

Diphone-Studio

Manuel de Resan

Première édition, novembre 2001



Copyright© 2001, Ircam. Tous droits réservés.

Ce manuel ne doit pas être copié, ni en entier ni partiellement, sans le consentement écrit de l'Ircam. Il a été rédigé par Romain Mules. Il a été produit sous la responsabilité éditoriale de Marc Battier, Département de la Valorisation, Ircam.

Diphone

Conception : Xavier Rodet

Programmation : Macintosh Adrien Lefèvre

Programmation Unix : Thierry Galas, Philippe Depalle

Etudiants-stagiaires : Guillaume Bouriez, Marteen de Boer, Xavier Hosxe, Gilbert Nouno

SuperVP

Conception : Philippe Depalle

Programmation : Philippe Depalle, Gilles Poirot, Chris Rogers, Jean Carrive

Pm

Conception : Xavier Rodet et Guillermo Garcia

Programmation : Guillermo Garcia, Diemo Schwarz

Chant

Conception : Xavier Rodet, Yves Potard

Programmation : Yves Potard, Gerhart Eckel, Francisco Iovino, Dominique Virolle

Sdif

Conception : Xavier Rodet

Programmation : Dominique Virolle, Diemo Schwarz

ModRes

Conception: Xavier Rodet

Programmation : Pierre-Francois Baisnée, Francesc Marti

Ce manuel Resan est destiné à la version de Diphone_Studio version 4r2n (novembre 2001)

Apple Macintosh est une marque déposée de Apple Computer, Inc.

Diphone, SuperVP, ModRes, Sdif et Pm sont des marques de l'Ircam.

Novembre 2001.

Ircam Centre Pompidou
1, place Igor-Stravinsky
F-75004 Paris
Tel. 01 44 78 49 62
Fax 01 44 78 15 40
E-mail ircam-doc@ircam.fr
<http://www.ircam.fr/forumnet>

Groupe d'utilisateurs Ircam

L'utilisation de ce programme et de sa documentation est strictement réservée aux membres des groupes d'utilisateurs de logiciels Ircam. Pour tout renseignement supplémentaire, contactez :

Département de la Valorisation
Ircam
1, Place Stravinsky
F-75004 Paris
France

Courrier électronique: bousac@ircam.fr

Veillez faire parvenir tout commentaire ou suggestion à :

