



Les espaces sonores des communications visiophoniques sur mobile

J. Morel & C. Licoppe

Objet

Distinguer les espaces sonores des communications visiophoniques sur mobile

Voix des participants. Comptent sur l'intelligibilité mutuelle de la parole, quelles que soient les « distorsions » générées par « l'artefact communicationnel » (Fornel, 1994).

Médiatisation du réseau 3G. Modifie, via des phénomènes de latence (délai de transmissions, décalage voix/image), les conditions de production et de réception des actions des participants /au face-à-face.

Dé-placement du mobile. Possibilité de visiophoner de presque n'importe où = exposition augmentée à différentes « espèces d'espaces » sonores se prêtant +/- à l'interaction distante.

Ressources non verbales. Rôle structurant pour l'organisation des échanges verbaux (via regards, orientations corporelles, gestes...) & pour favoriser l'écoute mutuelle /environnements sonores.

Pour une approche d'inspiration « naturaliste » (Goffman, 1973, Kendon 1990, etc.)

Décrire « *rigoureusement empiriquement et formellement* » (Schegloff, Sacks, 1973)

Le déroulement des communications

Les ressources, procédures utilisées

Les ancrages, engagements locaux des participants

Les composantes pertinentes des contextes

Comment mobilisent-ils des ressources conversationnelles, visuelles, gestuelles... pour organiser les communications ? Selon des ancrages spatiaux, sociaux, sonores... susceptibles de devenir des composantes des échanges ? Et d'in-former, dans une relation réflexive et indexicale, les lieux de communication ?

Méthodes



Constitution d'un corpus de communications visiophoniques à partir de avec 2 protocoles d'enregistrement vidéo.

1 Lunettes-caméra avec liaison sans fil vers enregistreur. Optique centrale & microphone intégré.

2 Exploitation de la sortie audio/vidéo disponible en série sur plusieurs modèles de mobile. Lorsque la sortie audio n'est pas active pour la visio, utilisation d'un micro-cravatte pour capter le flux, mixage au niveau de l'enregistreur.

Corpus actuel : près de 100 communications enregistrées par 8 personnes, depuis 1 an (2 mois d'enregistrement en moyenne par individu).



Thèmes développés

Des espaces sonores sensibles & de différentes natures

Attentes normatives concernant **l'écoute mutuelle** dans une interaction focalisée : compter sur un dispositif technique sensé supporter l'échange de parole.

Une **médiatisation par le réseau** et les temps de traitement des terminaux générant des **délais entre production/réception** des actions et une **désynchronisation voix/image** : les interactants se montrent particulièrement attentifs à leurs conditions participation, en particulier au niveau sonore (demandes de vérifications chroniques du son, Mondada, 2007).

Des environnements sonores « mobiles » : pas nécessairement de territoire pré-défini pour la vidéocommunication. **Expositions** à différents à **des lieux variables, sensibles** et plus ou moins aisément **maîtrisables** (lieux publics).

Juxtaposition de mondes sonores locaux/distants : si les participants d'une interaction visiophonique nourrissent des attentes quant à l'écoute mutuelle, celle-ci peut aussi concerner leurs **entours sociaux**.

Equation : garantir l'intercompréhension tout en ne devenant pas une source de nuisance sonore /co-présence publique notamment.

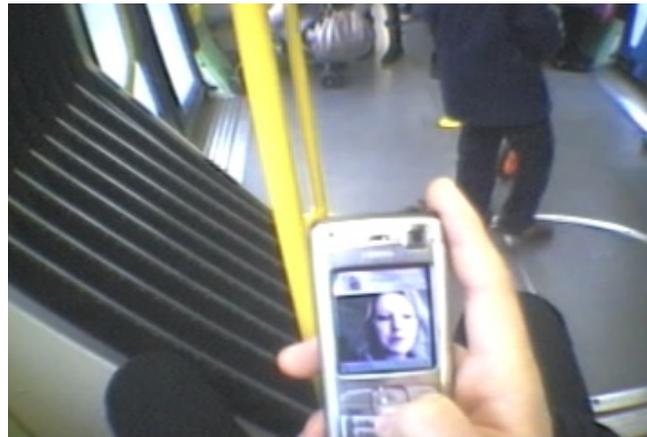
- I. Comment des bruits deviennent des interférences sonores (p. 6-8)
- II. La communication impossible (p.9-14)
- III. Seuils d'audition et effets de grossissement (p.15-18)
- IV. Conclusion (p.19)
- V. Bibliographie (p.20-21)
- VI. Convention de transcription (p.22)

I. Comment des bruits deviennent des « interférences sonores » (Goffman, 1973*)

Voici une « vue subjective » (lunettes-caméra) d'un visiophoniste (A) assis dans un tramway. Le champ de la caméra suggère qu'il regarde en direction de l'écran/la cam. Orientée de façon analogue, la jeune femme, visible sur l'écran principal du mobile, co-produit un « talking heads »** qui est le format standard des échanges visiophoniques sur mobile. En réalisant des cadrages visage, en entretenant des périodes de regard mutuel tout en étant engagés dans des séquences conversationnelles, ils élaborent une interaction focalisée via l'exploitation de différentes ressources verbales et non verbales entrelacées.

