

Ségrégation séquentielle et structure fine temporelle

Nicolas Grimault¹, Etienne Gaudrain²

¹UMR CNRS 5020 50 av. T. Garnier 69366 Lyon, nicolas.grimault@olfac.univ-lyon1.fr

²MRC Cognition and Brain Sciences Unit 15 Chaucer Road Cambridge, CB2 7EF, UK, etienne.gaudrain@mrc-cbu.cam.ac.uk

Utilisant une méthodologie spécifique et les signaux introduits par Hopkins et Moore [Hopkins & Moore (2007) J. Acoust. Soc. Am. 122, 1055-1068] consistant en des sons complexes dont les harmoniques ont été décalé d'une fraction de F0, de nombreux travaux se concentrent actuellement sur le rôle comparé de la structure fine et de l'enveloppe temporelle pour la perception de la hauteur. Le système auditif serait ainsi capable de percevoir l'enveloppe du signal (i.e. les fluctuations temporelles lentes,) et la structure fine du signal (i.e. la latence entre les pics de plus grande amplitude dans la forme d'onde du signal). La ségrégation auditive, basée sur des différences de hauteur et impliquée dans la perception de la parole en situation concurrentielle d'écoute, pourrait donc elle aussi tirer avantageusement profit d'indices de structure fine. Nous proposons ici de revisiter un certain nombre de données expérimentales adressant cette question ainsi que de confronter une tâche de ségrégation auditive aux signaux de Hopkins et Moore afin de déterminer plus précisément le rôle que pourrait revêtir la structure fine pour la ségrégation auditive.

1 Introduction

La ségrégation séquentielle réfère à notre aptitude à interpréter une séquence sonore *ABABAB...* constituée de sons *A* et *B* comme étant issue d'une ou au contraire de deux sources sonores distinctes (van Noorden, 1975). Dans une revue de littérature, Moore et Gockel, (2002) suggèrent que n'importe quelle différence acoustique entre *A* et *B* se traduisant pas une différence perceptive saillante peut potentiellement permettre la ségrégation perceptive de la séquence *ABABAB...* en deux séquences *A-A-A...* et *B-B-B...*

Une fois établi le rôle de la hauteur pour la ségrégation auditive, une littérature relativement abondante (e.g. Vliegen et Moore, 1999 ; Vliegen et al, 1999 ; Grimault et al, 2000 ; Roberts et al, 2002) a mis en évidence l'effet des indices temporels de hauteur pour cette mise en flux en utilisant pour *A* et *B* des sons complexes harmoniques non résolus par le système auditif (i.e. des sons dont les différences spectrales ne pouvaient être perçues). Ces études ne distinguaient pas alors entre les indices temporels d'enveloppe et les indices temporels de structure fine qui tous deux pouvaient contribuer à la ségrégation auditive observée.

Grimault et al. (2002), en utilisant pour *A* et *B* des bruits larges bandes modulés en amplitudes à des cadences différentes, ont, pour la première fois, mis en évidence que des indices temporels d'enveloppe (fluctuations entre 100 Hz et 300 Hz) pouvaient conduire à une ségrégation auditive en l'absence d'indices de structure fine.

A contrario, la contribution des indices de structure fine pour la ségrégation auditive reste à ce jour largement indéterminée. Grimault et Gaudrain (2008) ont réalisé une étude où une des conditions expérimentales consistait à mesurer la ségrégation auditive entre des sons complexes *A* et *B* non résolus de fréquences fondamentales respectives 100 Hz et 210 Hz et dont les indices de structure fine étaient dégradés (ou pas) en introduisant (ou pas) une légère modulation de fréquence fondamentale sur *A* et *B*.

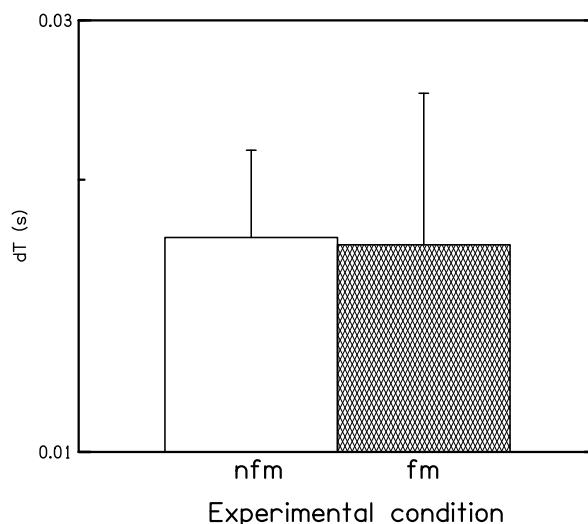


Figure 1 : Seuils de détection d'une variation de rythme dans des conditions de structure fine non dégradée (*nfm*) ou dégradée (*fm*). Ce seuil de détection est réputé être corrélé positivement à la ségrégation auditive.