M. Battier
Département de la Valorisation
Ircam
1, Place Stravinsky, F-75004 Paris, France

Courrier électronique : bam@ircam.fr

<http://www.ircam.fr/forumnet>

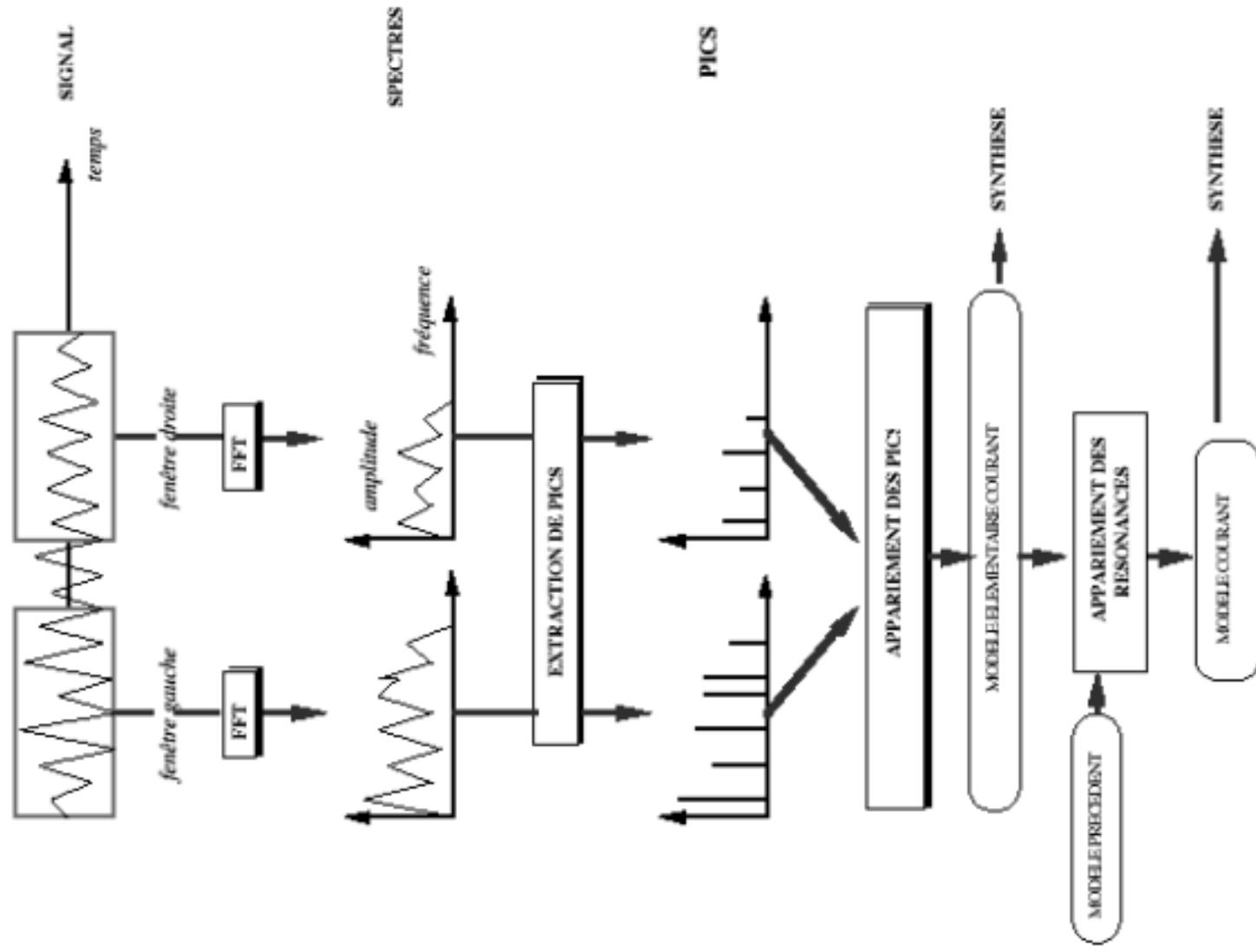
Resan

Le présent document présente les points importants dans la gestion des paramètres d'analyse du logiciel Resan, la création de presets. La maîtrise de Diphone et des différents paramètres de resynthèse de chant (fundamental, frequency, bandwidth...) est recommandée pour une mise en application des points soulevés.

1. Paramètres d'analyse du logiciel Resan.
2. Trois exemples sonores à travers différentes possibilités de paramétrage.
3. Quelques remarques additionnelles.
4. Documents annexes.

1. PARAMETRES D'ANALYSE DU LOGICIEL RESAN

Out Typing
Estimation spectrale
extraction des pics
appariement des pics
appariement des résonances



Out Typing

Out Typing										
Format Sdif	-Z	Yes	-	-	-	-	-	-	-	-
Format Asoll	-I	No	-	-	-	-	-	-	-	-
Format Max	-J	No	-	-	-	-	-	-	-	-

Pour chaque étape d'analyse, le logiciel Resan crée un ou plusieurs fichiers dans trois formats différents (sdif, ASCII et Max). Sauf utilisation spécifique des données collectées, changer les paramètres de cette section ne présente aucune nécessité.

Le format max peut par exemple servir à l'utilisation de l'objet resonators~ (CNMAT) dans Max/MSP. Le format ASCII peut donner des informations intéressantes à l'utilisateur curieux du détail de son analyse en format texte sur n'importe quel éditeur - Bbedit, Simpletext... On peut d'ailleurs en voir des extraits à la fin de ce document.

Le format Sdif est utilisé par Resan pour le traitement des données et ne doit donc pas être désactivé.

Note : La première colonne - celle dont le gris est plus dense que les autres - que l'on peut remarquer sur le tableau ci-dessus et sur tous les tableaux qui suivront - permet d'indiquer la valeur par défaut de toute la ligne concernée.

Pour exemple, sur le tableau out typing ci-dessus, le format sdif a par défaut la valeur « yes ». Tous les champs de la même ligne ayant pour valeur « - » ont la valeur de la ligne par défaut. Pour changer la valeur d'un champ, il suffit donc de cliquer sur le champ en question et d'en changer la valeur. Pour changer la valeur de la ligne entière, changer la valeur par défaut de la ligne et laisser dans chaque champs la valeur « - ».

Estimation spectrale

Spectral Estimation											
First Window Beginning	[sec]	-t	0.0	-	-	-	-	-	-	-	-
First Window Beginning Relative	[%]	-t	0.0	-	-	-	-	-	-	-	-
Second Window Beginning	[sec]	-T	0.0	-	-	-	-	-	-	-	-
Second Window Beginning Relative	[%]	-T	0.0	-	-	-	-	-	-	-	-
Window Duration	[sec]	-w	0.0	-	-	-	-	-	-	-	-
Window Duration Relative	[%]	-w	0.0	-	-	-	-	-	-	-	-
Window Type		-e	BH-92dB	-	-	-	-	-	-	-	-
FFT Duration	[sec]	-f	0.0	-	-	-	-	-	-	-	-

On peut voir dans cette section de l'analyse deux paramètres -t, -T (first & second window beginning) et -w (window duration). Il est en effet possible de spécifier ces temps et durées en secondes ou en pourcentage (l'un ou l'autre, mais pas les deux en même temps). L'avantage de l'option relative (%) est que les paramètres de l'analyse pourront ainsi s'appliquer à n'importe quel son, quelle qu'en soit la durée. Tandis que l'option en secondes permet une grande précision pour un son en particulier, que l'on voudrait traiter d'une façon vraiment spécifique. On verra dans les pages suivantes que l'option relative est plus sûre d'utilisation; il est en effet plus simple de la déboguer en cas de problème.

First Window beginning (ou) / relative (-t)

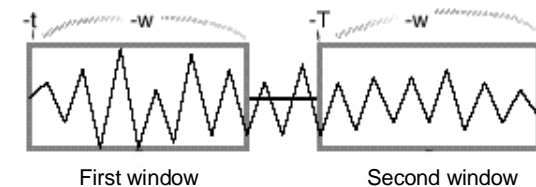
Sélectionne dans le son le temps de début de fenêtrage I pour chaque étape d'analyse.

(au choix en secondes ou en pourcentage de durée du son).

Second Window beginning (ou) / relative (-T)

Sélectionne dans le son le temps de début de fenêtrage II pour chaque itération.

Par défaut, cette valeur est égale à [First window beginning + Window duration].(au choix en seconde ou en pourcentage de durée du son).



Window Duration (ou) / relative (-w)

Sélectionne la durée de la sélection dans le son à analyser (dont le début est défini par les deux premiers paramètres). On peut donc déduire de First Window beginning et Window duration le début et la fin de la sélection à analyser dans le son.

(au choix en seconde ou en pourcentage de durée du son).

Window Type (-c)

Sélectionne le type de fenêtrage que l'on veut appliquer aux sélections définies plus haut. (pour plus de détails, se référer à la documentation du logiciel Audiosculpt). C'est généralement l'option BH -92 dB qui donne les meilleurs résultats.