Toutefois, les localisations des participants dans des environnements toujours sujets à variation les exposent à différentes composantes du contexte, susceptibles de « rentrer » dans le cadre interactionnel visiophonique.

Si une telle possibilité concerne en fait tous les lieux susceptibles d'être fréquentés au quotidien (domicile, travail, etc.), les lieux publics offrent certainement une plus grande diversité de co-présence, de mouvements, de bruits, etc., +/- aisément maîtrisables.



* Pour une discussion sur les usages possibles de la notion « d'interférence sonore » développée par Goffman, voir Morel, 2008.

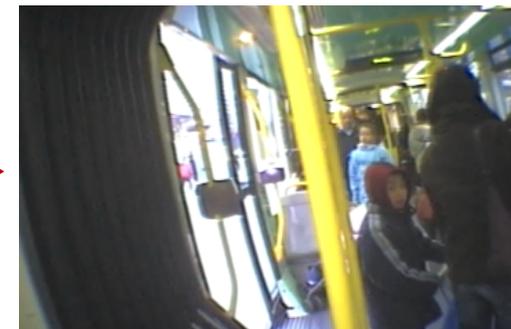
** Champ de la caméra frontale, port du mobile à bout de bras et entretien d'une interaction focalisée conduisent le plus souvent les participants à produire une réciprocité de portraits.

Pour attribuer le statut « d'interférence » à quelque bruit, il convient de se référer aux opérations et procédures à partir desquelles les participants rendent compte de leur perception selon leurs activités en cours.

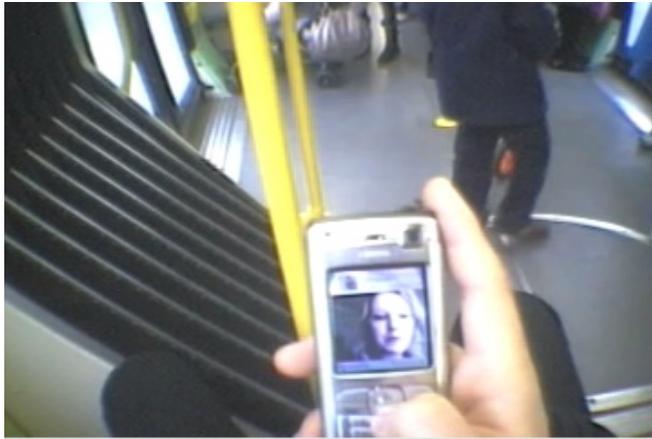
20. A : pourtant il a un téléphone trois g::: (0.8) et
 21. j'y arrive pas.
 22. (0.1)
 23. B : et est ce qu'il est en trois (g)/
 24. XXX
 25. []
 26. A : attend: deux s'condes (.) y a la sonnerie là
 A^(**)a-----m-----
 27. XXXXXXXXXXXXXXXXXXXXXXX
 A^(**)-----
 28. (1.0)
 A^(**)---r---
 29. A : j' suis bi entôt arrivé d'façon donc ° euh: bon. °
 A^(**)-----
 30. (0.7)
 31. B : (inc)
 32. A : bah voilà:

Grâce aux enregistrements vidéo, il est possible de transcrire finement – en fait, simplement à « l'échelle ordinaire » dit Schegloff, 1988 – l'ensemble des procédés conférant tel ou tel statut à des bruits (ou une somme de bruit) issus des environnements des interactants, d'une certaine façon « mutualisés » dans le cadre visiophonique. Un signal sonore (les X) de fermeture des portes du tramway retentit en l.24, pendant que B (la jeune femme) parle (l.23). Son tour de parole est chevauché par le signal sonore, et on constate qu'il n'est pas entièrement mené à son terme (« g » quasi imperceptible + interruption (/))

Si la fin de l.23 peut s'expliquer par la perception (distante) du signal, A chevauche également ce tour pour interrompre et suspendre « officiellement » l'interaction conversationnelle. Ceci est réalisé en l.26, via 2 unités de construction de tour (SSJ, 1974) : un premier segment réalisant et projetant l'interruption, suivi de l'identification de la source d'interférence sonore (« y a la sonnerie là »). Ces procédés conversationnels sont soutenus par une modification de l'orientation visuelle de A : Pendant l.26 il amorce (a) un mouvement ascendant de tête, atteignant son point maximum (m) → lorsqu'il énonce la source de nuisance. L'amorce du retour (r) en talking heads débute pendant le silence de l.28 (le bruit a cessé), et c'est au moment précis où il s'engage dans un nouveau tour de parole (l.29) que le regard mutuel est rétabli (dernier smiley).