Les résultats de cette étude, partiellement reproduits sur la figure 1, ne montrent pas de différence significative entre les deux conditions. Ces résultats ne permettent donc pas de mettre en évidence une quelconque utilité de la structure fine pour la ségrégation séquentielle.

2 Expérience

2.1 Hypothèses et objectifs

Il existe plusieurs limitations à l'expérience réalisée par Grimault et Gaudrain (2008). Premièrement, dans chacune des conditions expérimentales, les sons complexes *A* et *B* différaient en termes de fréquence fondamentale. Les indices de structure fine étaient donc toujours associés à des indices d'enveloppe, même si ces derniers étaient réduits. Deuxièmement, la modulation de fréquence utilisée pour

dégrader les indices de structure fine n'était peut être pas optimum et des indices de structure fine ont pu être utilisés même en condition dégradée. Troisièmement, et plus probablement, la bande passante utilisée (3500-4500 Hz) pour rendre les sons complexes non résolus ne permettait pas de garantir l'existence d'indices de structure fine perceptibles. L'absence d'effet dans cette étude pourrait donc s'expliquer par la dégradation d'indices de structure fine de toute façon non perceptibles.

Cette nouvelle étude propose (1) de tester le rôle intrinsèque de la structure fine pour la ségrégation auditive dans des conditions expérimentales optimum ou la structure fine était très probablement audible ; et (2) de déterminer l'existence éventuelle d'une interaction entre structure fine et enveloppe pour cette ségrégation.

2.2 Sujets

4 sujets ont participé à cette expérience. Tous, à l'exception du second auteur (S1), étaient indemnisés pour leur participation et ont signé un consentement de participation. Une audiométrie tonale réalisée chez chacun de sujets garantie une absence de perte auditive (seuils auditifs inférieurs à 15 dB HL entre 250Hz et 6kHz). Le protocole expérimental a par ailleurs été formellement autorisé par un comité d'éthique (CPP sud-est II, autorisation N°06035).

2.3 Signaux

Quatre sons complexes harmoniques notés *A*, *B*, *C* et *D* sont utilisés dans cette expérience. Tous durent 150ms y inclus des rampes temporelles *on/off* de 10ms (forme cosinus). Tous sont filtrés de façon à obtenir un plateau entre 450 Hz et 650 Hz et des jupes décroissants à 30 dB/oct. *A* et *C* sont des sons complexes harmoniques de fréquences fondamentales 50 Hz et 76 Hz. *B* et *D* sont des sons complexes harmoniques de fréquences fondamentales 50 Hz et 76 Hz dont les harmoniques ont été décalées vers les hautes fréquences de 25 Hz et 38 Hz respectivement (soit $0,5 \times F_0$) à la manière de Moore *et al* (2009) afin d'introduire entre *A* et *B* et entre *C* et *D* une différence de structure fine temporelle tout en préservant l'enveloppe temporelle. Les phases des composantes des sons complexes sont aléatoires. Tous les sons complexes sont présentés à un niveau fixe confortable dans les oreilles gauche et droite simultanément. Un bruit de type TEN (Moore *et al*, 2000) dont le niveau spectral à 1 kHz est ajusté à 20 dB sous le niveau spectral de la plus forte des harmoniques était présent en continue au cours de toute l'expérience afin de prévenir la perception de sons de combinaisons. Les sons complexes étant par ailleurs non résolus, aucun indice spectral ne peut conduire à discriminer entre *A*, *B*, *C* et *D*.

2.4 Protocoles expérimentaux

Deux tests ont été réalisés afin de garantir que les indices de structure fine étaient bien perçus lorsque la ségrégation était mesurée.

