

SUR LES QUALITES PRINCIPALES DES VOYELES SOUTENUES ET DES PAROLES CONTINUES

YOSHIYUKI OCHIAI, TERUO FUKUMURA et FUMIO ISHINO

Section Electrotechnique

(Reçu le 28 Mai, 1962)

Quand on jette un regard en arrière sur l'histoire de la téléphonie, particulièrement sur le développement du principe et de la méthode de l'évaluation des systèmes téléphoniques, on s'aperçoit qu'il y a deux principes donnant la base à l'évaluation des systèmes: le principe basé sur l'intensité sonore subjective et le principe reposant sur la qualité de netteté. Or, l'intensité sonore subjective et la qualité de netteté sont les attributs des sons de la parole, et c'est vraiment aux attributs des paroles que nous pouvons apprécier les systèmes de transmission.

Si la netteté, c'est-à-dire la qualité des phonèmes, mérite en effet d'être mise au service de l'évaluation des systèmes de téléphone, on peut réfléchir un peu, se basant sur le point de vue plus général, à une autre qualité, c'est-à-dire à la qualité des voix, parce que cette qualité est plus adéquate que la qualité de netteté pour l'évaluation haute et précise des systèmes de transmission des paroles en général (comprenant le système de téléphone). C'est ainsi que nous posons ici ces questions: quelles sont les caractéristiques des trois qualités des vocales ou des paroles? Quelle différence existe entre ces trois qualités, l'intensité sonore subjective, la qualité phonémique ou la qualité de netteté des phonèmes, la qualité vocale ou la qualité de clarté des voix, du point de vue des voyelles soutenues ou des paroles continues. C'est la considération la plus sérieuse quand on préconise l'idée de la qualité de la transmission au lieu du rendement de la transmission.

D'abord nous obtenons les caractéristiques des trois qualités des voyelles

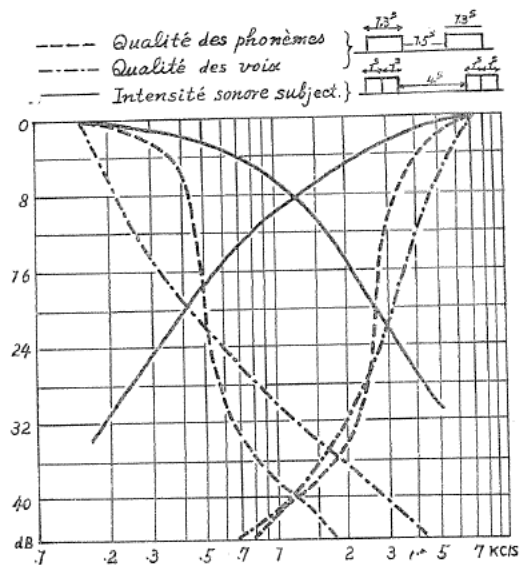


FIG. 1. Caractéristiques des trois qualités des voyelles soutenues, qualité phonémique, qualité vocale, qualité d'intensité sonore subjective, en fonction des fréquences de coupure. Trois qualités sont représentées par l'affaiblissement équivalent à la suite de la méthode indirecte.

soutenues quand les voyelles sont soumises en distorsion de bandes-éliminatoires de deux directions successives et systématiques. Comme unité de la qualité, l'affaiblissement équivalent est choisi communément pour les trois qualités afin que nous puissions examiner avec exactitude et précision les formes des caractéristiques différentes. Dans la figure 1, nous montrons ces trois paires de caractéristiques des voyelles: les courbes en traits pleins sont les caractéristiques de l'intensité sonore; les courbes en chaînes, les caractéristiques de la qualité vocale; les courbes en traits coupés, les caractéristiques de la qualité phonémique. Ces caractéristiques-ci sont obtenues au sujet des cinq voyelles japonaises émises par quatre speakers masculins sur une hauteur de 150 c/s, au niveau phonatoire de *mezzo-forte*, reçues par quatre auditeurs. Par cette représentation des caractéristiques des qualités, on peut atteindre à la conclusion suivante: la nature de l'intensité sonore subjective des voyelles est toute différente de leur qualité de netteté des phonèmes d'une part, et d'autre part, elle est aussi fort différente de leur qualité de clarté des voix; pour faire un pas en avant, le mode de distribution de la qualité de l'intensité sonore subjective en dimension de fréquence est très loin de celui de la qualité phonémique ainsi que de la qualité vocale. En d'autres termes, l'intensité sonore des voyelles ne peut se substituer tant à la qualité des phonèmes qu'à la qualité des voix, bien que le point de balance de l'intensité sonore—c'est le point d'intersection d'une paire de caractéristiques de l'intensité sonore—puisse presque identique à celui de la qualité phonémique.

