LA POLYSÉMIE DU TEMPS RÉEL ET DU TEMPS DIFFÉRÉ

Karim Barkati
CICM – Université Paris VIII
karim.barkati@free.fr

RÉSUMÉ

Au cours de l'interaction vivante avec les systèmes d'informatique musicale, on ressent avec évidence qu'il y a quelque chose qui relève du « tout de suite » et autre chose qui relève du « plus tard », et, dans ce sens-là, temps réel et temps différé constituent des catégories pertinentes, issues de l'expérience interactive, et délimitées par notre sensation d'immédiateté. La force de cette approche intuitive – la congruence entre notre expérience humaine des sensations d'immédiateté et de délai avec la dichotomie usuelle temps réel / temps différé – semble entériner ce modèle historique, et, ce faisant, dissimule ses faiblesses. De fait, un examen plus approfondi révèle des failles susceptibles de faire vaciller ce modèle robuste en apparence, dont l'évidence intuitive finit par s'évanouir. Plusieurs observations y contribuent : le caractère flou de la frontière entre temps réel et temps différé, la rapidité de l'évolution de l'informatique en général et de l'informatique musicale en particulier, et des contradictions conceptuelles majeures entre les niveaux technologique, pratique et musical des deux catégories.

1. UN SEUIL DISCUTABLE

1.1. Une frontière floue

La clé même de la définition des deux catégories du temps réel et du temps différé – le seuil de latence – repose en général sur des considérations d'ordre perceptif, c'est-à-dire nécessairement variables d'un individu à un autre, d'un contexte d'audition à un autre, ou d'un moment de l'interaction à un autre [1] . Il s'agit toujours de statistiques sur une population particulière, et l'unicité d'une valeur annoncée traduit toujours une *moyenne*. Puisque ces enquêtes subissent nécessairement une variabilité non négligeable, elles deviennent contestables quand elles décrètent une valeur unique et fixe – par exemple $s = 24$ ms. Autrement dit, devrait-on vraiment classer un logiciel ou un système numérique du côté du temps différé lorsqu'on mesure un temps de latence de 25 ms ? De 26 ms ? Ou de 30 ms ?

Ce questionnement, malgré sa légèreté, montre une première limite majeure de l'arbitraire de la définition d'un seuil *fixe* pour départager le temps réel du temps différé. Il suggère préférablement une frontière épaisse, ou bien une frontière floue. Or la solution d'élargir ce seuil ponctuel à un intervalle de la forme $]s_{min}; s_{max}[$ pour formaliser une frontière « épaisse », complique et fragilise le modèle sans apporter non plus de réponse totalement satisfaisante. En effet, d'une part le problème de l'arbitraire de la définition subsiste, puisque les deux bornes s_{min} et s_{max} de l'intervalle proposé restent fixées et donc contestables en tant que telles, et, d'autre part, des valeurs mesurées dans un contexte particulier ne peuvent être considérées comme pertinentes que dans ce contexte particulier. Pour illustrer ce dernier point, on peut aisément imaginer qu'une latence relativement importante mais acceptable dans un contexte musical où les textures prédominent ne le sera sans doute plus dans un contexte musical où la pulsation et les percussions prévalent.

En substance, le flou « indissipable » de la frontière entre temps réel et temps différé relève ici de deux ordres : l'arbitraire de fixer une valeur ou un intervalle plutôt qu'un autre, et l'impossibilité de généraliser toute valeur ou tout intervalle qui semble pourtant pertinent dans un contexte particulier.

1.2. Des définitions multiples

En informatique musicale, plusieurs définitions du temps réel coexistent, tout en différant selon le point de vue adopté. Pour montrer cette multiplicité, nous allons seulement remarquer qu'un même auteur parfaitement légitime dans notre domaine, Curtis Roads dans le cas présent [2] , peut proposer des définitions irréductibles l'une à l'autre :

Définition 1 [11, p. 109] *Le matériel rendait impossible de concevoir la synthèse du son haute-fidélité en temps réel. Par temps réel, nous voulons dire qu'on entendra le son en même temps qu'on verra les gestes du musicien.*

[1] Henri Piéron notait déjà en 1913 que « [...] l'allongement des temps de réaction au fur et à mesure qu'on se rapproche du seuil de sensation dépend essentiellement d'une augmentation du temps nécessaire pour que la transformation de l'excitant physique en phénomène cérébral de nature sensorielle s'effectue ; si l'on descend au-dessous du seuil, ce temps devient infini. [...] la loi de décroissance des temps de latence en fonction des intensités varie suivant les sensations [...] » [6, p. 17–18]

[2] D'autres auteurs auraient pu illustrer cette section.

Définition 2 [12, p. 69] *Qu'entendons-nous par temps réel ? Dans ce contexte [de la synthèse numérique], le temps réel signifie que nous pouvons achever les calculs pour un échantillon en un temps inférieur à une période d'échantillonnage.*

Précisons que cette multiplicité des définitions n'est pas du tout un phénomène isolé, mais que le temps réel a effectivement reçu de nombreuses définitions hétérogènes. Nous étudierons plus avant les aspects polysémiques des expressions « temps réel » et « temps différé » à la section 2 page suivante, néanmoins nous pouvons examiner dans cette présente section en quoi ces deux définitions précises divergent.

