

FONDAMENTALES DES QUALITES PHONEMIQUE ET VOCALIQUE  
DES PAROLES PAR RAPPORT AU TIMBRE, OBTENUES  
EN EMPLOYANT DES VOYELLES JAPONAISES  
VOCALISEES PAR DES SUJETS JAPONAIS

YOSHIYUKI OCHIAI

*Section Electrotechnique*

(Reçu le 31 octobre, 1958)

Le sens d'un mot peut être communiqué par le ton et le timbre de la voix de celui qui parle ainsi que par la façon dont le mot est employé. Le mot parlé (non le mot écrit) a deux qualités, phonémique et vocalique, qui sont les fondamentales de la parole. Le choix des mots pour exprimer une idée est différent de personne à personne. Le même texte qui exprime une définition quelconque se révèle entièrement différent suivant la disposition et l'état de l'émotion de celui qui parle. Il n'y a que certains degrés de variation dans la présentation logique des idées, tandis que les nuances exprimant les émotions sont illimitées.

Nous devons d'abord différencier les nuances dans la présentation logique des idées des nuances émotionnelles, même si elles sont, de fait, toujours en intime connexion.

Quand nous prenons un des attributs du son pour l'étude des voyelles soutenues, nous avons naturellement une différence entre la qualité phonémique et la qualité vocalique au point de vue du timbre. La qualité phonémique des voyelles correspond à une certaine qualité permettant l'identification et la discrimination des phonèmes au sens linguistique, tandis que la qualité vocalique des voyelles correspond à une autre qualité qui contribue à l'identification et la discrimination des voix de ceux qui parlent. La qualité phonémique est en connexion avec le texte, alors que la qualité vocalique est intimement liée à la différence des personnes ou à la différence d'expressions d'une même personne.

Si nous acceptons ce raisonnement, il n'y a pas de doute sur la méthode à suivre pour trouver le "type phonémique" (phonemic pattern en anglais) des voyelles et le "type vocalique" (vocal pattern) des personnes, parce que dans la parole humaine il peut y avoir plusieurs choses concernant l'élément de la voix du speaker ainsi que l'élément phonémique du texte. Le "phonème" est une idée trouvée et introduite dans la science phonétique comme résultat de la cristallisation graduelle d'une notion, abstraite d'une certaine façon de notre longue expérience de la vie du langage. C'est dans la même ligne que nous devons et pouvons trouver, dans le terme "voix," une idée très importante pour nous, abstraite aussi d'une autre façon de nos expériences accumulées de la vie et de l'habitude du langage parlé.

D'étudier la nature de la qualité phonémique et de déduire le type phonémique en structure du timbre des voyelles, c'est très important pour des connaissances de phonétique et aussi très utile à la technologie de la communication de la parole. Pour l'art et la science de la communication téléphonique, par exemple, c'est tout à fait essentiel de connaître les éléments des paroles suffisamment et nécessairement

utiles pour établir la netteté au point de vue d'un meilleur rendement de transmission, particulièrement quand on projette des circuits ou systèmes de communication. Eclairer le problème de l'élément vocalique des paroles qui n'est pas encore étudié, et clarifier le mystère des nuances des voix et le secret de l'individualité vocalique, c'est d'importance dans la science des voix et aussi dans la technique de la musique vocale. L'importance croissante de l'étude de la qualité vocalique et du type vocalique est nécessitée par la technologie de la communication de la parole.

Considérons, par exemple, la conversation au téléphone. Il y a une différence essentielle entre la conversation téléphonique et la conversation face-à-face (en présence de l'interlocuteur). L'échange des idées par le téléphone doit se baser uniquement sur la faculté auditive, parce que toute aide visuelle possible est entièrement refusée dans ce cas. Les usagers du téléphone ne peuvent se voir l'un l'autre dans la conversation téléphonique. Il en résulte que le jugement des interlocuteurs dépend de l'impression des voix transmises sur la ligne téléphonique. C'est une des raisons pour lesquelles la technologie du téléphone demande l'étude de la qualité vocalique des sons. Une autre raison de cette étude se trouve dans le problème de l'évaluation du système de communication. L'évaluation du système basée sur la qualité phonémique, c'est-à-dire l'évaluation du système au point de vue de la netteté ne peut plus avancer dans cette voie. Nous établissons donc un autre chemin pour déterminer, avec plus de précision et une meilleure échelle de qualités, l'évaluation du système de communication. L'évaluation du système basée sur l'intensité sonore psychologique (*loudness* en anglais) a peut-être quelque valeur; mais nous pouvons espérer davantage d'une méthode basée sur la qualité vocalique de la parole.

