

L'audiométrie vocale

Dr C Meyer-Bisch

Les signes les plus fréquents d'une atteinte fonctionnelle de l'audition sont une baisse des capacités de compréhension de la parole, notamment remarquée par les proches. En fait, la première évaluation fonctionnelle de l'audition est une observation de transmission vocale. Il n'est donc pas étonnant que l'audiométrie vocale soit un examen essentiel dans l'exploration de l'audition. Au départ plutôt empirique, cette audiométrie s'est perfectionnée au fil des ans, grâce au progrès technique et à une exigence méthodologique que l'on retrouve dans la norme ISO-EN-NF 8253-3¹ qui a récemment subi de nombreuses modifications, notamment pour harmoniser les pratiques.

1. Intérêts et principes de l'audiométrie vocale

L'audiométrie tonale, sur laquelle nous reviendrons, s'intéresse essentiellement à la sensibilité de l'oreille à certaines fréquences. Elle ne peut pas, à elle seule, apprécier quantitativement la gêne sociale d'un individu. Pour cela, cet examen doit être complété par l'audiométrie vocale, qui est une mesure globale de la fonction auditive, permettant d'objectiver et de quantifier les capacités de reconnaissance de la parole.

1.1. L'audiométrie vocale fait partie de l'examen audiométrique

L'audiométrie vocale occupe donc une place essentielle dans l'appréciation des facultés de communication orale d'un individu, aussi bien dans un but diagnostique pour objectiver ses difficultés de communication, pour proposer des modalités de prise en charge, que pour évaluer l'efficacité d'une réhabilitation.

La reconnaissance d'un message vocal n'implique pas uniquement l'appareil sensoriel auditif, mais aussi les capacités centrales de traitement du signal auditif, les capacités de suppléance mentale, la maîtrise parlée de la langue de communication et certaines connaissances culturelles.

Chez l'enfant, l'audiométrie vocale permet très tôt à la famille de prendre conscience du déficit auditif et de la nécessité d'y remédier rapidement. Les examens vocaux chez le jeune enfant requièrent des compétences spécifiques qui sortent du champ de la normalisation et ne sont pas traités ici.

1.2. Principes et définitions

Les résultats de l'audiométrie vocale dépendent des messages vocaux et de la méthode d'essai utilisés. La norme sur laquelle nous nous appuyons ici définit les conditions que doivent remplir les messages vocaux afin d'assurer un minimum de précision et de comparabilité entre différents tests. Elle spécifie également les modes opératoires à utiliser lors des essais de reconnaissance vocale.

Il existe en effet plusieurs stades de perception d'un message vocal, correspondant à des définitions normalisées bien précises :

- La perception proprement dite (*speech detection*), correspond à la sensation de stimulation auditive mais sans compréhension; En pratique, cette notion n'est pas utilisée en audiométrie vocale.
- La réception, qui permet de répéter le son, sans nécessairement le comprendre;

- **La reconnaissance** du message vocal (*speech recognition*) qui diffère de la réception dans le fait que le message est compris. Le seuil d'intelligibilité vocale repose sur ce critère de reconnaissance.

L'audiométrie vocale va consister, à l'aide de messages vocaux appropriés, à mesurer un score de reconnaissance vocal pour différents niveaux vocaux, de façon à tracer une courbe vocale ou, à tout le moins, de déterminer un seuil de reconnaissance ou d'intelligibilité vocale. La reconnaissance vocale s'apprécie généralement par le fait que le patient répète correctement le signal vocal (un mot par exemple) qu'on lui a fait entendre. Ces différents termes reçoivent les définitions suivantes :