1.2.1. Du point de vue de la perception

La première définition sous-entend une situation de concert et se fonde sur la perception humaine d'une simultanéité (chez le public qui voit les gestes du musicien et entend les résultats sonores de ces gestes). Cette définition implique clairement deux sens – la vision et l'audition –, or la perception de la simultanéité d'événements multisensoriels reste une question controversée et un sujet de recherche toujours d'actualité [3]. Elle fait appel à des mécanismes complexes, comme l'anticipation, la comparaison, la réinjection et la récursion [3], ainsi qu'à des notions récentes plus spécifiques, comme le « ventriloquisme temporel » ou la « fenêtre déplaçable » pour l'intégration multisensorielle [4].

L'article *The perception of cross-modal simultaneity* [3] rappelle d'abord l'importance des différences interindividuelles et l'asymétrie temporelle de la perception de la désynchronisation [5], puis revient sur plusieurs études qui lui précèdent à propos de la perception de la simultanéité audio/vidéo (cf. figure 1), et livre enfin les résultats d'une expérience nouvelle. Dans cette expérience, un « acteur » aux yeux bandés *frappe* une surface relativement muette avec le *Radio baton* de Max Mathews, tandis qu'il *entend* un impact préenregistré à travers son casque audio, l'impact étant avancé ou retardé d'une valeur prise aléatoirement entre -200 ms et +200 ms. Un « observateur » *regarde* l'acteur depuis une pièce isolée acoustiquement (double vitrage), à une distance de 2 mètres de l'acteur, et *entend* l'impact préenregistré à travers son casque, avec le décalage aléatoire. Après chaque essai, les sujets déclarent si le son a eu lieu *en même temps* que le geste ou non. Selon la ligne de partage des 75%, l'article conclut que les « acteurs » ont détecté des asynchronies entre le toucher et l'audition à partir de -25 ms et +42 ms, et que les « observateurs » ont détecté des asynchronies entre la vue et l'audition à partir de -41 ms et +45 ms :

> *A simple and intuitive way to consider the data is to plot the percentage of the time subjects judged the stimuli to be synchronous as a function of the sound-contact asynchrony (indicated by responses of "same"). If we consider the 75% line to indicate subjects' detection threshold (i.e., the asynchrony beyond which fewer than 75% of responses incorrectly judge simultaneous), we find that actors detected asynchronies at -25 and +42 msec, and observers detected asynchrony at -41 and +45 msecs.*

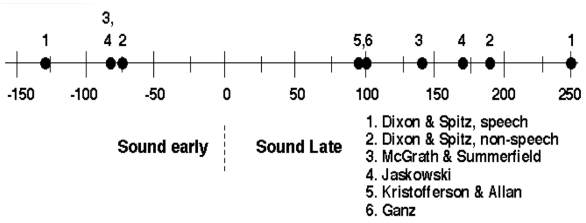

Figure 1. Estimations récentes du seuil de simultanéité audio/vidéo, d'après Levitin & al.

Dans le cas qui nous intéresse (la latence d'un système audionumérique perceptible par un public entre la vue du geste d'un musicien et l'audition du son produit par ce geste), nous pouvons retenir de ces quatre résultats uniquement celui qui concerne le seuil positif (plutôt que le seuil négatif) et perçu par l'observateur (plutôt que par l'acteur) : soit un seuil de +45 ms. Cette valeur, qui correspond à la première définition du temps réel susmentionnée, sera à comparer avec celle qui peut correspondre à la seconde définition.

1.2.2. Du point de vue de la fréquence d'échantillonnage

La seconde définition, liée à la puissance de calcul d'un ordinateur, est issue de l'analyse de Curtis Roads citée précédemment. Cette analyse démontre que pour qu'une synthèse sonore puisse avoir lieu en temps réel, elle doit réaliser les calculs nécessaires à chaque échantillon en un temps inférieur à la période

[3] « *An unsolved problem in cognitive science concerns the perception of simultaneous events, particularly when the information impinging on the sensory receptors comes from two different sensory modalities.* » [3] La perception de la simultanéité d'événements multisensoriels constitue d'ailleurs une des plus anciennes questions de la psychologie expérimentale, depuis le « problème de l'observatoire de Greenwich » en 1796 (concernant la vue et l'audition).

[4] « *We are rarely aware of differences in the arrival time of inputs to each of our senses. New research suggests that this is explained by a 'moveable window' for multisensory integration and by a 'temporal ventriloquism' effect.* » [14] Les auteurs notent aussi que la supériorité de la vitesse de la lumière par rapport à celle du son dans l'air (3 000 000 m/s contre 340 m/s) peut être partiellement compensée par la lenteur de la transduction chimique de la lumière à partir de la rétine comparée à la transduction mécanique des ondes sonores dans l'oreille.

[5] « *Past investigation of intermodal asynchrony has focused audio/video. There are two general findings. (1) There appear to be large individual differences in perception thresholds. (2) Thresholds are asymmetric : people are more likely to perceive events as synchronous when the audio* precedes *the video than vice versa.* » [3]

d'échantillonnage. Ainsi, pour un taux d'échantillonnage standard de 44 100 échantillons par seconde, les calculs doivent donc être achevés en moins de 23 *micro*secondes [6] [7] . . .

