Paris, France

^bInstitut de Recherche Biomédicale des Armées (IRBA), BP 73, 91220 Brétigny-Sur-Orge, France
vincent.isnard@live.fr



LE MANS

CFA2016/574**Nouveau paradigme pour l'étude du temps de traitement de sources sonores par le système auditif : RASP, Rapid Audio Sequential Presentation**V. Isnard^a, I. Viaud Delmon^a et C. Suied^b^aIRCAM, 1, Place Stravinsky, 75004 Paris, France^bInstitut de Recherche Biomédicale des Armées (IRBA), BP 73, 91220 Brétigny-Sur-Orge, France
vincent.isnard@live.fr

Reconnaître des sources sonores présentes dans notre environnement quotidien semble être une tâche facile pour le système auditif humain. Des preuves expérimentales de plus en plus nombreuses montrent que cette reconnaissance est robuste à des dégradations parfois très fortes du signal acoustique. En particulier, dans le domaine temporel, il a été montré que des sons extrêmement courts (quelques millisecondes) pouvaient être reconnus par des auditeurs. Ici, nous présentons un nouveau paradigme appelé RASP (pour Rapid Audio Sequential Presentation) qui permet de s'intéresser au temps mis par le système auditif pour traiter de courts extraits de sons. Le corpus de sons utilisé était composé de sons de voix chantées et d'instruments de musique, à 12 hauteurs différentes (entre La₂ et Sol#₃), et égalisés en sonie. Des séquences de sons courts (16 et 32 ms, courts extraits issus de notre corpus) ont été présentées aux auditeurs. Ces séquences étaient composées d'une cible à reconnaître au milieu d'un flux de distracteurs (voix cible et instruments distracteurs, et vice-versa). Pour toutes les séquences soit de durée fixe (500 ms), soit à nombre de sons fixe (7 sons), la performance restait meilleure que le hasard pour des taux allant jusqu'à 30 Hz. Il n'y avait aucune différence entre des séquences pour lesquelles la hauteur choisie des sons était fixe, ou aléatoire. Cela montre que le résultat ne peut pas être expliqué par un phénomène de masquage, qui aurait impliqué une meilleure performance pour la condition hauteur aléatoire. Enfin, une asymétrie dans les résultats en faveur de la voix a été mise en évidence dans toutes ces conditions expérimentales : une voix parmi des instruments est systématiquement mieux reconnue qu'un instrument parmi des voix. Les sons utilisés dans les deux types de séquences étant les mêmes, cet effet ne peut être expliqué par des différences acoustiques.