- Un **signal vocal** (*speech signal*) est un signal acoustique porteur d'information dans une langue donnée. Il peut être un signal de voix ou un signal acoustique simulant un signal de voix.
- Un **item** (*test item*) peut être un mot ou un logatome monosyllabique ou polysyllabique, une phrase ou fragment de parole enchaînée limité dans le temps, utilisé lors d'un essai d'audiométrie vocale selon des règles définies de comptabilisation (sur des items ou parties d'items).
- Un **message vocal** (*speech material*) est une série complète d'items utilisée pour des essais de reconnaissance vocale. Un message vocal est généralement subdivisé en plusieurs listes d'items.
- Une **spondée** (*spondee*) est un mot de deux syllabes également accentuées.
- Un **logatome** (*logatom*) est une syllabe n'ayant pas de signification pour l'auditeur, une « syllabe vide de sens ».
- Le **niveau vocal** (*speech level*) est le niveau de pression acoustique continu équivalent du message vocal, mesuré en utilisant la pondération fréquentielle C (voir Audio-Info N°...). Il faut noter que l'intégration ne doit comporter que les items et exclure les silences.
- Le **score de reconnaissance vocale** (*speech recognition score*), pour un niveau vocal spécifié, est le pourcentage d'items correctement reconnus (répétés) ou d'items susceptibles d'être comptabilisés si la méthode de comptabilisation n'est pas basée sur des items complets.
- Le **score maximal de reconnaissance vocale** (*maximum speech recognition score*) est la valeur maximale du score de reconnaissance vocale obtenue indépendamment du niveau vocal.
- Le **niveau liminaire d'intelligibilité vocale** ou **seuil d'intelligibilité vocale** (*speech recognition threshold level* ou **SRT**) est le niveau vocal le plus faible pour lequel le score de reconnaissance vocale est égal à 50 %. Le seuil d'intelligibilité vocale était autrefois dénommé « seuil de réception vocale ». Notons que cela ne simplifie pas la différence entre les notions de réception, de reconnaissance, d'intelligibilité...
- Le **niveau vocal optimal** (*optimum speech level*) est le niveau vocal auquel le score maximal de reconnaissance vocale est obtenu. Notons qu'il peut exister plusieurs niveaux vocaux optimaux.
- Le **niveau de confort optimal pour la parole** (*most comfortable level for speech*) est le niveau vocal auquel la sonie du signal vocal est considérée par le sujet comme correspondant au niveau de confort optimal.
- La **courbe d'intelligibilité vocale** (*speech recognition curve*) est la courbe de variation du score de reconnaissance vocale en fonction du niveau vocal.

Il convient de bien noter que le seuil d'intelligibilité vocale est un seuil statistique, comme le seuil tonal. Il possède donc son propre zéro de référence (pour un matériel vocal donné) ainsi que nous le verrons ci-après².

2. Le matériel vocal et ses caractéristiques

La norme précise que, sauf cas particuliers³, les messages vocaux utilisés en audiométrie vocale doivent provenir d'enregistrements qui garantissent la conformité aux spécifications qui sont maintenant très détaillées.

- **Contenu de l'enregistrement** : Chaque copie doit contenir le message vocal, un signal d'étalonnage (bruit aléatoire pondéré, bande de bruit d'un tiers d'octave centrée sur 1 kHz ou encore un son modulé en fréquence à 1 kHz de largeur au moins égale au tiers d'octave) d'au moins 60 secondes, et des signaux permettant de contrôler la réponse en fréquence de la chaîne de mesures. Ces signaux sont des bandes étroites centrées sur les fréquences de tiers d'octave recommandées par l'ISO 226 entre 125 Hz et 8 kHz.
- **Locuteur** : Il n'est pas précisé si le message vocal doit être produit par un locuteur ou une locutrice, mais il doit être prononcé « normalement », distinctement, sans accent prononcé et à un rythme normal. L'enregistrement ne doit pas être modifié avant d'en faire des copies.
- La **salle d'enregistrement** doit permettre d'obtenir un rapport signal/bruit d'au moins 40 dB et un temps de réverbération inférieur à 0,5 s pour toutes les fréquences entre 125 Hz et 8 kHz. Dans cette même gamme de fréquences, le matériel d'enregistrement doit avoir une réponse en fréquence plate à ± 2 dB.
- Le **niveau** de chaque item ne doit pas différer de ± 3 dB du niveau de l'ensemble des items de la liste et le niveau moyen d'une série d'items ne doit pas différer de ± 1 dB du niveau moyen de la liste.
- **Intervalles entre items** : Ils doivent être constants, au moins $4 \text{ s} \pm 10\%$.

Les listes d'items doivent être phonémiquement équilibrées⁴ et perceptuellement équilibrées, ce qui signifie que le résultat du test vocal doit être indépendant du choix de la liste d'items. Quelle que soit la liste d'items choisie, le seuil d'intelligibilité doit pouvoir être déterminé avec une incertitude ne dépassant pas 1 dB⁵.