Déterminer l'interdépendance du phonème et de la voix est le but ultime de notre étude; c'est peut-être là que nous trouvons le thème le plus intéressant. Pour étudier l'influence mutuelle du phonème et de la voix et pour déterminer la relation de leurs situations réciproques, nous avons besoin certainement d'une avance technique et d'un progrès dans l'expérimentation. Il suffit de voir ici que les études sur la confusion du phonème et de la voix sont les plus utiles pour détecter la relation subtile qui existe entre phonème et voix. De ceci nous reparlerons plus tard. Pour vérifier simplement ce point que la qualité vocalique au point de vue du timbre est très différente de la qualité phonémique, il convient de montrer que dans la nature essentielle de perception du timbre, par exemple, des vocales, la caractéristique de l'identification de la voix est entièrement différente de celle de l'identification du phonème. Pour différencier clairement la caractéristique de l'identification de la voix et la caractéristique de l'identification du phonème, l'adoption de la distorsion de bande éliminatoire, par exemple, est la plus appropriée, parce que le mode de distribution de la qualité vocalique de la parole en dimension de fréquence y est très différente du mode de distribution de la qualité phonémique. Nous pouvons illustrer ce point en employant des vocales soutenues par les speakers japonais, comme le signal du timbre.

Après avoir fait des expériences tant de fois répétées pendant plusieurs années (de 1953 à 1957) dans notre Laboratoire d'Audiologie, nous sommes amenés à la conclusion suivante: La forme de la caractéristique de l'identification du phonème, c'est-à-dire le mode de répartition de la qualité de netteté des voyelles est sûrement différente de la caractéristique de l'identification des voix, c'est-à-dire du

mode de répartition de la qualité de clarté des voix. Pour conclure cela, il suffit de noter deux points de différence: La différence de l'inclinaison de ces deux caractéristiques et la différence du point d'intersection que nous appelons le "point de qualité en balance." A ce point, la somme des qualités réparties dans le domaine de basse fréquence est en balance avec la somme des qualités distribuées dans le domaine de haute fréquence.

TABLEAU. Liste des Expériences répétées dans notre Laboratoire sur la Netteté des Voyelles et la Clarté des Vocales

Numéros des expériences	Années des expériences	Voyelles utilisées	Voix utilisées			Hauteurs musicales du ton	Nombre d'auditeurs
			(Homme)	(Femme)	(Enfant)		
(1)	1953	I, E, A, O, U	2	2	1	280 ~	4
(2)	1954	I, E, A, O, U	4	0	0	140 ~	4
(3)	1954	I, E, A, O, U	4	1	0	240 ~	4
(4)	1954	I, E, A, O, U	4	1	0	chuchotement	4
(5)	1955	I, E, A, O, U	3	2	0	240 ~	4
(6)	1956	I, E, A, O, U	2	2	0	220 ~	4
(7)	1957	I, E, A, O, U	4	0	0	150 ~	4

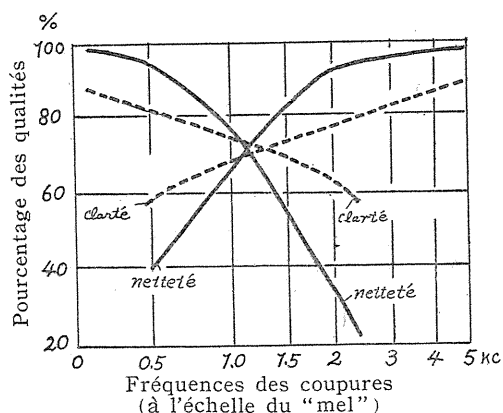


FIG. 1. Les caractéristiques de netteté des phonèmes et de clarté des voix dans la distorsion des bandes éliminatoires. Les 5 voyelles émises par 4 voix sur le ton de 140 p/s sont présentées aux 4 auditeurs. Chacune des caractéristiques est obtenue par 1200 observations pour chaque condition. Les allures de clarté en fonction de fréquence sont presque droites pour l'échelle du mel.

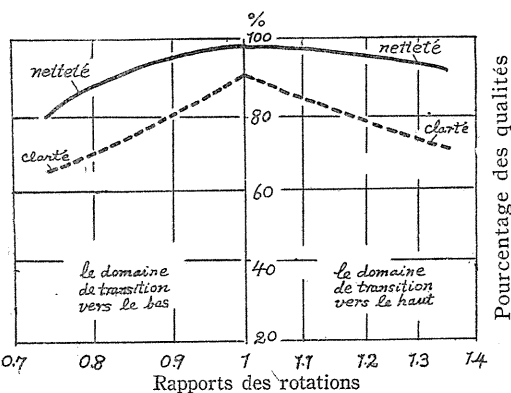


FIG. 2. Les caractéristiques de netteté des phonèmes et de clarté des voix dans la distorsion en synchronisation entre enregistrement et reproduction. Les 5 voyelles vocalisées par 4 voix sur le ton de 140 p/s sont présentées aux 4 auditeurs. Chacune des caractéristiques est obtenue par 1200 observations pour chaque condition.