Fft Duration (-f)

Les petites window size permettent d'extraire les résonances très amorties de l'attaque (généralement aïgues) et les résonances peu amorties les plus fortes; on obtient ainsi une meilleure résolution temporelle.

Les grandes fenêtres, quant à elles, permettent d'affiner l'estimation des paramètres des résonances peu amorties (les plus fortes et aussi les plus faibles), ainsi que de séparer les résonances très proches responsables de battements; on obtient ainsi généralement une meilleure résolution spectrale dans les fréquences moyennes-basses.

Les fenêtres de plus de 1 seconde ne sont utiles que pour des sons très résonants, pour séparer des résonances très proches.

Il est évident que les paramètres doivent être adaptés aux caractéristiques de chaque son.

(Pour plus de détails au sujet de la fft, se référer aux différentes documentations de Audiosculpt).

L'utilisation de ce paramètre dans Resan est particulièrement délicat et peut être à l'origine de certains problèmes dans l'analyse du son. Il est donc recommandé de ne pas toucher à sa valeur par défaut (plus proche puissance de 2 supérieure à la taille de la fenêtre). Si l'on désire malgré tout changer cette valeur, il faut ne pas oublier qu'elle être spécifiée en secondes.

ATTENTION :

1. Si le temps de début des fenêtres n'est pas spécifié, la première des deux débutera au temps 0, et la seconde à $-t$ (0) + $-w$ (c'est-à-dire que le temps de début de la seconde fenêtre coïncidera avec le temps de fin de la première).

2. Une erreur peut survenir si les paramètres ne sont pas correctement réglés :

fenetre droite:

fin du son atteinte avant lecture des echantillons demandes

- Kernel <modres> return error 1

Ce type de message d'erreur provient du fait que $[(2 * -w) + -t] > 100$.

Il faut donc toujours veiller que les durées des deux fenêtres additionnées au first window beginning n'excèdent pas 100 %. Sans quoi le logiciel ne peut achever son analyse sur la fenêtre droite (la seconde).

Exemple : si $-t = 15$ %, et $-w = 45$ %, alors l'analyse sera effectuée de la façon suivante :

Une première fenêtre, de 15 à 60 % de la taille du son, sera analysée. Puis une seconde, cette fois de 60 % (-T n'a pas été spécifié) à 105 %. Il est évident que le logiciel ne peut analyser au-delà de 100 % de la taille du son. Donc « fin du son atteinte avant lecture des échantillons demandés ».

Cette première section de l'analyse a permis d'observer le son pour en extraire des fréquences. Ces fréquences seront traitées dans la prochaine étape. L'enjeu sera alors de déterminer lesquelles de ces fréquences peuvent être considérées comme des pics.

Cette première section sera répétées sur des parties différentes du son de une à huit fois, selon les paramètres sélectionnés (chaque section d'analyse contenant huit colonnes). C'est seulement grâce à ces différentes étapes que le son pourra être balayé et analysé dans son intégralité. C'est aussi grâce à elles que les différents résultats d'analyse pourront être comparés, puis appariés.

Extraction des pics

Pics Extraction										
Extract Smaller	-P	All	-	-	-	-	-	-	-	-
Pics Width	-L	0.0	-	-	-	-	-	-	-	-
Pics Amplitude	[dB]	-I	2.0	-	-	-	-	-	-	-
Min Amplitude 1st Window	[dB]	-a	-80.0	-	-	-	-	-	-	-
Min Amplitude 2nd Window	[dB]	-A	-96.0	-	-	-	-	-	-	-

Nous sommes maintenant à la recherche des pics dans l'analyse que nous avons entamée. Il s'agit en fait, dans l'analyse précédemment obtenue, de séparer les pics - qui représentent les résonances importantes du son, des autres fréquences. Plusieurs paramètres doivent être renseignés; et plusieurs conditions remplies pour qu'une fréquence puisse être finalement considéré comme un pic.

Pics Width (-L)

Détermine le nombre de fréquences dont l'amplitude doit être inférieure à celle d'une autre fréquence pour que celle-ci soit considérée comme un pic. Si $-L = 3$, alors Resan tiendra compte des trois fréquences situées de part et d'autre de la résonance présumée. Si $-L = 0$, alors Resan conservera tous les pics (fréquences dont l'amplitude est inférieure à celle des fréquences directement adjacentes). Ainsi, plus le Pics Width a une valeur importante, moins le nombre de pics retenus est important.

Pics Amplitude (-l)

Détermine la différence d'amplitude (en dB) qui doit séparer les -L fréquences adjacentes au pic présumé pour que celui-ci soit extrait comme un pic.

Min Amplitude 1st window (-a)

Amplitude minimale en dB pour qu'un point soit potentiellement un pic; cette amplitude doit être atteinte dans la première fenêtre (-t + -w).

Min Amplitude 2d window (-A)

Amplitude minimale en dB pour qu'un point soit potentiellement un pic; cette amplitude doit être atteinte dans la seconde fenêtre (-T + w).

Ces deux options -a et -A sont préalables dans cette section d'analyse. C'est-à-dire que toutes les fréquences dont Resan tiendra compte pour l'extraction des pics doivent avant tout être inférieures à -a dB dans la première fenêtre, et à -A dans la seconde.

Si la valeur de -A est par défaut supérieure à celle de -a, c'est parce que l'amplitude du son décroît généralement dans le temps.

Augmenter ces deux valeurs d'une façon trop importante (par exemple à -40 et -50 dB) réduit le nombre de pics analysés. Les caractéristiques retenues dans le son sont alors les plus importantes. Le détail spectral est par conséquent amoindri (cette remarque est tout aussi valable pour l'option -l vue précédemment).