Or en comparant ces 23 μs (pour la seconde définition) aux 45 ms (admissibles comme limite du temps réel en regard de la première définition), on peut constater une différence colossale [8] d'un facteur 2 000 ! D'où peut venir un tel écart entre deux définitions du temps réel – une catégorie pourtant si familière en informatique musicale ? Les divergences des études sur la perception, bien qu'elles soient fréquentes, restent toujours bien en deçà d'un tel écart (cf. figure 1 page précédente). Une telle différence laisse à penser que derrière la même locution, il ne s'agit pas de la même signification : cet écart peut en effet s'expliquer par une confusion dans le langage courant entre les deux niveaux *technologique* et *perceptif* de la définition du temps réel.

Conformément à la démonstration de Curtis Roads [12, p. 69–70], un flux audionumérique doit être émis aussi rapidement que sa fréquence d'échantillonnage le requière pour pouvoir être audible en temps réel, sous peine d'artefacts audionumériques très gênants à l'audition. De fait, au bout de la chaîne de traitement numérique, si le convertisseur numérique-analogique (ou CNA) ne reçoit pas les échantillons à temps, il produira alors un signal erroné, ce qui se traduira typiquement par des saccades sonores intempestives, avec des « clics numériques » plus ou moins perceptibles selon les cas. D'un point de vue technique donc, les échantillons considérés individuellement doivent impérieusement avoir été calculés et arriver au CNA à une vitesse supérieure ou égale à la fréquence d'échantillonnage.

* * *

En résumé, on peut considérer que la notion de seuil qui est traditionnellement invoquée pour définir le couple temps réel / temps différé est finalement discutable, pour deux raisons principales : le flou de la frontière sensée les séparer, et la multiplicité des définitions.

2. LA POLYSÉMIE DU TEMPS RÉEL ET DU TEMPS DIFFÉRÉ

Le télescopage de plusieurs valeurs différentes pour définir le problématique seuil de latence censé séparer la catégorie du temps réel de celle du temps différé nous a amené à reconsidérer la définition du temps réel en informatique musicale dans le sens de la *pluralité*, pour proposer de distinguer trois champs sémantiques. Le découpage du sens reste un exercice délicat, car les sens dégagés possèdent évidemment des traits sémantiques communs, mais cette démarche qui, à notre connaissance, n'a pas encore été suivie, devrait permettre de tirer au clair certaines ambiguïtés persistantes.

2.1. Le champ technique

Le premier sens, le plus étroit, concerne le plus souvent le traitement numérique du signal (TNS) en rapport avec l'échantillonnage audionumérique, soit un seuil de latence proportionnel à la période d'échantillonnage du signal (de l'ordre de 23 μs pour un signal stéréophonique traditionnel échantillonné à 44 100 Hz, multiplié par la longueur des vecteurs audionumériques). Ce premier sens, impliqué dans les expressions telles que « synthèse (en) temps réel » ou « traitement (en) temps différé », nous proposons de l'indiquer à travers les expressions « temps réel *technique* » et « temps différé *technique* ».

2.1.1. Le champ sémantique le plus ancien

Dans ce champ sémantique, la notion de temps différé se définit en opposition simple avec celle de temps réel, en fonction de la capacité ou de l'incapacité d'un système numérique à *garantir le respect d'une contrainte temporelle donnée*. Il s'agit du sens le plus ancien, proche des définitions généralistes de l'informatique, comme cette définition de Gillies [1, p. 18–19] :

> Un système temps réel est un système dans lequel l'exactitude des applications ne dépend pas seulement de l'exactitude du résultat mais aussi du temps auquel ce résultat est produit. Si les contraintes temporelles de l'application ne sont pas respectées, on parle de défaillance du système. Il est donc essentiel de pouvoir garantir le respect des contraintes temporelles du système. Ceci nécessite que le système permette un taux d'utilisation élevé, tout en respectant les contraintes temporelles identifiées.

2.1.2. Une simple dichotomie technique

Ainsi, en tant que *locution adjective* (par exemple dans l'intitulé « synthèse en temps différé »), les deux expressions « temps réel » et « temps différé » caractérisent les applications et les processus selon qu'ils respectent ou non le seuil de latence fixé. Sous cette forme adjectivale, la préposition « en » peut être présente ou omise, de même qu'un trait d'union entre « temps » et « réel » ou « différé » ; ces deux remarques

[6] L'auteur de conclure en 1998 : « Ceci signifie qu'il est difficile à un ordinateur d'achever les calculs nécessaires pour plus de quelques oscillateurs simples en *temps réel*. » [12, p. 69]

[7] Cette période très brève est toutefois couramment élargie par la technique de vectorisation audionumérique pour les niveaux supérieurs à celui du convertisseur ; p. ex. à 1,45 ms pour un vecteur audio de 64 échantillons.

[8] Notons que cette période d'échantillonnage de 23 μs, qui correspond à la « qualité CD », ne vaudrait plus que 10 μs pour les taux professionnels actuels de 96 kHz et seulement 5 μs pour ceux qui travaillent jusqu'à 192 kHz. . .

sur la préposition et le trait d'union semblent s'appliquer en anglais, avec une forme supplémentaire fusionnée qu'on rencontre dans « *realtime synthesis* » par exemple.

En tant que *substantif* (avec le groupe nominal « le temps réel »), forme plus rare dans ce champ sémantique technique, les deux expressions « temps réel » et « temps différé » désignent simplement le principe technique (de respect ou de non-respect du seuil de latence).