Quant à l'inclinaison, la caractéristique de netteté des phonèmes en coupure haute ainsi que celle de netteté des phonèmes en coupure basse montrent, d'abord la pente douce et puis la pente très rapide au fur et à mesure que la distorsion devient plus grande. Les caractéristiques de la clarté des voix basée sur l'identification des voix, montrent, au contraire, une pente modérée et à peu près constante en dépit de la grandeur de la distorsion. Il est très intéressant de noter que le point de balance de la netteté des phonèmes est clairement un peu plus bas que celui de la clarté des voix et cette tendance devient de plus en plus évidente quand on vocalise des voyelles sur des tons plus élevés.

Il est d'une importance capitale de noter que le mode de distribution des qualités est différent suivant la nature de la qualité. La distribution de la qualité de netteté des phonèmes n'est pas uniforme et constante en dimension de fréquence, mais elle montre l'excentricité de la distribution, c'est-à-dire la répartition des qualités, très dense et concentrée, en une certaine région de fréquence déterminée. Le mode de distribution de la qualité de clarté des voix est, au contraire, relativement uniforme et étendue en dimension de fréquence. Puisque le point de balance de qualité correspond approximativement au point de vallée\* vocalique en structure du timbre des voyelles, la somme des qualités distribuées dans le domaine de haute fréquence, surpassant le point de vallée, est égale à la somme des qualités réparties dans le domaine de basse fréquence, au dessous du point de vallée. Il en résulte donc que pour l'évaluation du système de communication nous pouvons plus aisément nous baser sur la qualité qui montre la distribution la plus simple et la plus uniforme en dimension de fréquence.

Il n'est pas inutile d'ajouter ici: Le fait que le mode de distribution de la qualité est évidemment différente suivant la différence de la qualité, n'est pas sans rapport avec la nature de la faculté auditive. Pour illustrer cela, nous montrons un autre exemple de distorsion. En dehors de la distorsions des bandes éliminatoires par laquelle nous avons pu examiner le mode de distribution de la qualité dans le domaine de fréquence (c'est-à-dire la structure subtile de qualité en dimension de fréquence), nous devons utiliser une autre distorsion très importante: la distorsion en synchronisation entre enregistrement et reproduction. A l'aide de cette distorsion, nous pouvons examiner un autre type de la réponse à l'excitation du timbre lorsque tous le domaine des voyelles est soumis, dans son ensemble, à transition vers le haut ou vers le bas, sans avoir apporté dans ce cas aucune coupure de bande appréciable. En utilisant cette distorsion en synchronisation, nous pouvons positivement constater que la caractéristique de netteté des phonèmes peut supporter fortement la distorsion vers le haut ainsi que la distorsion vers le bas, tandis que la clarté des voix n'est pas forte contre la même distorsion et elle montre une pente de médiocre décroissance même pour la distorsion la plus infime. En d'autres termes, nous pouvons mentionner: Au point de vue de la netteté des phonèmes, les voyelles en leur ensemble sont très endurantes contre la distorsion de transition, mais ces mêmes voyelles sont assez faibles contre la même distorsion au point de vue de la clarté des voix.

En résumé: En utilisant des voyelles japonaises émises par des voix de sujets japonais et en adoptant la distorsion de bande éliminatoire ainsi que la distorsion en synchronisation entre enregistrement et reproduction, nous pouvons confirmer le fait que comme qualité du timbre des voyelles la netteté des phonèmes est entièrement différente de la clarté des voix, c'est-à-dire le fait que le mécanisme pour l'identification du phonème des voyelles est tout à fait différent du mécanisme pour l'identification de la voix des speakers. Ce fait se produirait sans aucun doute si l'on utilisait n'importe quelles voyelles et n'importe quelles voix en dehors de la langue japonaise et du sujet japonais.

#### Bibliographie

---

\* La vallée majeure trouvée aux environs de 1.2-1.6 kp/s qui est la marque frappante et commune des voyelles orales japonaises.