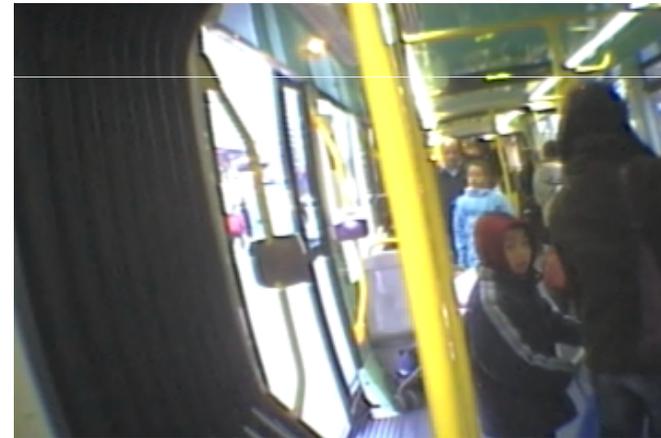
« y a la sonnerie là »



La nature d'un bruit, son intensité, son caractère répété, etc., sont des critères d'évaluation mesurables qui restent insuffisants pour statuer (à un niveau général) sur la perception et l'interprétation qu'en font des personnes à portée de champ auditif. Certains co-présents (du tramway) ont probablement ignoré le signal sonore, d'autres l'ont exploité en tant que ressource (pour se presser de rentrer/sortir par ex.), tandis que les visiophonistes lui ont pratiquement conféré le statut d'interférence sonore.

Altération de la conception de tours de parole, procédures d'interruption/suspension et reprise lors du retour à un (relatif) silence local, sont quelques-unes des techniques conversationnelles à partir desquelles les visiophonistes règlent et gèrent des bruits qu'ils dotent de statuts interférants.

Leur aptitude à se voir en commun, leurs attentes concernant le cadrage (talking heads) rentrent également en ligne de compte, dans un travail pratique d'ajustement constant à des mondes sonores sensibles et contingents.



Une échelle d'analyse suffisamment attentive au « détail » des actions verbales et non verbales des participants ne peut manquer de relever leur « liaison inextricable » (Heath & Luff, 1991), organisant l'interaction distante suivant différentes composantes locales.

II. La communication impossible

Il arrive que les conditions acoustiques soient plus continûment défavorables pour le développement d'échanges visiophoniques sur mobile.

La mobilité expose les participants à des mondes sonores se prêtant +/- bien à ce format d'interaction distante, et en particulier pour les personnes jointes, qui ne disposent que du « temps de la sonnerie » pour s'éloigner (lorsqu'elles le peuvent) de sources de nuisance sonore. Voici un exemple de conditions acoustiques « extrêmes » pour le développement d'une communication visiophonique, et une série de mesures pratiques visant à les surmonter.

Les lieux & les participants

Un appel passé depuis le RER (A, le jeune homme) et reçu dans le métro (par B). Les fonds sonores des interlocuteurs sont importants, en raison du bruit généré par les moyens de transport et une multitude de bruits d'origine humaine perceptibles sur l'enregistrement. Il s'agit d'un couple s'appelant souvent, brièvement, et notamment lors de périodes de transports (dès lors il ne s'agit pas vraiment de « temps morts »*).

Le silence (l.1) faisant suite à la mise en image est immédiatement exploité (via les pré-cadrages) pour créer une période de regard mutuel favorisant la reconnaissance de l'orientation des participants dans le cadre interactionnel partagé.

Comme souvent, c'est l'appelant qui commence à parler (l.3) avec une identification comprenant un terme d'adresse « recipient design », tour est chevauché par une demande de localisation projetant un cadre précis de passation d'appel : l'inférence est vraisemblablement liée à des savoir partagé (temporalité « encore ») + visualisation possible de l'arrière-plan de A (paysage défilant rapidement).

Le segment incompréhensible et la suspension du regard mutuel sont les premières manifestations de problèmes d'écoute et d'entretien d'une écologie « oeil-à-oeil » marquant finalement l'ensemble de la communication.

Une ouverture rapidement marquée par un problème d'écoute

1. (1.7)
2. ☺
3. A : mon ti bébé ?
4. []
5. B : t'es encore dans l'métro: ?
6. A : (0.5) mouais ça (inc) ☺
7.

En l.6, l'appelant fournit un début de réponse, mais la suite de son tour est incompréhensible (inc) ; on note également la fin de la période de regard mutuel, en raison d'une baisse subite de luminosité dans l'environnement de l'appelée (probable tunnel).

☺ = regard mutuel
☹ = A et/ou B ne regarde(nt) pas vers écran/cam

La durée des périodes de regards mutuels/non mutuels est exprimée via l'alternance des smileys et relativement aux tours.

* Pour une ethnographie soulignant comment les temps de transport perçus comme « perdus » sont en fait le moment de bien des activités, voir Meissonnier, 2001.

Des problèmes d'écoute limitant les possibilités de sortir de l'ouverture

Le segment incompréhensible et le silence qui suit sont traités par B comme une absence, puisqu'elle produit une sommation sur un mode téléphonique.