2.1.3. La garantie technique

Les deux expressions « temps réel » et « temps différé » dénotent finalement une qualité technique relative à la *garantie* du respect d'une contrainte temporelle forte et souvent de bas niveau, c'est-à-dire à un niveau proche de la machine. En généralisant, ces expressions relèvent du champ *technique* lorsqu'elles sont situées dans les contextes du Tns (éventuellement à d'autres échelles temporelles) ou de la gestion et du traitement des flux de données (soumis à des contraintes temporelles).

2.2. Le champ pratique

Le deuxième sens, relié aux seuils perceptifs de *l'interaction* (souvent indiqués entre 20 ms et 100 ms en fonction des modalités particulières de cette interaction) relève de la psychologie expérimentale tout autant que du langage courant et se retrouve par exemple dans les expressions « installation (en) temps réel » ou « jeu (en) temps réel ». Ce temps caractéristique de l'interaction et des systèmes de contrôle événementiels correspondants, nous proposons de le nommer « temps réel *pratique* ». De l'autre côté, les autres pratiques associées à l'écriture et à la composition acquièrent un statut autonome dans l'expression « le travail en temps différé » par exemple. Ce temps caractéristique de l'écriture, nous proposons de le nommer « temps différé *pratique* ».

Remarquons que ce champ sémantique se retrouve dans la description de Risset :

> Traditionnellement, la composition est une activité en temps « différé ». C'est l'interprète qui inscrira dans le temps la partition que le compositeur écrit « hors-temps », suivant l'expression d'Iannis Xenakis. La partition, code d'actions et représentation « idéale » ou « virtuelle », permet au compositeur de préparer à loisir un voyage dans le temps sans être assujetti au temps de la musique. Le « temps réel » est celui de l'interprétation – ou de l'improvisation : mais cette dernière activité repose sur une longue accumulation mémorielle de figures musicales associée à des réflexes gestuels. [10, p. 142]

Ou encore, dans la dichotomie entre « création directe » et « création indirecte » proposée par Cadoz [2, p. 168, c'est nous qui soulignons] :

> On peut donc comparer schématiquement deux échelles du processus de création :
> – d'un côté, une forme directe, celle du jeu instrumental, directement impliquée *dans le temps* et dans la matérialité. Le jeu instrumental peut être l'interprétation, mais aussi l'improvisation si l'on utilise ce terme pour désigner l'activité créatrice où l'acte physique (gestuel) s'applique directement aux outils matériels pour produire le résultat final (par exemple musical) ;
> – de l'autre, une forme indirecte, supposant une activité mentale, méditative, en interaction avec des outils de représentation symboliques, théoriques, abstraits, qui se déroule *en dehors du temps propre de l'expérience sensorielle*. Ce que l'on convient de nommer habituellement la « composition ».

2.2.1. Un temps réel pratique interactif

Dans ce champ sémantique, la forme *substantive* « *le* temps réel » prend sa valeur de modalité essentielle de l'interaction, comme temps de l'action ou de la *praxis*, où le geste humain rencontre un retour dans un délai suffisamment court pour permettre son ajustement éventuel. Il s'agit donc du temps intuitif et subjectif du « tout de suite », même si ce temps reste objectivement soumis à un temps de latence du point de vue technique pour permettre la rétroaction dans les faits.

Logiquement, la *locution adjective* (par exemple dans « installation temps réel ») caractérise son sujet en indiquant qu'elle exploite des modalités *interactives*. Cette locution se présente donc comme un synonyme métonymique de l'adjectif *interactif*, désignant avec une légère préciosité la cause – le temps réel technique – pour l'effet – l'interaction. En effet, on a aujourd'hui l'impression que « l'interactivité » a beaucoup perdu de son pouvoir de séduction au profit du « temps réel », dont les racines sémantiques technologiques sont sans doute plus à même de promettre quelque prouesse. . .

2.2.2. Un temps différé pratique révisable

Une *praxis* du temps différé n'est pas en reste et ne constitue pas cette fois une opposition simple, ni une définition négative (un temps de calcul trop lent). Au contraire, le temps différé, en tant que *substantif*, désigne ici positivement les logiciels et les pratiques musicales qui s'inscrivent dans la durée, typiquement les activités compositionnelles. Cela n'interdit d'ailleurs pas aux logiciels utilisés de répondre à des critères temps réel de bas niveau (temps réel technique) ni de posséder des capacités interactives (temps réel pratique), seulement l'usage qui en est fait n'est pas destiné à être écouté « tout de suite » par un auditoire. L'activité compositionnelle n'est en

effet généralement pas envisagée comme une improvisation, ni comme une interprétation, car pour laisser le temps à l'élaboration de se faire, les résultats de la composition sont réservés à être écoutés « plus tard », voire pas du tout, ou encore en creux [9] .

2.2.3. *La réversibilité comme frontière*

Finalement, dans le champ de la pratique, le critère distinctif majeur abandonne la notion technique du *seuil de latence* au profit de la notion pratique de *possibilité de retour* (de réversibilité, ou plus précisément de « révisabilité »). En effet, *l'impossibilité* de retour caractérise absolument le temps de la performance en tant que pratique, où ce qui est fait est fait, définitivement ; réciproquement, la *possibilité* de retour caractérise de façon toute aussi pertinente le temps de l'écriture en tant que pratique [10] .