- 1) Y. Ochiai: Mémoire sur les Sons des Voix Humaines, MFE (Memoirs of the Faculty of Engineering). Nagoya Univ. Vol. 4, No. 1 (1952).
- 2) Y. Ochiai: Transmission of Quality, MFE. Nagoya Univ. Vol. 6, No. 2 (1954).
- 3) Y. Ochiai and T. Yamashita: On Timbre Quality. MFE. Nagoya Univ. Vol. 7, No. 1 (1955).
- 4) Y. Ochiai and T. Fukumura: Preliminaries to Analysis of Quality in Speech-Communication, MFE. Nagoya Univ. Vol. 9, No. 2 (1957).
- 5) Y. Ochiai and T. Fukumura: Timbre Study of Vocalic Voices, MFE. Nagoya Univ. Vol. 5, No. 2 (1953).
- 6) Y. Ochiai and T. Fukumura: Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part I—Preliminary Studies on Naturalness and Articulation Qualities Actually and Directly Measured with Respect to Band-Eliminating Distortion, MFE, Nagoya Univ. Vol. 8, No. 1 (1956).
- 7) Y. Ochiai, T. Fukumura and T. Hattori: Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part II—Preliminary Studies on Timbre Confusion of Phoneme and Voice, MFE, Nagoya Univ. Vol. 8, No. 2 (1956).
- 8) Y. Ochiai and T. Fukumura: Timbre Study of Vocalic Voices Viewed from Subjective Phonal Aspect. Part III—Generalized Treatment of Timbre Confusion, MFE. Nagoya Univ, Vol. 8, No. 2 (1956).
- 9) Y. Ochiai and T. Fukumura: Study on Fundamental Qualities of Vocalic Timbre by Rotational Synchronous Distortion, MFE. Nagoya Univ. Vol. 8, No. 1 (1956).
- 10) Y. Ochiai und T. Fukumura: Beiträge zur Erkenntnis der Klangfarbestruktur bei vokalischen Klangbildern, MFE. Nagoya Univ. Vol. 8, No. 1 (1956).
- 11) H. Schulz: Die Fernsprechtbetriebsmasse. Ihre Entstehung, Bedeutung, kritische Würdigung und Verallgemeinerung. EFD Okt. (1934).
- 12) K. Braun: Die Bedeutung und Bestimmung der Übertragungsgüte in Fernsprechverkehr. TFT, 29, 5, Mai (1940).
- 13) H. Bornemann: Die Bewertung der Übertragungsgüte von Fernsprechsystem. EFD, März (1939).
- 14) F. W. McKown and J. W. Emling: A System of Effective Transmission Data for Rating Telephone Circuit. BSTJ, July (1933).
- 15) J. Collard: A Theoretical Study of the Articulation and Intelligibility of a Telephone Circuit. Elect. Comm. Jan. (1929).
- 16) J. Collard: A New Criterion of Circuit Performance. Elect. Comm. April (1933).
- 17) J. Collard: The Practical Application of the New Unit of Circuit Performance. Elect. Comm. April (1934).
- 18) K. Braun: Die Bezugsdämpfung und ihre Berechnung aus der Restdämpfungskurve (Frequenzkurve) eines Übertragungssystem. TFT. Aug. (1939).
- 19) K. Braun: Theoretische und experimentelle Untersuchung der Bezugsdämpfung und der Lautstärke. TFT, 29, 2, Feb. (1940).
- 20) K. Braun: Eine neuere Bezugsdämpfungsmesser mit objektiver Erregung und Anzeige. TFT, 29, 8, Aug. (1940).
- 21) P. W. Blye, O. H. Coolidge and H. R. Huntley: A Revised Telephone Transmission Rating Plan. BSTJ, Vol. XXXIV, No. 3, May (1955).
- 22) F. Lüschen und K. Küpfmüller: Die Entwicklung der Übertragungstechnik für den Nachrichtendienst über Leitungen. Jahrbuch der elektrischen Fernmeldewesens von F. Gladenbeck, Berlin (1938).
- 23) E. Schäfer: Über die Hörbarkeit von Frequenzbandänderungen bei der Übertragung von Sprache. ENT, Band 15, Heft 8, August (1938).
- 24) P. Schiaffino: Méthodes objectives de mesure de l'Equivalent de Référence et de l'Affaiblissement équivalent de Netteté en Téléphonométrie. Extrait des Annales des Télécommunications. Tome 12, No. 10, Octobre (1957).
- 25) Y. Ochiai: Onsei no Shizendo no Hyogen-Ho to Shakudo ni kansuru Kenkyu (in Japanese)-Study on Representation of Naturalness Quality of Speech and its Applicability to Quality Scale for Rating and Evaluation of Speech-Communication System. Special Report of Study under Contraction of Nippon-Denden-Kosha, March (1958).