Conformément à la sommation, B prend la parole, mais pour formuler à son tour une demande de vérification du son à laquelle ne répond pas B au tour suivant.

6. A : (0.5) mouais ça (inc) ☹️
7.
8. (0.6)
9. B : allô: ?
10. A : (1.9) ←
11. B : ti bébé: ? ←
12. A : : (0.6) tu m'entends ?
13. B : (3.2)
14. ☹️ ☹️
15. A : tu as ton/ (cahier des charges)
16. (2.5)
17. B : j't'endends ma:l

En l.10, A ne fournit pas la seconde partie de paire attendue et B reproduit une sommation comprenant un terme d'adresse terminé par un contour interrogatif.

En l.15, A s'engage vraisemblablement dans un autre type de tour de parole, très partiellement compréhensible.

C'est pourquoi le silence de l.16 n'est pas attribué à B, qui prend la parole (par auto-sélection ?) au tour suivant pour mettre explicitement en relief le problème d'écoute.

A ce stade, plusieurs segments sont incompréhensibles, les locuteurs tentent via des sommations ou des demandes de vérification, de rétablir des conditions d'écoute mutuelle, jusqu'à la mise en relief de problèmes de compréhension. Si une période de regard commun est établie pendant le silence de l.13, elle est rapidement suspendue.

De l'impossibilité d'aller plus loin

En fait, les environnements acoustiques des participants sont tels qu'ils ne parviendront pour ainsi dire « jamais » à s'engager dans des séquences traitant d'autre chose que le problème d'écoute. Voici quelques extraits montrant leurs difficultés, les procédés à partir desquels ils tentent de les surmonter, la lettre associée aux smileys indiquant que c'est surtout B qui ne fournit pas une orientation visuelle dirigée vers l'écran/caméra

53. A : t'as dit quoi ?

54. B : (9.6)

55. ☹️ ☹️ (B)

56. A : t'es où là exactement

57. B : (1.9) quoi ?

58. A : t'es où là exactement

Une demande de répétition suivie d'un long silence attribuable à B.

Une répétition d'une requête précédemment formulée, suivie d'une demande de répétition, et d'une répétition.

71. B : hein ?

72. A : (0.7) REPETE

73. B : (3.3)

74. A : T'ES OU LA

Une manifestation d'une incompréhension suivie d'une demande de répétition appuyée (PARLER FORT), aboutissant à une absence de répétition attendue, puis à une répétition d'une demande de localisation.

76. B : (3.6) franchement ti/ tu sais quoi ti coeur (.)

77. ☹️ ☹️ (B)

78. rappelle moi en::: en normal pace que là j'entends même

79. pas et ça sert à rien

80. (2.5)

81. B : ti bébé ?

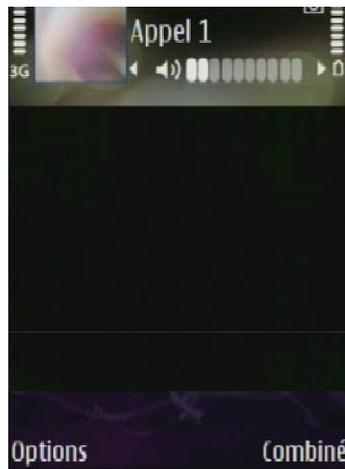
82. A : (0.3) ((A met fin à la communication))

En l.76, B s'engage dans un tour relativement étendu, où elle fournit un premier élément permettant une orientation vers la clôture, en proposant un rappel « en normal » et en l'associant aux nombreux problèmes d'écoute rencontrés.

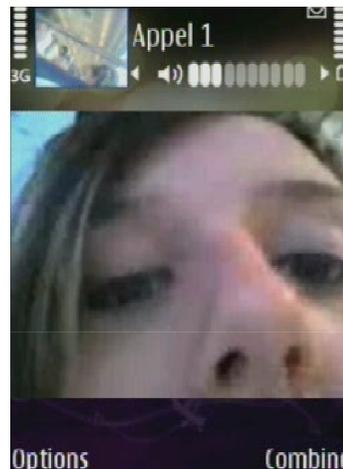
En l.80, A ne se sélectionne pas pour valider une telle orientation, mais il la réalise de façon « brutale » en l.82, en ne participant pas aux séquences collaboratives permettant de mettre un terme à l'interaction, et alors que B a tenté en l.81 d'obtenir une contribution verbale de B. Tout cela se fait sans que B regarde vers l'écran...