Autrement dit, il s'agit d'un certain rapport à la flèche du temps et surtout à son irréversibilité : un rapport en phase pour le « temps réel », ce qui lui confère effectivement en ce sens une propriété capitale de la *réalité* avec l'attribut de l'irréversibilité, contre un rapport qui laisse une place à la réversibilité, en différant la diffusion ou l'exécution de l'œuvre, ce qui confère là aussi un sens profond à l'expression « temps différé », sous l'angle d'une *praxis* dont le résultat est effectivement différé, remis à « plus tard ». La *possibilité de retour* devient alors une des modalités qu'on s'autorise ou qu'on s'interdit pour l'activité musicale, et constitue ainsi un point de rupture valide vis-à-vis des pratiques musicales, entre la performance et la composition.

Le cas du *live-coding* pourrait apparaître comme un cas-limite puisqu'il s'agit à la fois d'une écriture (programmation) et d'une performance (en direct). Or, à la lumière du critère de la *possibilité de retour*, ce cas rejoint sans ambiguïté le temps réel (pratique), puisque toute modification validée durant la programmation en direct est irréversiblement donnée à entendre.

La métaphore du couple signal / code illustre schématiquement la distinction entre temps réel et temps différé dans le champ pratique. D'une part, le *signal* est une information éphémère, prise dans un flux – ce qui correspond au temps réel pratique, de l'interaction à la performance. D'autre part, le *code* est une information inscrite, enregistrée et susceptible d'être consultée, utilisée ou révisée ultérieurement – ce qui correspond au temps différé pratique, du paramétrage à la composition.

[9] Par exemple, on peut considérer qu'après plusieurs étapes de traitement successives, un son de 1re ou 2e génération qui n'apparaît pas dans l'œuvre finale existe tout de même « en creux » si un autre son de génération suivante, formé à partir de ce son d'origine, y apparaît.

[10] La pratique du *work in progress*, chère à Pierre Boulez entre autres, franchit d'ailleurs un pas supplémentaire vis-à-vis de cette possibilité de retour dans l'écriture, entraînant un changement d'échelle temporelle dans la notion d'*œuvre*.

2.2.4. *Différenciation évolutive des champs technique et pratique*

Historiquement, si aux débuts du temps réel (technique et pratique, dans les années 80 pour la musique) les programmes et les usages associés à ces programmes pouvaient se confondre dans une correspondance plus ou moins bijective entre le domaine *technique* et le domaine *pratique*, dès que les programmes dédiés au temps différé (pratique) ont bénéficié de fonctionnalités temps réel (technique et pratique, comme la possibilité d'écouter directement un son), cette bijection relative fut logiquement rompue de fait. Ainsi, la croissance exponentielle de la puissance des systèmes informatiques aidant, l'écart entre le temps différé pratique et le temps différé technique n'a cessé de croître, jusqu'au stade actuel où la plupart des programmes d'écriture (au sens large) sont équipés de nombreuses fonctionnalités temps réel (*Csound* en est un exemple typique [11]). Cette évolution des logiciels traditionnellement considérés « temps différé » vers le temps réel suggère bien d'opérer une distinction entre les niveaux *pratique* et *technique*, puisque le temps réel a tendance à se généraliser au niveau technique. Réciproquement, les pratiques temps réel comportent en réalité presque toujours des phases de *préparation* directes ou indirectes, qui relèvent par définition du temps différé (pratique).

2.3. Le champ musical

Le troisième sens, le plus large, exprime le glissement sémantique vers la fonction musicale en tant que choix esthétique. Lorsqu'on lit des expressions comme « une œuvre avec temps différé » ou « improvisation et électronique en temps réel », on comprend qu'il ne s'agit pas de caractériser une technologie ni une pratique mais bien *une musique* dans ses choix esthétiques. Nous proposons de nommer cette partie de notre typologie, selon le contexte : « temps réel *esthétique* » et « temps différé *esthétique* », ou bien « temps réel *musical* » et « temps différé *musical* ».

2.3.1. *Un temps réel musical non-superposable aux autres champs*

Cette dimension *esthétique* ou *musicale* ne s'appuie pas sur la dimension pratique de façon simple ou univoque, de même que la dimension *pratique* identifiée précédemment comporte déjà en elle-même des niveaux d'imbrication entre temps réel et temps différé. Bien sûr, le temps réel *esthétique* ou *musical* implique à la fois les temps réels *pratique* et *technique* (l'expression « une œuvre avec temps réel » fait communément référence aux idées d'interaction, de per-

[11] *« As fast or realtime software synthesis becomes increasingly possible, the free exchange of* instrument *and* score files *becomes worthwhile. »* [15, p. 211, c'est nous qui soulignons].

formance, et aux moyens techniques logiciels et matériels nécessaires à cette interactivité), mais il implique aussi du temps différé ! Et ce à plusieurs niveaux : au niveau *pratique* dans la préparation de la performance, au niveau *technique* aussi selon les programmes utilisés dans cette préparation, voire même au niveau *esthétique*, par exemple lorsque des séquences sonores pré-enregistrées doivent être déclenchées durant la performance ou l'exécution de l'œuvre (ce qui reste très fréquent).