Extract Smaller (-P)

Ce paramètre est le plus important. Il définit la méthode qui sera utilisée pour l'interprétation des données qui précèdent (-l, -L, -a, -A).

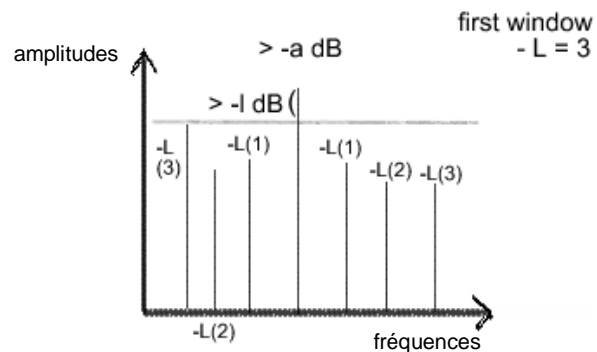
Trois possibilités sont offertes :

1. $P = 0$ (all smaller)

Les $-L$ fréquences se trouvant directement de part et d'autre du pic présumé (à l'exception des fréquences adjacentes $-L(1)$) doivent lui être inférieures de $-l$ db.

Attention : Si $-L = 1$, on prend en compte tous les pics (c'est-à-dire toutes les fréquences entourées de deux fréquences d'amplitude inférieure). On ne tient pas compte des fréquences adjacentes $-L(1)$ pour $-l$. Donc, si $-L = 1$ et $-l = 5$ dB, alors même un pic dont les fréquences adjacentes lui seraient inférieures de seulement 1dB serait extrait.

L'amplitude de ce pic présumé doit être supérieur à $-a$ pour la première fenêtre; et à $-A$ pour la seconde.



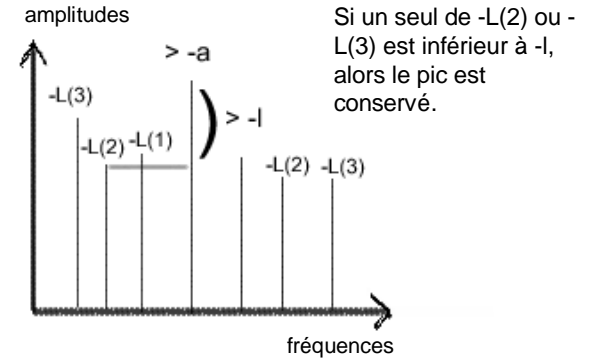
Cet exemple est caractérisé par les paramètres suivants :

1. First Window (donc $-a$ plutôt que $-A$);
2. $-L = 3$, donc contrôle sur les pics $-L(2)$ et $-L(3)$;
3. $-L(1)$ ne sont pas pris en compte puisque ce sont les pics directement adjacents; c'est-à-dire que $-L(1)$ doit seulement être inférieur au pic.
4. $L(1)$ équivaut pour all smaller à $L(0)$.

2. P = 1 (one smaller)

Au lieu que toutes les $-L$ fréquences (adjacentes exceptées) doivent être inférieures de $-l$ dB au pic présumé, il suffit maintenant qu'une seule de ces fréquences le soit pour que le pic soit extrait. Comme pour $-P = 0$, si $-L = 1$, on prend en compte tous les pics (c'est-à-dire toutes les fréquences entourées de deux fréquences d'amplitude inférieure). Car on ne tient pas compte des fréquences adjacentes ($-L(1)$) pour $-l$.

Donc, si $-L=1$ et $-l = 5$ dB, alors même un pic dont les fréquences adjacentes lui seraient inférieures de seulement 1dB serait extrait.



3. P = 2 (mean smaller)

Pour ce calcul, on tient compte de toutes les $-L$ fréquences de part et d'autre du pic présumé, fréquences adjacentes incluses.

On fait la moyenne des $-L$ pics qui se trouvent de part et d'autre du pic présumé; cette moyenne doit être inférieure au pic présumé de $-l$ dB.

Si $-L = 1$, on ne prend donc plus en compte tous les pics. Tout pic présumé doit alors être supérieur à la moyenne de ses deux pics adjacents de $-l$ dB.

Dans cette section, l'analyse des fréquences a été traitée. Seules les fréquences importantes ont été retenues (selon les critères d'analyse).

Il s'agira, dans la section suivante d'observer les informations collectées dans chacune des deux fenêtres, et de les apparier. C'est-à-dire que chaque pic retenu pour la première fenêtre sera comparé à chaque pic retenu pour la seconde fenêtre. De cette manière, Resan pourra observer comment le son évolue dans le temps, en fréquence et en amplitude. Il pourra en déduire les caractéristiques qui, dans le son, évoluent d'une façon cohérente, séparant ce qui tient plutôt de l'accidentel de ce qui tient de l'essentiel.

Appariement des pics

Pics Matching										
Matching Mode	-m	Closest	-	-	-	-	-	-	-	-
Matching Threshold	[Hz]	-s	0.0	-	-	-	-	-	-	-
Coefficient Threshold	[Coef]	-y	2.0	-	-	-	-	-	-	-

Nous voici maintenant à l'appariement des pics. L'objectif est de de comparer les deux première et seconde fenêtres et d'observer les corrélations entre les différents pics extraits.

Matching Mode (-m)

Les deux jeux de pics sont parcourus par ordre de fréquences croissantes. Pour la valeur 0 (first match), le premier appariement possible est retenu. Pour la valeur 1 (closest), le meilleur appariement possible est recherché (ce sont les pics les plus proches qui sont appariés).

Cet appariement est effectué suivant un critère de proximité en fréquence. Peuvent être appariés deux pics dont la distance fréquentielle est inférieure à -s hz.