Conséquences sur l'entre-deux visuel

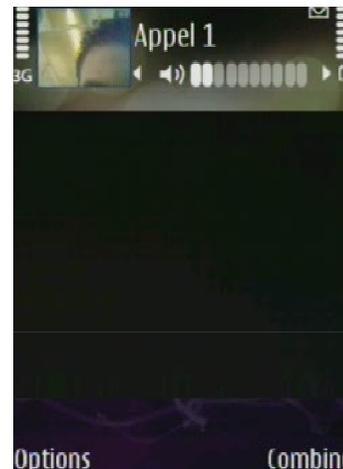
Les difficultés rencontrées par les participants pour garantir la production/réception de leurs tours de parole ne sont pas sans conséquences quant à l'orientation de la partie haute de leur corps, leur activité visuelle et leur capacité à entretenir un regard commun. Pour faire face à leurs environnements sonores particulièrement défavorables, ils mettent en place différentes parades corporelles visant à maximiser le passage ou la réception de la parole. 4 configurations sont possibles :



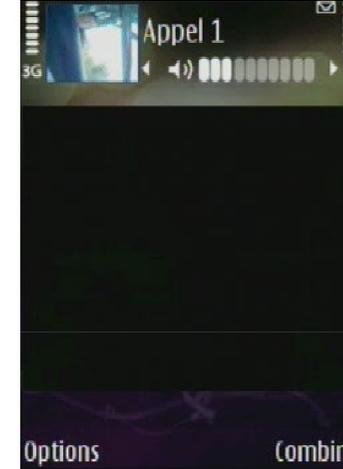
Un haut, A « prête l'oreille » pour accéder à un tour de parole actuellement délivré par B. Cette dernière est également « collée » à son mobile pour se faire attendre, et l'image noire est ici générée par ses cheveux ôtant toute luminosité.



Lorsque c'est la jeune femme qui cherche à rétablir un contact visuel, ce peut être son interlocuteur qui a disparu du cadre pour mieux se faire entendre, écouter ou pendant un silence.



La configuration inverse, où A parle et revient progressivement en talking heads, tandis que B est « collée » à son mobile.

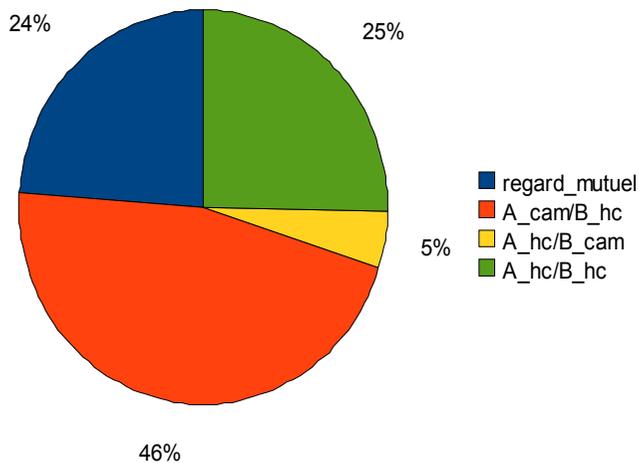


Il arrive également que ni l'un ni l'autre ne soit présent dans le cadre pour maximiser la production/réception de la parole.

Distribution des regards et des temps de parole

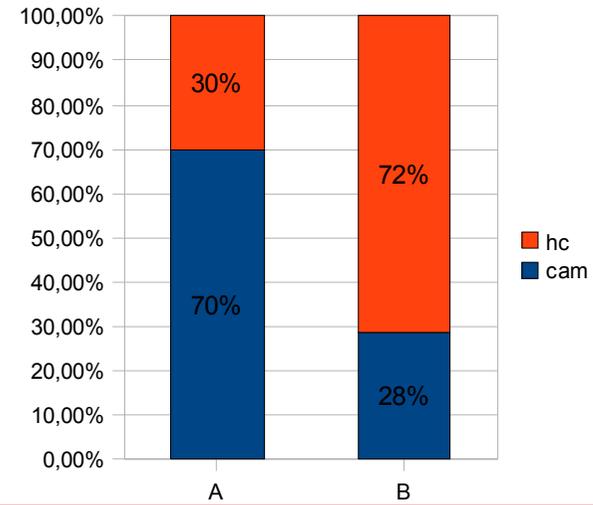
Le logiciel d'indexation vidéo utilisé (ELAN) permet de produire des indicateurs statistiques montrant autrement les effets d'expositions sonores.

1 Ci-dessous, le graphique montre que les périodes cumulées de regard mutuel ne représentent ici qu'1/4 (■) du « temps de communication disponible » (TCD), une proportion équivalente à la somme des moments où A et B regardent simultanément hors champ ou ne sont simplement pas visibles (■). Les deux autres proportions rendent compte de l'asymétrie l'un regarde l'écran/pas l'autre, et c'est surtout B qui tend à ne pas être orientée visuellement vers le mobile (■)