L'extrait suivant peut illustrer ce champ musical, au sens où les idées de temps réel et de temps différé s'affrontent ici manifestement sur un terrain esthétique :

> Le temps réel n'est pas une panacée [...] Le travail d'écriture – au sens large – permet d'approfondir et de personnaliser sa démarche : les propositions du temps réel manquent de variété, et elles incitent à des réactions réflexes et donc irréfléchies, donnant souvent lieu à des stéréotypes vite usés ou périmés. [...] Je tiens à défendre ici la musique sur support, trop souvent critiquée, voire même enterrée dans certains cercles, alors qu'elle permet de composer les sons avec soin et de présenter un résultat musicalement travaillé. [10, p. 145]

En effet, ce sont bien les limites *musicales* du temps réel et les possibilités tout aussi *musicales* du temps différé qui sont pointées par Risset.

2.3.2. Un temps différé musical « sur support »

Curieusement, le temps différé *esthétique* ou *musical* n'implique pas nécessairement le temps différé *pratique*. Il existe en effet quelques cas extrêmes, mais pas si rares, de diffusion d'un enregistrement où ce qui a été enregistré provient directement d'une unique improvisation interactive (par exemple avec des magnétophones). Il s'agit là d'une situation musicale apparemment temps différé et le plus souvent déclarée comme telle (diffusion d'une musique « sur support ») alors qu'aucun travail d'écriture, de réécriture ou de montage n'a été effectué, ce qui tendrait au contraire à justifier d'une pratique « purement » temps réel (ici, une improvisation). . .

Au-delà de ces cas particuliers, le champ *esthétique* de l'expression « temps différé » recouvre souvent l'ensemble des musiques dites « sur support », à comprendre comme les musiques savantes utilisant les technologies numériques et analogiques sous leurs différentes dénominations (concrète, électronique, électroacoustique ou acousmatique), et rejoignant parfois l'idée d'un art des sons *fixés* telle qu'elle a été développée par Michel Chion.

L'utilisation de ce sens *esthétique* transgresse de plus en plus l'ontologie informatique [12] du temps différé, en annexant progressivement les technologies analogiques (pour désigner les concerts de musique concrète par exemple). Cette extension du champ sémantique de la notion de « temps différé musical » va parfois même jusqu'à désigner la musique instrumentale qui utilise des partitions, considérant que ces partitions sont effectivement des « supports » de la musique (mais pas directement supports du son, puisque l'interprétation ne se confond jamais avec la transduction).

* * *

La polysémie du temps réel et du temps différé est avérée, il suffit pour s'en convaincre de lire quelques définitions apparemment incompatibles de la notion de temps réel. Aussi, nous avons proposé de distinguer trois champs sémantiques : le champ technique, le champ pratique et le champ musical. Ces trois champs ne peuvent pas se superposer d'une façon simple, car les notions de temps réel et de temps différé y acquièrent un sens suffisamment différent, individuellement et en rapport, à chaque niveau.

3. DES CONTRADICTIONS CROISÉES

De plus en plus d'observateurs relèvent des contradictions dans différents discours sur le temps réel ou le temps différé, et en particulier sur les logiciels d'informatique musicale. Cette section rappelle donc en premier lieu la valeur paradigmatique qu'ont pris certains logiciels pour le temps réel ou pour le temps différé ; puis elle revient séparément sur des contradictions fréquemment évoquées à propos du temps réel et du temps différé, des idées typiquement proches des énoncés « le temps réel nécessite du temps différé » ou « le temps différé s'opère de plus en plus en temps réel » ; elle propose enfin une explication de ces contradictions à travers l'enchevêtrement des niveaux polysémiques.

3.1. La valeur paradigmatique de certains logiciels

Comme on peut le constater d'après les sections précédentes, les notions de temps réel et de temps différé restent relativement floues en général, jusqu'à ce qu'elles s'incarnent dans des logiciels particuliers, dont certains prennent alors une valeur paradigmatique. Ainsi, *Max/MSP* et ses cousins *jMax* et *PureData* – ou plus simplement le paradigme « Max [13] »

[12] Il ne s'agit pas ici de l'ontologie informatique « technique » telle que l'a définie Thomas Gruber en 1993 dans son article *Towards Principles for the Design of Ontologies Used for Knowledge Sharing*, comme la spécification d'une conceptualisation d'un domaine de connaissance, mais bien d'une ontologie au sens philosophique, en posant l'informatique comme condition préalable à l'existence des notions de « temps réel » et de « temps différé ».

[13] « *I have worked for many years on a computer environment for realizing works of live electronic music, which I named in honor of Max Mathews. Three currently supported computer programs – Max/MSP, Jmax, and Pd – can be considered as extended implementations of a common paradigm I'll refer to here as "Max." [...] The Max paradigm*

– viennent souvent s'opposer à *OpenMusic*, son ancêtre Patchwork ou encore la série des Music N dans le discours, précisément par rapport à la notion de temps réel :

> *The semantic of OpenMusic (and Patchwork) is one of demand-driven dataflow. Each object is essentially a function call, which recursively evaluates its inputs, precisely as a Lisp form is evaluated. Compared to Pd, the relationship between data and process is reversed. There is no notion of real-time events or even of real time itself ; rather, the contents of a patch are static data.* [8]

Notons qu'il en va de même à propos des logiciels du commerce. Ainsi, *Live* est régulièrement associé au temps réel dans le discours, à commencer par son nom qui fait évidemment référence au « *live* » tel que ce mot s'est intégré dans le vocabulaire, y compris en France, à partir des expressions « *live music* » et « *live electronic(s)* ». En opposition, les logiciels de montage (comme ProTools) ou encore d'édition de partition (comme Finale) sont volontiers associés au temps différé.