Le *test 1* consistait à mesurer la discriminabilité (*d'*) entre le son *A* et les sons *B*, *C* et *D*. Ce test consistait en 10 répétitions de 3 blocs présentés dans un ordre aléatoire. Chaque bloc était spécifiquement dédié à la comparaison entre *A* et un des son *B*, *C* ou *D*. Au cours d'un bloc, des paires de sons (par exemple *A/B*, *B/A*, ou *A/A*) étaient

présentés. Le sujet, qui était informé de la condition testée dans le bloc, devait alors indiquer si les sons composant la paire étaient identiques ou différents. Un feedback était communiqué au sujet après chaque réponse. Chaque bloc était composé de 10 paires identiques et de 10 paires différentes. Les résultats de ce test, exprimés en terme de *d'*, sont calculés à partir des 300 paires identiques et des 300 paires différentes pour chaque condition (*A/B*, *A/C* et *A/D*). Ce test a été effectué deux fois de suite chez chacun des sujets afin de saturer l'entraînement, et seules les valeurs collectées la seconde fois sont rapportées ici.

Des seuils de détection d'une irrégularité rythmique (Roberts *et al.*, 2002) ont ensuite été mesurés au cours du *test 2* afin de révéler un possible effet des indices temporels sur la ségrégation séquentielle. Les séquences présentées étaient toutes constituées d'un même motif (*A-X*) se répétant (*A-X-A-X-A-X-A-X...*). Quatre motifs différents, correspondant aux 4 conditions de ce test, ont été utilisés : (1) *A-A*, (2) *A-B*, (3) *A-C* et (4) *A-D*. En condition 1 (*A-A*), tous les sons sont identiques. En condition 2 (*A-B*), une différence de structure fine est introduite. En condition 3 (*A-C*), des différences d'enveloppe et de structure fine coexistent. La condition 4 (*A-D*) vient ajouter des indices de structure fine supplémentaires aux indices temporels existant dans la condition 3.

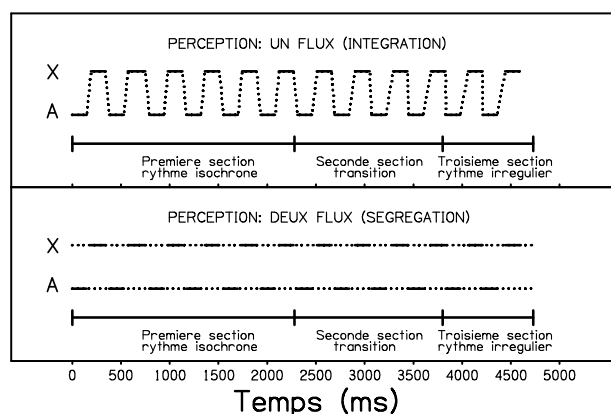


Figure 2 : Séquence *A-X-A-X...* utilisée dans le *test 2*. La séquence, isochrone dans la première section devient irrégulière dans la troisième section en décalant progressivement les sons *X* lors d'une section de transition. Cette irrégularité rythmique, à la fin de la séquence, est plus facilement détectable lorsque un seul flux auditif est perçu (en haut) que lorsque deux flux auditifs sont perçus (en bas).

Les seuils de détection de l'irrégularité rythmique étaient estimés au moyen d'une procédure à choix forcé à deux intervalles. Le délai temporel introduit pour altérer la régularité rythmique variait suivant une règle adaptative 3-down/1-up. La mesure du seuil prenait fin après 6 changements de direction. Le seuil est défini comme étant la moyenne géométrique des délais temporels lors des 4 derniers changements de direction. Chaque essai consistait en la présentation, dans un ordre aléatoire, d'une séquence isochrone (tous les intervalles entre *A* et *X* et entre *X* et *A* sont égaux à 40 ms) et d'une séquence irrégulière (un délai temporel est introduit et le rythme de présentation du son *X* varie au cours de la présentation de la séquence : détails figure 2). À chaque essai, le sujet devait identifier le stimulus présentant une irrégularité rythmique. Le délai

temporel variait géométriquement au cours de la procédure adaptative par pas de $2^{1/2}$ (pas réduit à $2^{1/4}$ après 2 changements de direction). La valeur initiale du décalage temporel était fixée à 10 ms. La valeur maximale du décalage était de 40 ms afin d'éviter toute superposition de *A* et de *X*. Une seconde de silence séparait les deux intervalles à comparer.

