



HAL
open science

Une station de test automatisée pour la mesure de l'intelligibilité et de la qualité de la parole

Christian Cavé, Philippe Di Cristo, Alain Ghio, Bernard Teston

► To cite this version:

Christian Cavé, Philippe Di Cristo, Alain Ghio, Bernard Teston. Une station de test automatisée pour la mesure de l'intelligibilité et de la qualité de la parole. Travaux interdisciplinaires du Laboratoire Parole et Langage, 1995, 16, pp.123-140. hal-01665238

HAL Id: hal-01665238

<https://hal.science/hal-01665238>

Submitted on 18 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TRAVAUX DE L'INSTITUT DE PHONETIQUE D'AIX
Volume 16, 1995, pp. 123-140

**UNE STATION DE TESTS AUTOMATISEE
POUR L' EVALUATION DE LA QUALITE
ET L' INTELLIGIBILITE DE LA PAROLE**

**Christian CAVE
Philippe DI CRISTO
Alain GHIO
Bernard TESTON**

UNE STATION DE TESTS AUTOMATISEE POUR L'EVALUATION DE LA QUALITE ET L'INTELLIGIBILITE DE LA PAROLE

Cavé Christian, Di Cristo Philippe, Ghio Alain et Teston Bernard

Institut de phonétique d'Aix-en-Provence
Laboratoire "Parole et Langage" URA 261, CNRS
29, Av. R. Schuman, 13621 Aix-en-Provence, France

Résumé

L'évaluation de la qualité et de l'intelligibilité de la parole exige le recours à des jurys d'auditeurs. Les procédures utilisées sont délicates à mettre en oeuvre et prennent un temps considérable. Il est donc intéressant d'automatiser autant que possible toutes les opérations nécessitées par ces mesures et de procéder à une passation collective des sujets. Nous présentons une station de test organisée autour d'un micro-ordinateur de type IBM PC qui répond à ces objectifs.

Abstract

Assessing the intelligibility or the quality of speech requires the use of psychophysical methods. Recourse to panels of listeners and perception tests is indispensable.

This type of testing procedure is very time-consuming, hence the need to automate as many test operations as possible, such as stimulus generation, procedure control, and data processing, and to increase the number of subjects that can be tested at once.

The system we are presenting here meets these objectives. It is comprised of a computer-driven control station run by an operator and response stations for up to sixteen subjects. It automatically presents the auditory stimuli (words or other sounds) to the subjects, displays the instructions and/or response choices on video screens, and records each subject's responses and response times. The considerable amount of time gained by this multilisteners test administration procedure makes it possible to obtain very large quantities of perceptual data.

UNE STATION DE TESTS AUTOMATISEE POUR L'EVALUATION DE LA QUALITE ET DE L'INTELLIGIBILITE DE LA PAROLE

Cavé Christian, Di Cristo Philippe, Ghio Alain et Teston Bernard

Institut de phonétique d'Aix-en-Provence
Laboratoire "Parole et Langage" URA 261, CNRS
29, Av.R.Schuman, 13621 Aix-en-Provence, France

Introduction

L'évaluation de l'intelligibilité ou de la qualité de la parole, exige l'intervention de procédures psychophysiques. En effet, aucune méthode physique actuelle ne peut s'appliquer à toutes les conditions d'écoute ou de transmission des divers types de parole, et même les méthodes objectives nécessitent toujours, au moins dans un premier temps, une validation par comparaison avec des méthodes subjectives (Steeneken, 1987 ; Morellini, 1994). Pour cela, le recours à des jurys d'auditeurs, sous la forme de tests de perception, reste indispensable.

Ce type de procédure de mesure exige toujours un temps considérable, du fait de l'utilisation de plusieurs locuteurs et auditeurs pour tenir compte de la variabilité individuelle, et à cause du grand nombre de répétitions des mesures nécessaires à la validation statistique des résultats. On comprend dès lors l'intérêt d'automatiser autant que possible toutes les opérations de ces tests d'évaluation, telles que la génération des stimulus, le contrôle des procédures et le traitement des résultats. Le dispositif intégré sous la forme d'une station de travail que nous présentons répond à ces objectifs.

1. Description de la station de test

1.1. Présentation générale

La station de test est constituée par un micro ordinateur de type IBM PC et divers matériels de diffusion des stimulus et de contrôle permettant de tester jusqu'à 16 sujets simultanément. Cet ensemble de matériel et les logiciels associés réalisent les opérations suivantes:

- Génération des stimulus sonores qui sont présentés aux sujets.
- Edition de messages qui sont affichés sur des écrans face aux sujets.
- Recueil des réponses des sujets par l'intermédiaire de claviers.
- Conduite de toute la procédure de déroulement des tests.

- Sauvegarde, traitement et présentation des résultats.