3.2. Le temps réel nécessite du temps différé

L'utilisation du logiciel *Live* mise en avant est incontestablement la situation de concert, même si ses capacités de montage sont parfois évoquées ; *Live* serait en quelque sorte « le » logiciel temps réel par excellence. Or, des observateurs de plus en plus nombreux font malicieusement remarquer que la part du temps différé dans *Live* est *primordiale*, puisque son fonctionnement repose essentiellement sur l'utilisation de fichiers audio pré-enregistrés ou de séquences Midi. . . Ce cas particulier souvent repris reflète une critique bien plus large, qui dénonce en réalité un abus de langage quant au « temps réel » :

> Ce qu'on appelle alors « temps réel » dans la composition musicale serait un abus de langage puisqu'une part des composants musicaux est souvent déjà fixée, et n'a pas pour vocation à varier d'une interprétation à l'autre. [4, p. 8]

Cette critique de Philippe Manoury se fonde sur son analyse du temps réel à partir de l'interprétation musicale, et pointe la *fixité* des éléments préenregistrés, qui entre en contradiction avec l'interprétation comme *variabilité*.

On peut ajouter à cette remarque le fait que ces logiciels dits « temps réel » demandent paradoxalement un temps de préparation souvent important. Dans le cas de *Live*, il s'agit de la préparation des configurations. Dans celui de *Max/Msp*, il ne s'agit de rien de moins que de la *programmation* d'un patch [14] , ce qui peut demander un temps tout à fait considérable. . . Or, de toute évidence, cette préparation relève du temps différé (au sens pratique), en se rapprochant de la composition : le choix de séquences et la constitution de configurations verticales potentielles dans le cas de *Live*, comme la programmation de circuits sonores réactifs et la mise en correspondance des paramètres dans le cas de *Max/Msp*, sont autant de choix compositionnels et d'écriture, c'est-à-dire de codes enregistrés. Surtout, le résultat de ces codes préparés n'est pas destiné à être entendu sur-le-champ, mais bien *ultérieurement*, ce qui les rend *révisables* et constitue une marque indubitable du temps différé.

3.3. Le temps différé s'opère de plus en plus en temps réel

Concernant la synthèse sonore, il ne fait aucun doute que la synthèse en temps différé se raréfie au profit de la synthèse en temps réel. De fait, *Csound*, l'héritier des langages *Music* N à forte valeur paradigmatique temps différé, est déjà passé à sa version temps réel en 1990. Les bénéfices de cet accès au son en temps réel ont d'ailleurs été largement plébiscités :

> *As performance increases, things that were impossible become possible. [...] When you can create sound in real time you can interact with it, and interaction makes possible things that never could have been achieved by typing in lists of commands.* [5]

Mais le temps différé ne se limite pas à la synthèse. Par exemple, Miller Puckette a expliqué récemment comment, contrairement à la synthèse, la Cmao [15] n'est pas devenue temps réel :

> *In general, the cost of CGM per audio sample has not remained constant, but has not grown quickly. The best CGM of the seventies, thirty years ago say, probably cost less than ten thousand arithmetic operations per sample of output. The speedup in computing in the intervening years has allowed us the luxury of running the classic CGM algorithms in real time, thus changing the nature of the pursuit fundamentally. The algorithms themselves have grown somewhat more complex as well,*

can be described as a way of combining pre-designed building blocks into configurations useful for real-time computer music performance. » [7]

[14] Le préjugé selon lequel programmer en *Max* serait facile s'avère souvent faux : cela dépend du type d'application à réaliser, et surtout de la taille du patch. Miller Puckette, s'il ne reconnaît pas *Max* comme un véritable langage informatique, reconnaît néanmoins que « faire un patch » s'apparente bien à une activité de programmation ; seul le niveau d'abstraction change : « *a patch is itself a program, and the job of connecting simple functions together to make larger ones can be thought of as programming. In this sense, the trend toward patch-based software can be seen as shifting the level on which the user programs the computer away from the C or Lisp code itself and into the "language" of patches. It may be that this is fundamentally a better level at which to operate a computer, than either that of code or that of the user of a large, monolithic program such as a database application.* » [9]

[15] Miller Puckette emploie *CGM* (*Computer Generated Music*) pour « synthèse sonore » et *CAC* (*Computer Aided Composition*) pour Cmao (Composition Musicale Assistée par Ordinateur).

but the universal preference for real-time synthesis and processing has put a lid on that growth.

CAC has not become real-time at all. Since it appeals to the composer in us (whereas CGM appeals to the performer in us), it seems reasonable to expect that CAC software will continue to absorb all the computing resources that can possibly be brought to bear on it. Anything allowed to grow will naturally do so. [9]

Si actuellement les ressources de calcul ne permettent pas à la CMAO de se faire en temps réel, comme l'indique Miller Puckette, il n'en reste pas moins qu'un degré d'interaction de plus en plus important apparaît dans les applications temps différé. Signalons par exemple les travaux de l'équipe « Représentations musicales » de l'IRCAM sur l'intercommunication entre *OpenMusic* et *Max/MSP*, dont la bibliothèque OSC *Moscou* pour *OpenMusic* [13, p. 36, avec Carlos Agon] et le projet *OMax* [16] , qui fait collaborer *Max/MSP* et *OpenMusic* pour former un « co-improvisateur » par « réinjection stylistique ».