Enfin, chaque copie d'enregistrement doit obligatoirement être accompagné d'une documentation comprenant la transcription écrite du message enregistré, les méthodes de comptabilisation, les courbes de références et la description détaillée⁶ des conditions de mesurage, la reproductibilité des résultats, toutes les précisions relatives aux signaux d'étalonnage... ainsi que les caractéristiques des éventuels bruits de fond associés.

En France, on utilise le plus couramment les listes dissyllabiques de Fournier et parfois de Lafon, ou celles cochléaires de Lafon. Selon les pays et les langues, le consensus est plus ou moins établi pour utiliser tel ou tel matériel vocal. Il y a une trentaine d'années, j'ai souvenir qu'un groupe de travail *ad hoc* de l'ISO a essayé de créer des listes internationales de logatomes. Malgré cet échec prévisible, il existe toujours des tentatives çà et là.

2.1. Les niveaux de référence dépendent du message vocal

Pour chaque message vocal, il doit exister une courbe d'intelligibilité de référence établie au moyen d'essais sur au moins 25 sujets otologiquement normaux de 18 à 25 ans⁷, dont la première langue est celle du message vocal. Quelques définitions supplémentaires :

- Le **seuil d'intelligibilité vocale de référence** (*reference speech recognition threshold level*) est la valeur médiane des seuils d'intelligibilité vocale d'un nombre « suffisamment grand de sujets otologiquement normaux des deux sexes, âgés de 18 ans à 25 ans inclus, pour lesquels les messages vocaux utilisés sont appropriés ».
- La **courbe d'intelligibilité vocale de référence** (*reference speech recognition curve*) est tracée à partir de seuils d'intelligibilité de référence pour des scores de 30 à 90%.

- Ceci permet de déterminer le **niveau d'audition de la parole** (*hearing level for speech*) qui est la différence entre le niveau vocal et le niveau liminaire (ou seuil) d'intelligibilité vocale de référence approprié.

Cette notion, très importante, permet de dissiper la confusion habituelle entre le niveau vocal (niveau à référence purement physique, SPL) et le niveau d'audition de la parole, dont la référence est physiologique et qui est à la vocale ce que le niveau HL (*hearing level*) est à la tonale. Il est à noter que la norme précise que le niveau du signal peut être exprimé en niveau vocal ou en niveau d'audition de la parole pour le message vocal utilisé. Les deux systèmes peuvent donc exister.

2.2. A propos du bruit de fond

L'audiométrie vocale peut être réalisée en présence d'un bruit de fond (*competing sound*), bien caractérisé et maîtrisé. Il faut alors définir :

- Le **niveau de pression acoustique du bruit de fond** (*competing sound pressure level*). C'est le niveau de pression acoustique continu équivalent d'un bruit de fond, mesuré en utilisant la pondération fréquentielle C selon la CEI 61672-1 ((L_{ceq})).
- Le **rapport parole/bruit** (P/B) (*speech-to-noise ratio (SNR)*) est la différence entre le niveau vocal et le niveau de bruit de fond.

Ce test est encore plus réaliste (et sensible) que la vocale dans le silence. Les Scandinaves ont mis en évidence, dans les années 90 qu'il était possible de réaliser des tests de dépistage vocaux dans le bruit par téléphone.

Pour l'audiométrie vocale dans le bruit, l'enregistrement de référence doit inclure également le ou les bruits de fond associés, sur les mêmes canaux ou sur des canaux séparés. Il est possible d'utiliser comme bruit de fond un bruit aléatoire pondéré en fréquence. Aucune valeur de référence valable n'existe lorsque des bruits de fond différents de ceux qui sont contenus dans l'enregistrement de référence sont utilisés. On peut aussi utiliser des enregistrements de voix humaines (bruit émis lors d'un cocktail) ou une parole enchaînée d'un locuteur unique. La variabilité des résultats peut être plus élevée avec ce type de bruits qu'avec un bruit aléatoire pondéré.

3. Méthodes d'audiométrie vocale

Généralement, une audiométrie tonale précède l'audiométrie vocale. Elle permet d'évaluer le niveau de départ, de mettre en évidence une différence interaurale. Les méthodes décrites dans la norme le sont dans un langage complexe qui, il faut le dire, décourage rapidement le lecteur. Nous allons tenter d'en extraire l'essentiel, de façon simple mais sans concession à la rigueur.