2.5 Matériel

Les sons complexes étaient présentés via une carte son de type RME (HDSP, Multiface II). Le bruit, préalablement enregistré sur CD, était présenté via un lecteur CD audio (Technics SL PG590). Les sorties de ces modules étaient indépendamment atténuées (TDT, PA5) puis mixés (TDT, SA1). La sortie de ce module de mixage était dupliquée et connectée aux entrées gauche et droite d'un pré-ampli casque (TDT, HB7) connecté à un casque de type Sennheiser HD 250 linear II. Les sujets étaient confortablement installés dans une cabine insonorisée devant un clavier, une souris et un écran.

3 Résultats

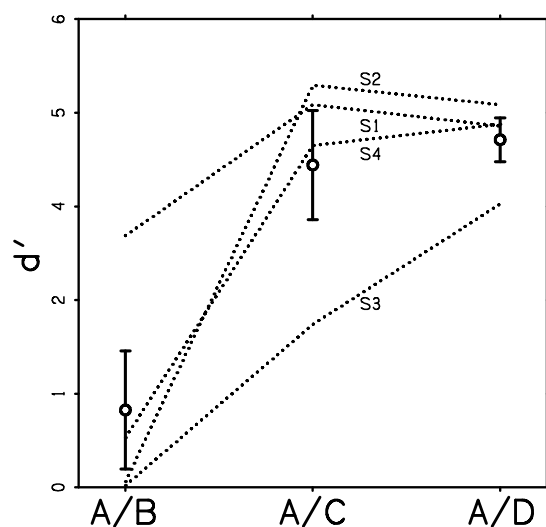


Figure 3 : Performances de discrimination (d') entre les sons *A* et *B*, *A* et *C*, et *A* et *D*. Les résultats individuels sont tracés en lignes pointillées. Les sujets sont identifiés (S1, S2, S3, S4). Les résultats moyens ainsi que l'erreur standard sont reproduits sur ce graphique.

Les résultats de discrimination du *test 1* sont tracés sur la figure 3. Une ANOVA à mesures répétées appliquée à ces données révèle un effet significatif de la condition expérimentale [$F(2,6) = 16,44$, $p=0,0037$]. Des tests post-hoc (Tukey HSD) indiquent que la condition *A/B* est différente de la condition *A/C* [$p<0,000$] et *A/D* [$p<0,000$]. Les conditions *A/C* et *A/D* ne sont pas significativement différentes entre elles [$p=0,878$]. Une analyse des résultats individuels montre par ailleurs que deux des sujets (S1 : $p<0,0001$; S4 : $p=0,007$) discriminent au dessus de la chance entre *A* et *B* au contraire des deux autres sujets (S2 : $p=0,60$; S3 : $p=0,85$).

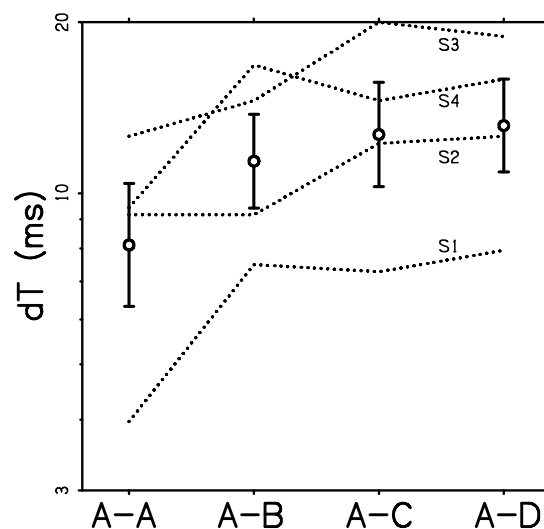


Figure 4 : Seuils de détection d'un délai temporel (dT) exprimés en ms pour chacun des motifs (*A-A*, *A-B*, *A-C* et *A-D*). Les résultats individuels sont tracés en lignes pointillées. Les sujets sont identifiés (S1, S2, S3, S4). Les moyennes géométriques ainsi que l'erreur standard logarithmique sont reproduits sur ce graphique.

Les résultats de détection du *test 2* sont tracés sur la figure 4. Une ANOVA à mesures répétées appliquée au logarithme de ces données révèle un effet significatif de la condition expérimentale [$F(3,9) = 7,25$, $p=0,009$]. Des tests post-hoc (Tukey HSD) indiquent que la condition *A-A* est différente de toutes les autres conditions expérimentales (*A-B* : $p=0,039$; *A-C* : $p<0,001$; *A-D* : $p<0,001$). Les autres conditions (*A-B*, *A-C*, *A-D*) ne sont pas significativement différentes les unes des autres.