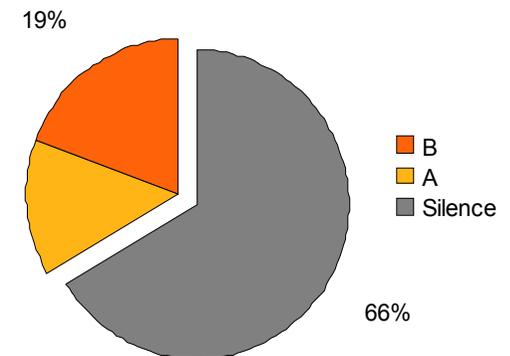
V =	regard_mutuel	27,24	23,87%
	A_cam/B_hc	52,62	46,11%
	A_hc/B_cam	5,24	4,59%
	A_hc/B_hc	29,02	25,43%
	com_dispo	114,12	

Les proportions de ces différents indicateurs sont généralement plus importantes au regard du corpus, en particulier pour les périodes cumulées de regard mutuel. Toutefois, il n'est pas pertinent d'avancer de valeurs moyennes, ces « temps » variant beaucoup selon les développements thématiques, leur absence, des échanges non verbaux « remplissant » de longs silence par ex.



2 Ceci se vérifie au niveau de la distribution des regards cam/hors champ par individu : pendant moins du tiers de la communication, B regarde vers l'écran, la proportion étant inverse pour A.

3 Les temps de parole cumulés des participants sont également faibles puisqu'ils parlent pendant 34% du TCD



III. Seuils d'audition et effets de grossissement

Lorsqu'on examine l'ensemble du corpus, les conditions de la communication précédentes sont, d'une certaine façon, ce qui se fait de pire : visiophonistes en déplacement rapide (transports en commun), fonds sonores respectifs bruyants à très bruyants, absence d'oreillette, luminosité changeante, variations du réseau, etc., et bien qu'ils privilégient visiblement le canal sonore (Tang & Isaacs, 1993), ils éprouvent de grandes difficultés pour converser.

Les risques d'expositions sonores ne se réduisent pas à l'entre-deux visiophonique. En vertu de leur localisation, ce sont parfois les interactants distants qui ont à limiter le volume de leurs contributions verbales selon les régimes de co-présence publique. L'analyse suivante permet de montrer cette autre facette des risques « d'interférences sonores ». Ces dernières concernent différents cadres de participation, et génèrent parfois des modifications du cadre préférentiel de la visiophonie mobile, le talking heads.

Les lieux & les participants

Un appel passé depuis le lieu de travail et reçu dans le bus.

Le fond sonore de l'appelante (A) est totalement calme, et sensiblement moins pour la personne jointe dans le bus (B). Il s'agit d'amis habitant à quelques kilomètres l'un de l'autre, se voyant assez souvent, et se passant régulièrement des appels visio « pour se voir » (à distance) dira l'appelante. Pas d'oreillette pour la personne jointe.

Ouverture

La relation indexicale entre certaines variables environnementales et la structuration de l'interaction s'observe dès l'ouverture de la communication.

Le premier regard mutuel est important pour manifester une orientation dans le cadre visiophonique. Par le rire (l.3), l'appelante rend compte de la réception de l'image, qui est ensuite traité par B avec la délivrance d'une salutation. La structuration de cette ouverture repose beaucoup sur la dimension visuelle.

En l.8, formule une demande de localisation ne recevant pas immédiatement de réponse B faisant à ce moment les « gros yeux ».

Après un nouveau rire, A formule 2 demandes de localisation basée sur la catégorie transport en commun en vertu d'un arrière-plan déchiffrable (fenêtre).

L'inférence est s'avère exacte et...

c'est alors que A propose immédiatement un rappel en vertu de la connaissance dont elle dispose quant à des lieux +/- favorables de communication visio. Comme nous l'avons vu *supra*, les bruits générés par certains transports en commun (+ les co-présents), peuvent considérablement affecter la compréhension des tours de parole.

1. dring
2. (1.2)
3. A : oua::::: ah ah ah ah ah ah ah ah ah hh hh hh hh hh hhhhh
4. []
5. B : (salut)
6. A : hhhhh hhhhhh (0.5) ça va ?
7. B : (2.4)
8. A : t'es où: ?
9. B : (0.7)
10. A : hu hu hu hu hu (.) hhhhhh (0.2) t'es dans l'métro ?
11. B : (0.6)
12. A : t'es dans l'bus ?
13. B : (0.5) allô ?
14. A : (0.7) t'es dans l'bus ?
15. B : (1.4) oui: je suis dans l'bus là h.h.
16. (0.3)
17. A : bah alors je raccroche alors
18. []
19. B : (t'en) rates pas u:ne !



Le recadrage ascendant initial de B, à 0.2 et 1.2 sec après la mise en image mutuelle, faisant rire A.



Les « gros yeux » de l.9.

On notera enfin que c'est surtout B qui met fin aux périodes de regard mutuel, à proximité du début de ses tours de parole...

Quand le parler « près/bas » génère des effets de grossissement

Le constat précédent de désengagements visuels de B, généralement réalisés à proximité du début de ses tours, sont en fait liés à une double contrainte : favoriser la réception de la parole tout en limitant leur caractère publiquement audible (co-présence du bus). Pour ce faire, il se rapproche du microphone juste avant de parler, ce qui produit un effet de grossissement traité par son interlocutrice.