3.1. Avant de commencer

Le sujet doit être détendu et ne pas avoir été exposé au bruit dans les 5 minutes qui précèdent l'examen pour éviter la fatigue auditive (*TTS*). L'otoscopie vérifie la vacuité des conduits. Il faut vérifier l'aptitude du sujet à comprendre et répéter les messages vocaux.

Le sujet doit recevoir des instructions précises, notamment sur l'oreille testée en premier, le type d'items à entendre et le mode de réponse (orale, écrite, choix sur un écran...), la nécessité de fournir une réponse aussitôt que l'item a été perçu, aussi faible soit-il, en sachant qu'une seule réponse est admise.

Si la réponse est orale, l'examineur doit pouvoir la comprendre.

En cas d'audiométrie en voix naturelle, il faut respecter un intervalle entre items de l'ordre de 4 à 4,5 s pour que la courbe de référence reste valable.

3.2. Mesure du seuil d'intelligibilité

Il faut utiliser des listes d'items complètes, et familiariser le sujet avec la méthode de réponse (répéter l'item en général).

1. Commencer à un niveau d'audition de la parole supérieur de 20 à 30 dB à la moyenne des seuils tonals du sujet à 500 Hz, 1 et 2 kHz (moyenne des fréquences conversationnelles – MFC), et présenter au moins deux items (en pratique, 3 sont suffisants).
2. Réduire le niveau de 5 dB en présentant au moins deux items, jusqu'à l'apparition d'une réponse incorrecte.
3. A ce niveau, présenter une liste complète d'au moins dix items et noter le score.
 - a. Si le score dépasse 50%, baisser de 5 dB et recommencer jusqu'à obtention d'un score < à 50%. Le seuil d'intelligibilité se détermine par interpolation.
 - b. Si le score n'atteint pas 50%, augmenter de 5 dB et recommencer jusqu'à obtention d'un score > à 50%. Le seuil d'intelligibilité se détermine par interpolation.

Remarque : La norme précise qu'il est possible d'utiliser un pas de progression de 2 dB au lieu de 5.

Enfin, la norme propose une méthode alternative que nous ne décrivons pas ici, car elle est assez complexe et ne semble pas présenter plus d'intérêt.

3.3. Détermination des scores de reconnaissance vocale

Ceci est nécessaire pour tracer la courbe d'audiométrie vocale complète, ou simplement pour déterminer le score maximal de reconnaissance vocale.

Détermination du score maximal : si le seuil d'intelligibilité a été déterminé au préalable, on présente une liste d'items à 25 ou 30 dB au-dessus de ce seuil.

Si le maximum (100%) n'est pas atteint, augmenter de 5 en 5 dB jusqu'à obtention du score maximal ou jusqu'à une sensation d'inconfort.

Détermination du niveau de confort : On cherche à obtenir une appréciation sur la sonie du message reçu (par exemple, beaucoup trop faible, confortable, trop fort, beaucoup trop fort). En général, le point médian entre trop faible et trop fort constitue un bon choix du niveau de confort optimal.

Courbe d'intelligibilité vocale. On procède comme pour la détermination du seuil d'intelligibilité, en commençant 20 à 30 dB au-dessus de la MFC, mais en présentant des listes complètes avec un intervalle de 5 dB (ou de 2 dB) en augmentant et en baissant le niveau. Généralement, 5 à 7 points sont suffisants.

Présentation de l'audiogramme vocal. Si on décide de présenter les résultats sous forme graphique, on doit porter en ordonnée le score de reconnaissance vocale (en %) et en abscisse le niveau vocal (SPL) ou le niveau d'audition de la parole (correspondant à un niveau HL pour la vocale ou SHL). Le rapport d'échelle doit être de 20% pour 10 dB (Fig. 1).

Note : La norme ne donne pas de préférence pour l'échelle des dB, mais il est logique de ne faire figurer sur l'audiogramme vocal que l'échelle de niveau d'audition de la parole en dB SHL.

Il est possible de masquer l'oreille controlatérale, notamment quand le test est effectué à l'aide d'écouteurs. Cet aspect sera décrit en détails dans une prochaine mise au point consacrée à l'audiométrie tonale.

4. Audiométrie vocale dans le bruit

Cet examen, très intéressant, est sensible à la nature du bruit de fond choisi. Les valeurs de référence fournies avec l'enregistrement utilisé ne peuvent pas être appliquées à un autre bruit⁸.