Exemple :

-s = 20 hz; -m = 0; (1) (2) (3) (4) (5) (6)

Pics extraits sur la 1ère fenêtre : 100, 130, 300, 350, 360, 700.

Pics extraits sur la 2de fenêtre : 110, 132, 275, 357, 383, 708.

Alors les pics appariés seront : (1: 100 et 110 hz), (2: 130 et 132 hz), (3: pas apparié), (4: 350 et 357 hz), (5: pas apparié), (6: 700 et 708 hz).

Avec les même paramètres et -m = 1 : (1: 100 et 110 hz), (2: 130 et 132 hz), (3: pas apparié), (4: 360 et 357 hz), (5: pas apparié), (6: 700 et 708 hz).

Matching Threshold (-s)

Seuil en hz pour l'appariement (c'est la valeur en hz maximale pour que l'appariement entre deux pics soit possible). La valeur par défaut est -y fois la résolution de l'estimation spectrale.

Coefficient Threshold (-y)

Facteur pour calcul de -s si non spécifié.

Les pics ont maintenant été appariés. Des informations sur l'extinction des résonances ont été collectées; ce premier résultat d'analyse va maintenant être confronté aux autres résultats obtenus lors des étapes suivantes (le nombre des étapes peut varier de une à huit pour une analyse complète Resan). Dans cette dernière section, tous les résultats d'analyse seront conservés. Ainsi, les résonances qui pourront être appariées le seront, et les autres seront laissées sans changement. Il sera aussi possible de corriger sensiblement certains défaut qui auraient pu survenir lors des différents pas dans l'analyse; notamment en ajustant les résultats d'extinction des résonances dans le son.

Appariement des résonances

Merged Analysis											
Resonances Matching		-M	Closest	-	-	-	-	-	-	-	-
Resonances Doubling		-e	Off	-	-	-	-	-	-	-	-
Matching Checking		-V	No	-	-	-	-	-	-	-	-
Matching Threshold	[Hz]	-S	0.0	-	-	-	-	-	-	-	-
Coefficient Threshold	[Coef]	-Y	2.0	-	-	-	-	-	-	-	-
Max Bandwidth For Doubling	[Hz]	-b	12.0	-	-	-	-	-	-	-	-
Min DeltaBandwidth For Doubling	[%]	-B	5.0	-	-	-	-	-	-	-	-

Étape n° 1 2 3 4 5 6 7 8

L'appariement des pics concernait le résultat des deux fenêtres d'analyse sur le son. L'appariement des résonances travaille maintenant sur les résultats des modèles élémentaires précédent (celui obtenu après appariement des pics) et courant. La méthode est globalement la même que pour l'étape précédente.

Resonances Matching (-M)

Même paramètres que -m; mais une différence : alors que, pour -m, les pics n'étaient conservés que si l'appariement était possible, avec -M, ils sont conservés, et rajoutés sans modification lorsque l'appariement n'est pas possible.

Matching Threshold (-S)

voir -s.

Coefficient Threshold (-Y)

voir -y.

Matching Checking (-V)

Lorsque l'option -V (vérification du taux de décroissance des résonances) est activée (Yes), le mécanisme suivant est activé : lorsqu'une résonance modèle produite par l'étape précédente n'est pas retrouvée dans l'analyse élémentaire de l'étape courante, on vérifie que le taux de décroissance de cette résonance est compatible avec la disparition observée. Si elle ne l'est pas, le taux de décroissance est augmenté dans le nouveau modèle (une résonance peut très bien échapper à une étape intermédiaire de l'analyse sans pour autant qu'elle soit "éteinte" dans le son).

Resonances Doubling (-e)

Lorsque l'option -e est activée, on cherche à modéliser les résonances dans le son dont la décroissance n'est pas exponentielle par deux filtres dont les taux de décroissance diffèrent, suivant le mécanisme suivant : si une résonance du modèle de l'étape précédente est retrouvée dans l'analyse élémentaire de l'étape courante avec un taux de décroissance significativement plus faible, la résonance est doublée. Une première résonance, plus amortie, représente l'attaque; la seconde, moins amortie, mais d'amplitude plus faible, représente la seconde partie de la réponse impulsionnelle.

Pour que le doublement intervienne, deux conditions doivent être remplies : la largeur de bande de la résonance dans l'étape courante (qui mesure le taux d'amortissement) doit être inférieure à une valeur spécifiée par l'option -b. Et le changement constaté par rapport à l'étape précédente doit être supérieur en pourcentage à la valeur spécifiée par l'option -B (uniquement si -e est positionné sur Min&Max). Ce mécanisme est utile pour certains sons (de piano par exemple), mais délicat d'utilisation.

L'analyse est maintenant achevée. Il ne reste plus qu'à créer un dictionnaire contenant toutes les informations de fréquence, d'amplitude, de largeur de bande... qui sera exploitable dans Diphone. Toutes les transformations y seront possibles.

Pour la création du dictionnaire, simplement cliquer sur « dico » dans la fenêtre principale. Le logiciel Diphone est automatiquement démarré, et le dictionnaire ouvert.

2. Trois exemples sonores
à travers différentes
possibilités de
paramétrage.

Chacun des trois sons qui vont être analysés est un son de percussion. Steel Drum, Marimba et Udu.

Ils ont été préparés avant l'analyse. C'est-à-dire que le fichier débute au premier échantillon de son de la percussion. Ainsi, l'attaque est la plus précise et l'analyse plus fine (un « blanc » de 0.1 seconde au début du fichier pourrait dénaturer les résultats obtenus).