C'est en l.72 que A traite pour la première fois le grossissement qui ne correspond pas au mode portrait standard.

Il s'ensuit des tours mêlant rire, « menaces » et

contestation de la « distorsion » (Fornel, 1994) aggravée par la technique du parler près/bas.

71. B : où ça:: à nation: ?
 72. A : (0.5) nan nan:: (.) ah j'vois qu't'on nez: (.) c'est
 73. trop bon: hu hu hhhhhh hu hu hu
 74. (0.5)
 75. A : hu hu hu hu >hu hu hu< °hu hu hu° hum hu hu
 76. [] []
 77. B : **quoi ! h.h.h.** **ça** te fait marrer:
 78. et tu/ (.) éh mumu tu vas voir:
 79. (.)
 80. A : hu hu hu !
 81. []
 82. B : ah ah ah ah hhhhhh et un p'tit (.) ah::: p'tit (nez:)



Ci-contre, le rétablissement (normativement attendu) du mode talking heads, après le traitement du grossissement. La durée de ce mode portrait est figurée en rouge sur la transcription. Seulement, lorsqu'il poursuit son tour (2^{ème} UCT de l.77), B se rapproche de nouveau du mobile pour en améliorer le caractère audible, et de nouveaux grossissements succèdent.

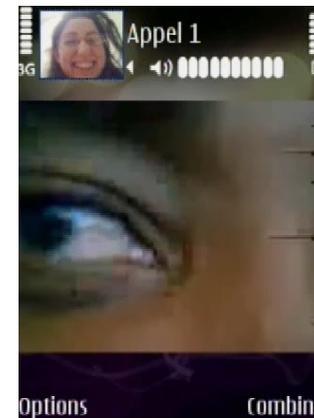
Quand le parler « près »...

B semble ne pas pouvoir faire autrement que de se rapprocher de la caméra pour favoriser le passage de la parole, et parfois écouter, ce qui conduit A à traiter une seconde fois mais différemment l'effet de grossissement.

Le placement du tour où est traité une seconde fois le grossissement n'est plus uniquement d'ordre humoristique. Il s'inscrit dans une logique de pré-clôture initiée en l.91, et il se présente comme un élément de justification pour une telle orientation conversationnelle.

En l.101, B valide la proposition de rappel et pendant que A la répète pour obtenir une confirmation définitive, B chevauche son tour (l.105) pour apporter une autre justification concernant ses conditions de participation : la co-présence publique fait qu'il ne peut « pas parler » dit-il, ceci expliquant conjointement les grossissements « anormaux » mis en relief par A.

91. A : J'TE RAPPELLE TOUT A l'heure yeu:
92. (1.2)
93. B : pourquoi ?
94. A : (0.2) béh parc'que en visio comme ça c'est nul huhuhu
95. hhhh j'te vois pas là: c'est °bon° hh j'vois
96. qu'ton oeil
97. []
98. B : tu m'vois pas ?
99. A : (0.5) hu (0.5) hu (.) hu hu
100.(0.4)
101.B : bah oui:: bah appelle moi:
102.(.)
103.A : hhhhh j't'appelle tout à l'heure/
104. []
105.B (inc) parce que j'peux pas
106. parler::
107.(0.7)
108.A : hein ?



Juste avant la délivrance du segment « j'vois qu't'on oeil ».

IV. Conclusion

Les analyses précédentes permettent d'évaluer les façons dont des environnements sonores, se modifiant constamment au sein de cadres sociaux, sont susceptibles de devenir des composantes à part entière des interactions visiophoniques sur mobile.

Que les fonds sonores – mutualisés dans l'entre-deux visiophonique – soient continûment importants, ponctuellement marqués par des bruits d'intensité variable, ou relativement calmes, les interactants distants rendent évident un travail d'ajustement constant pour faire face aux interférences sonores ou pour ne pas en devenir les producteurs /écoute publique.

Ce travail pratique est réalisé à partir de différentes modalités interactionnelles (parole, regard, orientation de la partie haute du corps, distance variable bouche/microphone, selon qu'on parle/écoute, etc.) à la recherche du maintien de l'intelligibilité mutuelle, et en particulier dans sa dimension verbale : si les circonstances sonores le réclament, ils mettent en oeuvre différentes « parades » visant à maximiser la production/réception de la parole. Lorsqu'ils procèdent ainsi, ils abandonnent ou modifient le talking heads, format standard des échanges visiophoniques sur mobile.

Le dé-placement du mobile dans presque tous les lieux du quotidien, et peut-être plus particulièrement dans les lieux publics, expose les visiophonistes à des mondes sonores extrêmement variés, changeants, et +/- appropriés pour ce format d'interaction. Ce principe d'exposition varie également selon que les participants utilisent ou non des oreillettes, peuvent ou non (bus, tramway) exploiter l'espace pour s'éloigner de sources, avérées ou simplement potentielles, de nuisance sonore ; lorsqu'ils sont en mode haut-parleur, la distance le microphone tend à capter sur un pied d'égalité l'acoustique des lieux de communication et la voix des participants en raison de la distance /à la bouche : ils réalisent alors des arbitrages entre écoute mutuelle, production du portrait et aptitude à entretenir un regard commun.