Ensuite nous étudions les trois qualités des paroles continues. Par adoption des voyelles statiques ou des voyelles soutenues, bien adaptées à une certaine hauteur et à un certain niveau, nous pouvons, avec aisance et précision, observer et déterminer les qualités différentes qui sont inhérentes aux voyelles elles-mêmes, et de plus nous pouvons rechercher, non sans difficultés, la relation entre le modèle spectral du timbre des voyelles et l'allure des caractéristiques de leurs qualités. C'est vraiment là, dans la correspondance entre le modèle comme stimulus et la qualité comme réponse, que nous pouvons trouver un grand sens à l'emploi des voyelles statiques. Pourtant les voyelles actuelles parues dans la conversation courante ne sont pas ensemble avec telles voyelles statiques que nous employons dans nos expériences, mais elles varient très rapidement et brusquement tant en niveau qu'en hauteur. Relativement aux voyelles qui changent, non seulement leurs phonèmes, mais leurs hauteurs et niveaux d'une manière dynamique et continue dans le déroulement des conversations vivantes, quels caractère et allure sont observés du point de vue des trois qualités que nous venons de voir au sujet des voyelles soutenues? C'est ainsi qu'est amenée la thèse: quelles sont l'intensité sonore subjective, la qualité de la naturalité des voix exprimée par les paroles, et la qualité de l'intelligibilité des idées entendues dans les paroles?

Dans ce but, nous utilisons aussi la distorsion de bandes-éliminatoires de deux directions afin de saisir la distribution des trois qualités des paroles dans la dimension de fréquence. Comme qualité vocale des paroles continues, nous pouvons utiliser la même idée sur l'identification des voix que nous avons utilisée pour les voyelles soutenues. Quant à la qualité phonémique, nous ne pouvons utiliser, pour les paroles variables, la même conception de la netteté des sons que pour les voyelles soutenues. Nous devons utiliser ici la notion de l'intelligibilité des idées comprises dans l'expression des paroles considérées.

Concernant la méthode de mesurage des qualités, intelligibilité et naturalité,

de la parole continue, il y en a deux : l'une est la méthode indirecte, l'autre, la méthode directe. La méthode indirecte consiste dans le processus que nous avons déjà employé à propos des voyelles soutenues. C'est la méthode où nous effectuons deux expériences sur la distorsion de bandes-éliminatoires et sur la distorsion d'affaiblissement, et au moyen de la comparaison des deux résultats, nous pouvons compter l'affaiblissement équivalent correspondant à la perte de la qualité en pourcent. La méthode directe repose sur le processus où l'on examine la même condition de la qualité correspondante en comparant directement la parole découpée par la distorsion de bandes-coupures et la parole affaiblie par la distorsion d'affaiblissement réglable. Quant à l'intensité sonore subjective, la mesure consiste toujours dans la méthode de balance entre la parole découpée et la parole affaiblie, les deux présentées aux oreilles d'auditeurs en en faisant une paire de combinaisons successives, pendant 2 ou 3 secondes, avec une pause de la même durée.

Nous présentons, dans la figure 2, l'allure des caractéristiques des qualités des paroles obtenues par la méthode indirecte, et dans la figure 3, l'allure des caractéristiques des qualités obtenues par la méthode directe. Un regard sur la figure 2 suffit pour comprendre que les deux qualités fondamentales, c'est-à-dire l'intelligibilité des idées et la naturalité des voix, viennent se rapprocher l'une de l'autre, toutes les deux étant très loin de la qualité de l'intensité sonore subjective. Le point de balance de l'intensité sonore des paroles se situe dans la région de 1.0 kc/s. Les points de balance de

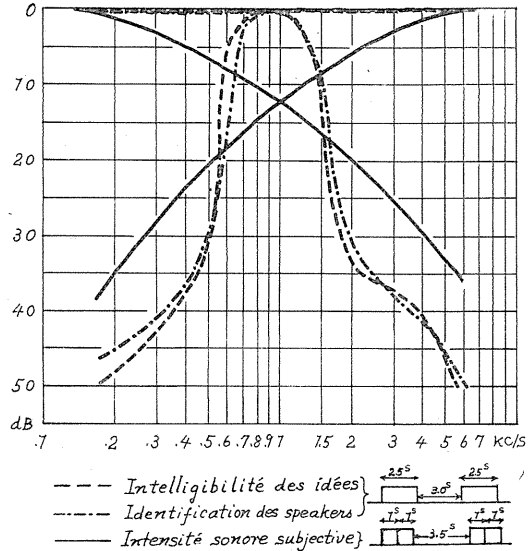


FIG. 2. Caractéristiques des trois qualités des paroles continues: l'intelligibilité des paroles, la clarté des voix, l'intensité sonore subjective des paroles, en fonction des fréquences de coupure. Trois qualités sont représentées par l'affaiblissement équivalent mesuré au moyen de la méthode indirecte.