Steel
Drum



Udu



Marimba

Huit analyses différentes ont été faites sur chacun de ces trois sons. Il est recommandé de placer les huit fichiers d'analyse fournis dans le dossier presets de Diphone. Les paramètres d'analyse seront alors accessibles par nom de preset dans Resan (toute fenêtre d'analyse Resan dont les paramètres seraient à conserver pour de futures utilisations peuvent ainsi être consignées dans le dossier preset; c'est une façon simple de construire / échanger ses banques de paramètres d'analyse).

Analysis type 1 est relativement courant et simple.

Les fenêtres d'analyse se succèdent d'une façon efficace; les deux premières étapes d'analyse se concentrant sur l'attaque du son; les 4 dernières sur ses résonances. Le choix d'utiliser « first window beginning relative » et « window duration relative » (en %) plutôt que les deux autres modes (en secondes) s'explique ainsi : cette suite de paramètre devant s'appliquer à l'analyse de plusieurs sons de taille différente, on a ainsi la possibilité d'utiliser le même preset; il est en effet inutile de calculer pour chaque son les valeurs exactes en secondes pour chacune des étapes d'analyse.

Pics Width étant à 0, tous les pics vont être retenus.

Voici maintenant le résultat de la resynthèse de l'analyse obtenue pour chacun des trois sons :

Comparer le résultat avec les fichiers sons originaux (page précédente) :



Steel
Drum



Udu



Marimba

Il est intéressant de constater que Resan a détecté 154 résonances (ou formants) pour le Steel Drum, 38 pour le marimba et 52 pour le Udu.

Passons maintenant à l'analyse suivante :

Les paramètres sont les mêmes que pour l'analyse 1, exception faite de l'estimation spectrale. Ici, les window duration relative sont plus longues. Écoutons le résultat :



Steel
Drum



Udu



Marimba

Les sons sont dans l'ensemble moins riches dans les médium/hautes fréquences. On peut aussi remarquer plus de grave dans la deuxième resynthèse du Steel Drum. Ceci est dû au fait que, plus la fenêtre d'analyse est grande, plus importante la précision dans le registre des basses fréquences. Dans cette analyse, 129 résonances pour le Steel Drum, 30 pour le Marimba et 32 pour le Udu ont été détectées.

Pour l'exemple 3, on reprend les paramètres de l'exemple 1; mais cette fois, on donne au pics width la valeur de 3. Ainsi, il faut maintenant, pour qu'un pic soit retenu, qu'il soit supérieur de -1 dB aux trois pics adjacents (à l'exception des deux - de part et d'autre - **directement** adjacents). En conséquence, moins de pics seront retenus; ainsi, 79 formants pour le Steel Drum, 18 pour le Marimba et 7 pour le Udu :



Udu et Marimba ne sont presque plus reconnaissables; trop peu de formants viennent les dessiner.

Nous passerons les résultats de l'analyse 4. Il vous est maintenant facile de les créer par vous même en utilisant le preset 4 et les trois sons originaux fournis.

Analyse 5. Allons immédiatement au résultat :



L'option resonance doubling utilisée pour cette analyse a un résultat tour à tour intéressant et surprenant. Le son de Marimba devient plus timbré (avec 102 formants contre 18 pour l'analyse 3 et 28 pour l'analyse 4 - si vous l'avez effectuée). D'autre part, un formant s'échappe du son de Steel Drum, comme le produit des deux filtres -b et -B de -e (resonance doubling). Cette option, pour être efficace, doit être réglée avec beaucoup de précision.

Pour le cas des analyses 6, 7 et 8, l'objectif était de montrer comment l'option extract smaller peut avoir une incidence importante sur le nombre de formants retenus par l'analyse.

	Udu	Steel Drum	Marimba
Analyse 1	7	79	7
Analyse 6	20	102	26
Analyse 7	15	99	26

Le son obtenu en est bien évidemment altéré :

Analyse 6 :
Extract smaller
ALL



Analyse 7 :
Extract smaller
ONE



Analyse 8 :
Extract smaller
MEAN



L'objectif de cette deuxième partie était moins de montrer des liens évidents entre paramètres d'analyse et résultat à la resynthèse, que l'utilité d'expérimenter des combinaisons. Les résultats les plus intéressants sont parfois le produit de technique d'analyse « hasardeuses ». On peut aussi constater que les presets qui donnent les meilleurs résultats peuvent avoir des résultats médiocres sur d'autres sons. Il n'existe donc pas de règle stricto sensu. L'expérimentation reste la meilleure ressource, surtout dans le cas où le résultat d'analyse/resynthèse désiré n'est pas une reproduction fidèle du son d'origine, mais une variation enrichissante de ce son.

Pour conclure, si une bonne compréhension des paramètres est indispensable à la création de paramètres d'analyse efficaces et non-buggés, la richesse sonore peut aussi être le produit d'analyses « aventurières ».

Pour clore ce chapitre, il serait intéressant d'écouter le résultat de resynthèse des 5 premières étapes d'une analyse (il s'agit ici de l'analyse du Steel Drum avec Analysis type 5).