Ces éléments montrent comment se juxtaposent et se font à la fois différents mondes sonores. Ceux-ci mêlent attentes d'intelligibilité mutuelle dans le cadre d'une interaction focalisée et maintien de « l'ordinaire acoustique » des lieux de communications /co-présence publique. La gestion de ces deux « fronts d'interaction » et d'attentes normatives propres ne se réduisent pas à la dimension sonore, qu'il s'agisse de garantir la production/réception de la parole ou de se fondre dans le paysage sonore : la production de l'image de soi, sa réception, et différentes ressources corporelles, visuelles, rentrent également en ligne de compte.

Bibliographie

Argyle M. (1975) *Bodily communication*. New York, NY: International Universities Press.

Bruce V. (1996) "The role of the face in communication implications for videophone design", *Interacting with computers*, 8(2).

De Fornel M. (1994) « Le cadre interactionnel de l'échange visiophonique », *Réseaux* n°64.

Goffman E. (1963) *Behaviour in public places. Notes on the social organization of gatherings*, New York, Free Press.

Goffman E. (1973) *La mise en scène de la vie quotidienne - Les relations en public*, Paris, Les Editions de Minuit.

Goodwin C. (1981) *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.

Heath C., Luff P. (1991) "Disembodied Conduct: Communication Through Video in a Multi-media Office Environment". In *Proceedings of CHI '91*, ACM.

Isaacs E., Tang J.C. (1993) "What Video Can and Can't Do for Collaboration: a Case Study". In *Proceedings of Multimedia '93*, ACM.

Kendon A. (1990) *Conducting Interaction Patterns of behaviour in focused encounters*, Cambridge University Press.

Kies J. (1997) *Empirical Methods for Evaluating Video-Mediated Collaborative Work*. Phd, faculty of the Virginia Polytechnic Institute.

Messonnier J. (2001), *Provinciliens : les voyageurs du quotidien, entre capitale et province*, L'Harmattan.

Mondada L. (2007) « Imbrications de la technologie et de l'ordre interactionnel. L'organisation de vérifications et d'identifications de problèmes pendant la visioconférence », *Réseaux* n°144.

Morel J. (2008) « Les interférences sonores vues de la conversation téléphonique », à paraître.

Relieu M. (2007) « La téléprésence ou l'autre visiophonie », *Réseaux* n°144.

Richter A. (2007) Mobile videotelephony – Test of 3G telephones, Hjälpmedelsinstitutet (SHI).

Sacks H., Schegloff E., Jefferson G. (1974) “A simplest systematics for the organization of turn taking for conversation”, *Language*, n°50.

Schegloff E. (1988) “Goffman and the analysis of conversation”, in P. Drew & T. Wootton (eds) *Erving Goffman: Exploring The Interaction Order*, CUP.

Schegloff E. (1972) “Sequencing in conversational openings”, in Gumperez J., Hymes D. (ed.), *Directions in Sociolinguistics: The Ethnomethodology of Communication*, NewYork: Holt Rinehart and Winston.

Tang J.C., Isaacs E. (1993) “Why do users like video?”, *CSCW*, 1.

Whittaker S., Scلابassi R. (1993), “Turning Away from Talking Heads: The Use of Video-as-Data in Neurosurgery”, *ACM, INTERCHI '93*.

Convention de transcription

a) Chevauchements

1 A : c'est ce que je disais
2 []
3 B : moi aussi

b) Rythmes d'élocution

1 A : >parler vite<
2 B : <parler lentement>

c) Silences

(.) : silence de moins d'un dixième de seconde
(1.2) : silence d'une seconde et deux dixièmes

d) Accentuation/atténuation

1 A : parler fort
2 B : ENCORE PLUS FORT
3 A : °parler bas°
4 B : °°encore plus bas°°

e) Enoncés liés

1 A : continuité des=
2 B : =tours de paroles

f) Interruption

all/ : item dont la fin est interrompue
all/(ô) : fin probable ou suggérée

g) Intonations

ah ! : ton vif
ah. : intonation descendante
ah, : intonation descendante puis montante
ah ? : contour interrogatif

h) Allongements vocaliques

1 A : allô:.....

i) Respiration

hhhhhhh : inspiration
h.h.h.h.h. : expiration

j) Commodités de transcription

[inc] : item, segment voire tour de parole incompréhensible
(pas sûr) : transcription est incertaine
((sifflement de train)) : indication de certains événements
xxxxxx : façon de représenter un bruit important

Convention adaptée de M. Atkinson et J. Heritage, *Structures of social action...*, p. ix-xvi et de B. Conein et alii., *Les formes de la conversation...*, vol. 2, p. 11-13)