Le son vocal et le bruit doivent être présentés dans le même écouteur.

En champ acoustique, il est recommandé de présenter le son vocal à l'aide d'un haut-parleur situé en position frontale à 0° d'incidence. Le bruit est alors émis par deux haut-parleurs, sans cohérence acoustique, situés symétriquement à 45° en avant ou en arrière. Toute autre disposition est possible, mais elle doit alors être précisée.

4.1. Niveaux du signal vocal et du bruit

Les niveaux doivent être mesurés avec la même pondération et le même temps d'intégration (L_{ceq}).

Le niveau vocal recommandé est de 65 dB (niveau d'une conversation).

La valeur recommandée pour le niveau du bruit de fond, s'il est constant, est de 60 dB. S'il est variable, le pas de variation doit être au maximum de 5 dB.

4.2. Méthodes de mesure

On peut vouloir déterminer des scores de reconnaissance vocale pour plusieurs valeurs de P/B, ou la valeur du rapport P/B qui correspond au seuil d'intelligibilité vocale. Dans les deux cas, on peut faire varier le niveau du bruit ou celui de la parole. Dans ce qui suit, on choisit de fixer le niveau de bruit à 60 dB et de faire varier le niveau vocal, comme on le fait dans le silence.

En pratique, il est souvent suffisant de mesurer le **seuil d'intelligibilité dans le bruit**. Pour ce faire, on procède comme précédemment, en fixant le rapport P/B de départ à zéro, après un test de familiarisation bien sûr.

Noter qu'il peut être préférable, ici, de procéder par pas de 2 dB. Le seuil d'intelligibilité vocale dans le bruit servira éventuellement de repère pour tracer une courbe complète, qui peut se présenter sous forme graphique (Fig. 2).

Noter que la norme ne précise pas le rapport d'échelle pour une audiométrie vocale dans le bruit. Il semble difficile d'utiliser le même que pour la vocale (20% pour 10 dB), d'autant que l'exemple donné en annexe avec des phrases de Göttingen donne un rapport d'échelle 100% pour 5 dB.

Conclusion. Quand les conditions minimales sont réunies : bonnes conditions matérielles de mesure, langue utilisée bien comprise, bons enregistrements, audiométriste expérimenté... le seuil d'intelligibilité vocale dans le bruit, est le test auditif le plus efficace, d'après Hagerman, avec un écart-type de 1,02 sur la reproductibilité, comparé à presque 3 pour la tonale et le SRT avec spondées.

Point de vue de l'audioprothésiste

L'audiométrie vocale a pour objectif principal de tester l'aptitude d'un sujet à percevoir et à comprendre la parole dans une langue donnée.

Les problèmes d'intelligibilité de la parole éprouvés par un malentendant sont le cœur de l'activité de l'audioprothésiste. Le bilan et les tests en audiométrie vocale prothétique sont donc la clé de voute de notre métier. Nous n'appareillons pas une perte auditive, mais nous nous intéressons à la gêne sociale et fonctionnelle qu'elle peut engendrer en contribuant à sa réhabilitation.

La grande pluralité et diversité qui existent en matière de matériel vocal (listes de mots, de chiffres, de logatomes, de phrases,...) ou en matière de méthode utilisée, sont des éléments qui peuvent être à l'origine d'une variabilité de mesures non négligeable.

En effet, les tests subjectifs d'audiométrie vocale prothétique possèdent de grandes variabilités intra-individuelles, ils sont testeurs-dépendants et même « lieux et heure dépendant » !

Comment comparer ou suivre l'évolution de telles mesures si elles ne sont pas réalisées dans les mêmes conditions et qu'elles ne répondent pas à une standardisation ? C'est à partir de ce postulat que les normes prennent toute leur place.

L'existence de normes et leurs correctifs conditionnent donc une approche rigoureuse permettant d'obtenir un niveau de qualité élevé en matière d'audiométrie vocale prothétique.

La connaissance, le respect et l'application stricte de ces normes sont des éléments primordiaux pour limiter et encadrer la variabilité de mesures précédemment décrites (reproductibilité) inhérentes aux caractères subjectifs de ces tests.

Morgan POTTIER, Audioprothésiste – Narbonne.