Étape n° 1
50 résonances



Étape n° 2
112 résonances



Étape n° 3
174 résonances



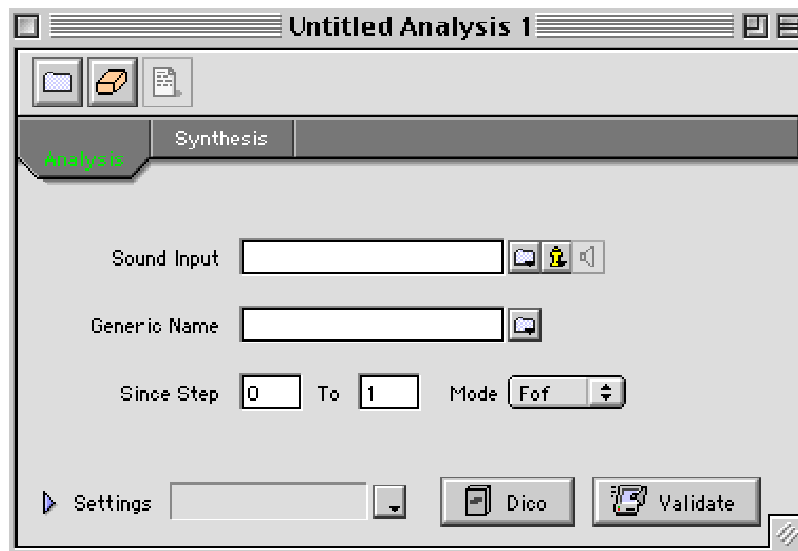
Étape n° 4
200 résonances



Étape n° 5
210 résonances

3. Quelques remarques additionnelles

- Toutes les analyses faites jusqu'ici étaient en mode fof.



Un second mode est disponible : le mode filtre. Sans rentrer dans le détail, une analyse en mode filtre permet, lors de l'exploitation des données dans Diphone, non pas de générer des formants (mode fof), mais de filtrer une source (au choix bruit blanc, fichier son, formants) avec un banc de filtres - produit de l'analyse.

Pour exemple, ce son, très sommaire, d'un enregistrement de cymbale filtré dans un banc de filtres produit par une analyse de marimba :



Son de marimba avec preset 4 filtrant une cymbale; on a l'enveloppe de la cymbale et le timbre du marimba.

- Il est possible, directement dans Resan, de resynthétiser un résultat d'analyse pour se rendre compte immédiatement de la qualité de l'analyse. Pour ce faire, simplement aller sur l'onglet Synthesis; sélectionner le fichier sdif désiré et spécifier le nom du fichier de sortie et son taux d'échantillonnage.

Attention : les fichiers reconnus sont du type :

x.m1.FOF.sdif : 1ère étape d'analyse du son x dans le mode fof;

Y.m5.FIL.sdif : 5ème étape d'analyse du son y dans le mode filtre.

« m » indique que ce fichier est le résultat d'une **merged analysis**.

- Une fois l'analyse terminée et concluante, la dernière étape jusqu'au travail dans Diphone consiste à créer un dictionnaire. Pour ce faire, simplement cliquer sur Dico dans la fenêtre principale et valider la création du dictionnaire.

4. Documents annexes

Une habitude qui peut s'avérer très profitable serait de consulter, lorsque cela paraît nécessaire, les fichiers ASCII générés par Resan lors des différentes étapes d'analyse. Pour ce faire, il suffit, dans la partie out typing des paramètres de sélectionner la sortie d'analyse en fichier texte. Les résultats sont ensuite consultables dans tout éditeur de texte. Voici, par exemple, quelques résultats pour une analyse du Udu :

```
numero total de resonances 41
model_type MODELE_MERGE
son UduASCII.temp.aiff
date Wed Nov 7 15:55:28 2001
sample_rate 44100.000000
niveau_analyse 1
```

Etape d'analyse 1;
41 résonances.

```

niveau 0
t1 0.002000 t2 0.050000 wsize 0.048000 fftsize 0.092880
window type BLACKMAN_HARRIS_3_(-92_dB) left_dB_floor -80.000000
right_dB_floor -96.000000 peak_width 1 peak_height 2.000000 peak_mode 0
peak_match_type 1 peak_threshold 21.533203
double_mode 0 max_d_band 0.000000 min_d_reldeltaT 0.000000 check_mode 0
res_match_type 0 res_threshold 0.000000
psize1 54 psize2 135 ele_model_size 25 model_size 25
```