4 Discussion

Les performances de discrimination entre des sons qui se distinguent par leur seule structure fine (sons *A* et *B*, figure 3) sont inférieurs au d' moyen de 3.7 (environ) trouvé par Moore *et al* (2009) dans la même condition expérimentale. Cette différence de d' peut sans doute s'expliquer par un entraînement moindre de nos sujets (1h versus 3h) et par le fait que nos sujets n'avaient pas de pratique musicale particulière contrairement aux sujets de Moore *et al* (2009). Toutefois, en moyenne, les indices de structure fine étaient perceptibles et ont permis (au moins à certains sujets) de discriminer entre *A* et *B*.

La dégradation significative des seuils de détection entre la condition *A-A* et la condition *A-B* qui est manifeste chez trois de nos sujets (sur 4) indique pour la première fois un effet intrinsèque de la structure fine sur la ségrégation séquentielle irrépressible.

Les indices d'enveloppe présents dans les autres conditions expérimentales ne semblent pas contribuer à la ségrégation auditive. Deux hypothèses peuvent être avancées pour expliquer l'absence d'effet de l'enveloppe temporelle. Premièrement, il est possible que la ségrégation auditive ait été saturée par les seuls indices de structure fine et que l'effet de l'enveloppe sur cette ségrégation existe mais ne soit pas observable dans cette expérience. Deuxièmement, le nombre trop restreint de sujets peut être seul responsable de cette absence d'effet. En effet, l'observation détaillée des résultats individuels semblent

indiquer que les sujets ne discriminant pas la structure fine (S2 et S3) puissent utiliser l'enveloppe temporelle pour la ségrégation auditive (cf. saut qualitatif pour ces sujets dans la tâche de détection entre les conditions *A-B* et *A-C*). Le nombre de sujets participant à cette étude devra donc être augmenté afin de pouvoir conclure sur ce point.

5 Conclusion

Malgré un nombre restreint de sujets qui limite la portée des résultats, cette étude apporte pour la première fois des éléments convainquant quant à l'utilité de la structure fine temporelle sur la ségrégation auditive irrépressible.

Remerciements

Cette étude a bénéficié d'un soutien de l'Agence Nationale de Recherche (ANR Multistap).

Références

- Grimault N., Bacon S.P., Micheyl C. (2002) "Auditory stream segregation on the basis of amplitude-modulation rate", *J. Acoust. Soc. Am* 111, 1340-1348.
- Grimault, N., Gaudrain, E. (2008) Temporal fine structure cues in auditory stream segregation of complex tones, *J. Acoust Soc. Am.* 123, 3050.
- Grimault, N., Micheyl, C., Carlyon, R. P., Arthaud, P., and Collet, L. (2000) "Influence of peripheral resolvability on the perceptual segregation of harmonic complex tones differing in fundamental frequency," *J. Acoust. Soc. Am.* 108, 263–271.
- Hopkins, K., Moore B.C.J. (2007) "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information," *J. Acoust. Soc. Am.* 122, 1055-1068
- Moore, B. C. J., and Gockel, H. (2002) "Factors influencing sequential stream segregation," *Acta Acust. Acust.* 88, 320–333.
- Moore, B. C. J., Hopkins, K., Cuthbertson, S. (2009) "Discrimination of complex tones with unresolved components using temporal fine structure information," *J. Acoust. Soc. Am.*, 125, 3214-3222.
- Moore, B. C. J., Huss, M., Vickers, D.A., Glasberg, B.R., Alcantara, J.I. (2000) "A test for the diagnosis of dead regions in the cochlea," *Br. J. Audiol.*, 34, 205-224.
- van Noorden, L. P. A. S. (1975) "Temporal coherence in the perception of tone sequences," doctoral thesis, Eindhoven University of Technology, The Netherlands.
- Roberts, B., Glasberg, B. R., Moore, B.C.J. (2002) Primitive stream segregation of tone sequences without differences in fundamental frequency or passband *J. Acoust. Soc. Am.* 112, 2074-2085.
- Vliegen, J., and Oxenham, A. J. (1999) "Sequential stream segregation in the absence of spectral cues," *J. Acoust. Soc. Am.* 105, 339–346.

- Vliegen, J., Moore, B. C. J., and Oxenham, A. J. (1999) "The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task," *J. Acoust. Soc. Am.* 106, 938–945.