3.4. La confusion des niveaux polysémiques

Deux types de contradictions ont été relevées : un temps réel qui comporte en réalité une grande part de temps différé, et réciproquement un temps différé qui comporte à son tour une part importante de temps réel. Or ces contradictions peuvent être levées après un examen des niveaux sémantiques.

Les reproches avancés à l'encontre de *Live* par exemple, c'est-à-dire la fixité des séquences employées et le temps important de préparation, appartiennent au niveau du temps différé *pratique*, ce qui n'empêche pas le logiciel *Live* d'appartenir au niveau du temps réel *musical*, en accord avec son utilisation musicale de type « performance ».

De la même façon, les nouvelles fonctionnalités interactives d'*OpenMusic*, appartenant au niveau du temps réel *pratique*, n'invalident pas la catégorisation consensuelle de cet environnement dans le niveau du temps différé *musical*, en accord avec son utilisation musicale de type « composition ».

En réalité, à bien y regarder, un logiciel destiné à la performance aura d'autant plus de chances d'être opérationnel et efficace en situation de direct (temps réel technique, pratique et musical) qu'il aura bénéficié d'une préparation importante au préalable (temps différé pratique). Inversement, un logiciel destiné à la composition aura d'autant plus de chances d'être opérationnel et efficace dans une situation d'écriture (temps différé pratique et musical) qu'il offrira des fonctionnalités interactives (temps réel technique et pratique). . .

* * *

On peut retenir deux points importants de cette analyse des contradictions. D'une part, les contradictions apparentes entre le temps réel et le temps différé proviennent souvent d'une confusion croisée entre les différents niveaux polysémiques de ces deux notions. D'autre part, on peut constater que c'est le niveau *musical* qui fait finalement le plus consensus dans notre domaine de l'informatique musicale, c'est-à-dire le niveau le plus abstrait.

[16] *« OMax is a software environment which learns in real-time typical features of a musician's style and plays along with him interactively, giving the flavor of a machine co-improvisation. OMax uses OpenMusic and Max. It is based on researches on stylistic modeling carried out by Gerard Assayag and Shlomo Dubnov and on researches on improvisation with the computer by G. Assayag, M. Chemillier and G. Bloch (Aka the OMax Brothers) in the Ircam Music Representations group. »* `http://recherche.ircam.fr/equipes/repmus/OMax/`

4. REFERENCES

[1] Christian Bonnet and Isabelle Demeure. *Introduction aux systèmes temps réel*. Hermes, Paris, 1999.

[2] Claude Cadoz. Continuum énergétique du geste au son — simulation multisensorielle interactive d'objet physiques. In *Interfaces homme-machine et création musicale*, chapter 8, pages 165–181. Hermes, Paris, 1999.

[3] D.J. Levitin, K. MacLean, M. Mathews, L. Chu, and E. Jensen. The perception of cross-modal simultaneity. *Proceedings of International Journal of Computing Anticipatory Systems*, 1999.

[4] Philippe Manoury. Considérations (toujours actuelles) sur l'état de la musique en temps réel. *L'étincelle*, 3, nov 2007.

[5] James McCartney. A few quick notes on opportunities and pitfalls of the application of computers in art and music. In Hatje Cantz, editor, *CODE – The Language of Our Time*, Linz, sep 2003. Ars Electronica.

[6] Henri Piéron. Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices. *L'année psychologique – Revue de psychologie cognitive*, 20, 1913.

[7] Miller Puckette. Max at seventeen. *CMJ*, 26(4) :31–43, 2002.

[8] Miller Puckette. A divide between 'compositional' and 'performative' aspects of Pd. *First International Pd Convention*, 2004.

[9] Miller Puckette. Computing while composing. In Carlos Agon, Gérard Assayag, and Jean Bresson, editors, *The OM Composer's Book*, volume 1. Delatour France / IRCAM, 2006.

[10] Jean-Claude Risset. Temps en musique numérique. In *Le temps en musique électroacoustique*, volume Actes V, pages 141–145, Bourges, 1999. Académie Bourges, Mnemosyne.

[11] Curtis Roads. Des instruments pour un son organisé. In *Composition et environnements informatiques*, pages 107–125. IRCAM – Centre Georges-Pompidou, Paris, automne 1992. Traduit de l'anglais par Serge Grünberg.

[12] Curtis Roads. *L'audionumérique*. Dunod, 1998. Version française : Jean de Reydellet.

[13] Carl Seleborg. Interaction temps-réel/temps différé – Élaboration d'un modèle formel de Max et implémentation d'une bibliothèque OSC pour OpenMusic. DEA ATIAM, Université Aix-Marseille II, Paris, jun 2004.

[14] Charles Spence and Sarah Squire. Multisensory integration : maintaining the perception of synchrony. *Current Biology*, 13(13) :R519–21, jul 2003.

[15] B. Vercoe and D. Ellis. Real-time csound : Software synthesis with sensing and control. In *Proc. ICMC 1990*, pages 209–211, Glasgow, 1990.