On retrouve ici les différents paramètres spécifiés dans Resan.

```

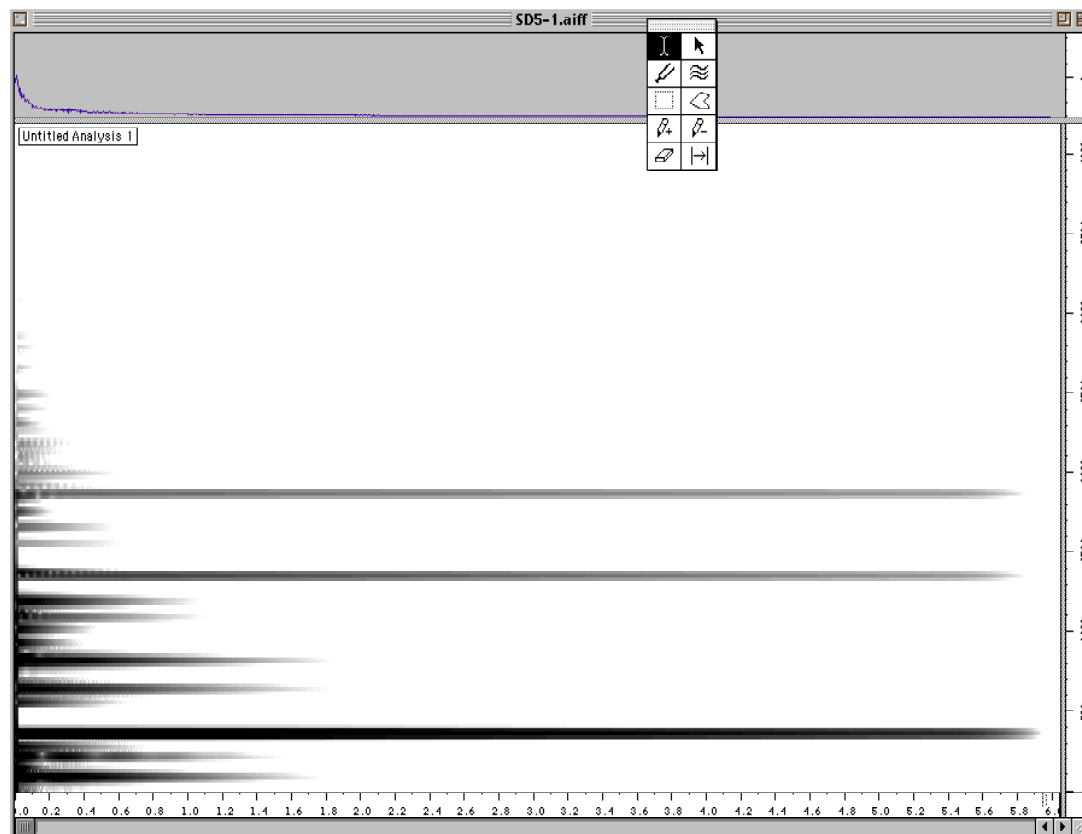
model_size 41
1    70.698   -8.956 (0.357)   1.63  0  0
2   210.367  -24.377 (0.060)   12.7  0  0
3   307.557  -33.027 (0.022)   16.6  0  0
4   351.982  -41.405 (0.009)   13.5  0  0
5   408.359  -45.995 (0.005)   12.4  0  0
6   494.339  -40.855 (0.009)   13.1  0  0
7   512.530  -38.705 (0.012)   14.1  0  0
8   759.002  -55.427 (0.002)   5.45  0  0
9   881.917  -44.875 (0.006)   19.1  0  0
10  1258.529  -39.112 (0.011)   10.3  0  0
11  1624.863  -63.245 (0.001)   6.27  0  0
12  1673.909  -53.980 (0.002)   4.24  0  0
13  2233.213  -45.020 (0.006)   12.8  0  0
14  2445.879  -52.504 (0.002)   8.25  0  0
15  2514.719  -59.728 (0.001)   7.69  0  0
16  2892.407  -56.693 (0.001)   7.74  0  0
17  3174.141  -50.249 (0.003)   16.4  0  0
18  3601.518  -55.964 (0.002)   10    0  0
19  3673.491  -59.410 (0.001)   9.11  0  0
20  3718.520  -55.418 (0.002)   9.98  0  0
21  4009.334  -54.055 (0.002)   15.7  0  0
22  4440.112  -62.324 (0.001)   9.29  0  0
23  5383.748  -65.590 (0.000)   13.6  0  0
24  5530.458  -64.432 (0.001)   10.9  0  0
25  5695.836  -67.374 (0.000)   14.2  0  0

```

La première colonne indique le numéro de la fof; le 2ème sa fréquence; la troisième son amplitude puis sa largeur de bande (bandwidth).

Pour exemple de l'utilité de cette démarche, reprenons notre son de Steel Drum analysé avec Analysis type 5. On pouvait très distinctement percevoir trois résonance qui, dans ce son, s'échappaient. Il suffit d'aller dans le fichier ASCII de cette analyse (ne pas oublier de spécifier dans les paramètres de out typing comme valeur par défaut la sortir en ASCII de chaque étape d'analyse).

On peut noter directement, si l'on se réfère à la colonne bandwidth que trois formants ont une largeur de bande particulièrement étroite. Ceci a pour effet - il est important de le savoir pour aborder diphone - de donner plus de résonance aux formants. Il s'agit des lignes 10, 23 et 29; aux fréquences de 367, 1357 et 1875 hz.



```

numero total de resonances 50
model_type MODELE_MERGE
son SD5iterations.temp.aiff
date Thu Nov 8 09:52:40 2001
sample_rate 44100.000000
niveau_analyse 1
(...)
model_size 50

```

1	10.263	-53.379 (0.002)	4.89	1	0
2	12.665	-55.945 (0.002)	3.33	-1	0
3	52.880	-75.741 (0.000)	1.59	0	0
4	88.649	-17.170 (0.139)	4.04	1	0
5	91.845	-22.760 (0.073)	1.17	-1	0
6	140.659	-23.661 (0.066)	8.01	0	0
7	222.840	-32.185 (0.025)	1.75	1	0
8	219.627	-33.576 (0.021)	1.09	-1	0
9	277.636	-35.740 (0.016)	2.17	0	0
10	367.085	-26.721 (0.046)	0.118	0	0
11	560.169	-38.829 (0.011)	2.31	0	0
12	610.448	-54.039 (0.002)	1.27	0	0
13	648.093	-33.980 (0.020)	0.918	0	0
14	690.479	-51.032 (0.003)	2.11	0	0
15	836.718	-15.935 (0.160)	8.08	1	0
16	826.510	-25.279 (0.054)	1.06	-1	0
17	934.180	-32.772 (0.023)	4.44	0	0
18	1027.044	-33.444 (0.021)	3.52	0	0
19	1096.699	-43.709 (0.007)	1.2	0	0
20	1114.259	-29.801 (0.032)	7.02	0	0
21	1137.559	-34.836 (0.018)	11.1	0	0
22	1196.853	-31.909 (0.025)	1.6	0	0
23	1357.807	-46.499 (0.005)	0.109	0	0
24	1380.043	-36.914 (0.014)	2.37	0	0
25	1562.194	-60.814 (0.001)	1.05	0	0
26	1668.627	-52.006 (0.003)	1.71	0	0
27	1764.969	-43.263 (0.007)	5.98	0	0
28	1864.260	-39.582 (0.010)	5.01	0	0
29	1875.784	-54.365 (0.002)	0.0566	0	0