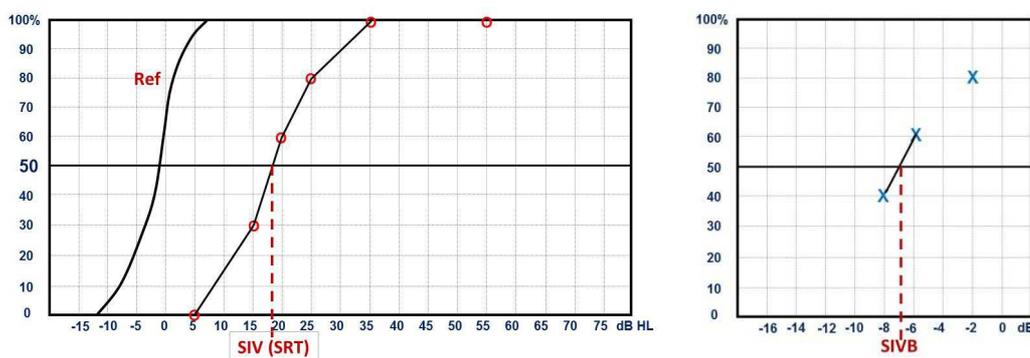
Point de vue de l'ORL

A venir

Dr Martine OHRESSER, ORL - Paris

Fig. 1- Représentation graphique d'un audiogramme vocal de l'oreille droite (○). Pour l'oreille gauche, les marqueurs seraient des X bleus. On note la courbe de référence dépendant du matériel vocal. Le seuil d'intelligibilité vocale (SIV) est défini par interpolation. Notons que nous avons choisi de représenter l'échelle des niveaux en « niveau d'audition de la parole » (*hearing level for speech*) que nous avons noté simplement dBHL.

Fig. 2 – Représentation graphique d'un audiogramme vocal dans le bruit de l'oreille gauche (X). L'échelle des niveaux représente le rapport parole/bruit (P/B). Il s'exprime en dB sans dimension particulière. Chez le sujet normoentendant, le seuil d'intelligibilité dans le bruit (SIVB), déterminé avec des spondées, varie entre -6 et -4 dB selon les études.



¹ ISO EN NF 8253 : Acoustique — Méthodes d'essais audiométriques — Partie 3: Audiométrie vocale.

² Il faut noter que la norme ne définit pas le score de discrimination (SDS), malgré son intérêt et la fréquence de son utilisation dans les publications scientifiques. C'est le score de reconnaissance vocale à un niveau supérieur de 35 dB au seuil d'intelligibilité (SRT + 35).

³ Les tests vocaux à voix nue, contrôlée par un vumètre restent utiles pour les personnes âgées, les enfants, ceux qui ont des difficultés de compréhension de la langue. Mais il faut savoir que l'incertitude de mesure augmente alors considérablement, dépendant beaucoup de l'expérience de l'audiométriste.

⁴ Un équilibrage phonémique parfait peut parfois s'avérer impossible à obtenir. L'équilibrage phonémique des listes d'items doit alors être réalisé sur la base des classes de phonèmes, c'est-à-dire des plosives et fricatives voisées et non voisées et des voyelles nasales, longues et courtes. La distribution phonémique du message vocal doit correspondre, dans la mesure du possible, à la distribution phonémique de la langue utilisée.

⁵ Pour satisfaire cette exigence, on doit valider les enregistrements en répétant les tests sur un échantillon de dix sujets otologiquement normaux de 18 à 25 ans inclus, de même langue maternelle que la langue du message vocal. Notons que la norme n'évoque pas le sexe des sujets.

⁶ Par exemple le nombre, l'âge et les niveaux liminaires d'audition des sujets, ainsi que le type d'audiomètre vocal utilisé et la disposition des hauts-parleurs pour le bruit de fond, le cas échéant...

⁷ La norme précise que, pour les enfants, il est nécessaire d'établir des courbes d'intelligibilité vocale par classes d'âge, ce qui peut présenter des difficultés conduisant à une plus grande incertitude de mesure.

⁸ Ceci est particulièrement vrai pour les bruits de fonds contenant des voix humaines (cocktail). La plupart des enregistrements disponibles sont obtenus à partir de prises de sons effectuées dans des pays anglo-saxons. L'effet de masque, ou pouvoir de confusion, est bien inférieur à un bruit de cocktail en français, pour des Français. Cela est notamment lié à un pouvoir de masquage temporel différent.