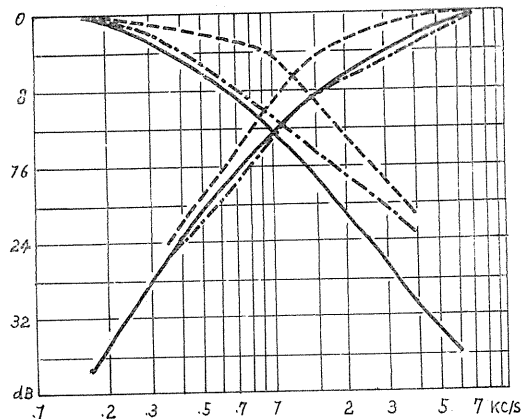


FIG. 3. Caractéristiques des trois qualités des paroles continues: intelligibilité des paroles, naturalité des voix, intensité sonore subjective des paroles, en fonction des fréquences de coupure. Trois qualités sont mesurées par l'affaiblissement équivalent par la comparaison directe.

l'intelligibilité des idées et de la naturalité des voix, semblent se trouver aussi dans la même région, ne s'écartant pas beaucoup de celle de l'intensité sonore. Il nous semble très significatif que la qualité sur l'identification des idées et la qualité sur l'identification des voix sont beaucoup plus similaires dans les paroles continues que ne sont la qualité phonémique et la qualité vocale dans les voyelles soutenues, quoique les qualités soient obtenues suivant par la même méthode indirecte par la même équipe d'auditeurs sur les deux objets, paroles continues et voyelles soutenues.

Remettant en comparaison la figure 2 avec la figure 1, nous pouvons voir, concernant la forme des caractéristiques, que l'intensité sonore des paroles continues ne diffère que peu de celle des voyelles soutenues, tandis que les qualités fondamentales des paroles sont toute différentes en forme de celles des voyelles. Ce fait signifie que l'intensité sonore des paroles est déterminée exclusivement par celle des voyelles contenues dans les paroles, mais la naturalité des voix exprimée dans les paroles et l'intelligibilité des idées entendues dans les paroles ne sont pas déterminées exclusivement par la netteté des voyelles et par la clarté des voix comprises dans les paroles. A parler d'une manière inverse, il y a, dans les paroles continues, divers facteurs qui gouvernent l'intelligibilité et la naturalité des paroles en dehors de la netteté des voyelles et de la clarté des voyelles au sens statique.

Il va sans dire que l'intelligibilité des paroles continues dépend, non seulement des voyelles, mais aussi des consonnes; le rôle de ces dernières devient d'autant plus grand que la fréquence d'occurrence des consonnes est plus grande dans la langue considérée. Relativement à la clarté des voix dans la perception des paroles continues, nous pouvons considérer qu'en dehors des structures du timbre des voyelles, il y en a plusieurs facteurs qui ont de l'influence sur l'identification et la discrimination des sujets parlants. Par exemple, les parties d'attaque et d'extinction des sons vocaux, le mode du changement des voyelles au sujet tant du niveau que de la hauteur, le mode de combinaisons des consonnes avec les voyelles et la prononciation des consonnes elles-mêmes. Ces divers facteurs semblent influencer, d'une manière séparée ou combinée, la formation de la qualité de naturalité des voix dans les paroles continues.

Un autre fait, qui est communément le même dans ces deux représentations des figures 1 et 2, est que la forme des caractéristiques des qualités est très symétrique à l'égard de leur point de balance. En d'autres termes, une paire de caractéristiques des qualités dans la distorsion de bas en haut ainsi que dans la distorsion de haut en bas devient plus symétrique dans l'expression des qualités par l'affaiblissement équivalent que dans l'expression des qualités par le pourcentage.

Ce qui est montré par la représentation de la figure 3, c'est que les trois paires de caractéristiques des trois qualités des paroles continues deviennent assez similaires, se rapprochant au moins l'une de l'autre, quand les paroles découpées sont mises directement en comparaison avec les paroles affaiblies du point de vue de la qualité individuelle. Cette expérience est faite après avoir fini les diverses expériences sur les différentes qualités des voyelles soutenues et même l'expérience sur les qualités des paroles continues montrées dans la figure 2. Par conséquent, les auditeurs sont déjà suffisamment habitués aux notions et aux mesurages des qualités différentes des voyelles et des paroles. Malgré cela,

voilà le résultat. C'est très significatif et très important.

Par la comparaison de la figure 3 avec la figure 2, nous pouvons voir que la mise en balance des deux paroles sous les deux différentes distorsions à propos d'une qualité quelconque est très difficile quand la qualité considérée est la qualité supérieure et complexe, et davantage qu'il est presque impossible de différencier strictement les trois qualités par lesquelles on met en balance les deux paroles différemment déformées par les deux différentes distorsions.