Toutes ces opérations sont mises en oeuvre au moyen de programmes de pilotage et de gestion, qui fonctionnent dans l'environnement WINDOWS de Microsoft (Levingston, 1993) en exploitant au maximum ses potentialités comme le multimédia audio et ses grands outils tels que tableurs et gestionnaires de bases de données.

1.2. Fonctionnement général

Chaque sujet dispose d'un écran monochrome alphanumérique sur lequel sont présentées des informations telles que des consignes ou des réponses possibles, d'un clavier de réponse à cinq touches et d'un casque audiophonique.. Le déroulement d'un test consiste à envoyer un stimulus sonore aux sujets, soumis au préalable à des suggestions de réponses possibles sur l'écran. Les sujets donnent leur réponse en appuyant sur l'un des cinq boutons du clavier à leur disposition. Il est possible d'effectuer la mesure du temps de réponse,

Dans sa configuration actuelle, la station peut tester huit sujets simultanément; elle est extensible jusqu'à seize. Cette passation collective de tests permet un gain de temps considérable et une uniformisation des conditions expérimentales par rapport à une installation où un seul sujet est soumis au test. De plus, le recueil des réponses étant automatique, leur traitement ultérieur ne nécessite que très peu de manipulations.

2. L'architecture matérielle de la station

La station de test est composée des éléments suivants (Figure 1) :

- Un ordinateur de type PC 486 DX 33, qui pilote l'ensemble du dispositif, un écran graphique SVGA, 8 Mo de mémoire vive et un disque dur de 400 Mo.
- Trois cartes additionnelles spécialisées:
 - Une carte écran monochrome "Hercules"
 - Une carte d'entrées/sorties numériques National Instruments PC-DIO-96
 - Une carte audio multimédia de type Sound Blaster 16 SP
- Une interface, développée spécialement pour la station de test, qui sert de liaison entre l'unité centrale et les écran monochromes, les claviers de réponse et les casques audiophoniques.
- Huit claviers de réponse.
- Huit moniteurs monochromes.
- Huit casques audio.

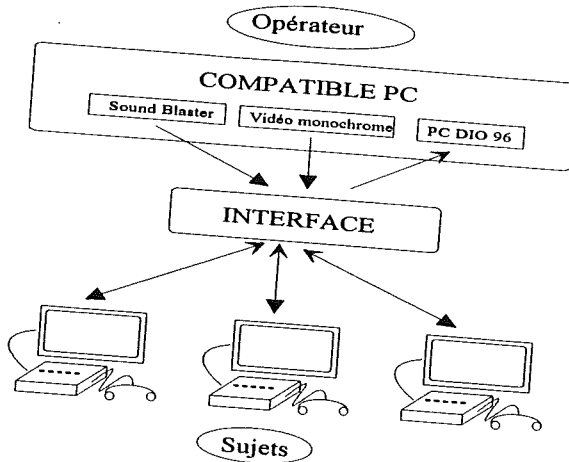


Figure 1: Architecture matérielle de la station

2.1. Le poste de pilotage

- D'un point de vue matériel, l'ordinateur est chargé des opérations suivantes :
- Générer les stimulus sonores qui sont diffusés aux sujets
 - Editer les messages alphanumériques qui sont affichés sur les écrans monochromes.
 - Recueillir les réponses des sujets par l'intermédiaire des claviers.

2.1.1. La source sonore

La source sonore est assurée par une carte audio de type Sound Blaster 16 SP dont les caractéristiques techniques permettent un codage mono ou stéréo sur 8 ou 16 bits, à des fréquences d'échantillonnage de 11, 22, ou 44 KHz (Stolz, 1994). Les stimulus présentés aux sujets sont stockés sous la forme de fichiers "échantillons" en mémoire de masse. Au moment du test, ils sont adressés à la carte, qui les convertit en signaux analogiques diffusés aux casques après amplification de puissance.

Les fichiers "échantillons" peuvent être acquis directement avec la carte Soundblaster, soit par l'intermédiaire de l'entrée microphone, soit à partir de l'entrée ligne en connectant un magnétophone ou un lecteur D.A.T. Les stimulus peuvent également provenir de n'importe quel support informatique (CD ROM, bandes magnétiques, disques..) pourvu que les fichiers importés soient, ou puissent être convertis au format '.WAV'. Ce format répond aux normes RIFF (Resource Interchange File Format) développé par Microsoft (Microsoft, 1991) et

apparaît comme le standard audio sous Windows. Les fichiers de type .WAV contiennent non seulement les échantillons proprement dits, mais aussi un entête décrivant les caractéristiques du signal: fréquence d'échantillonnage, nombre d'octets par échantillon, mode mono/stéréo...

2.1.2. La source d'affichage

Une carte vidéo monochrome de type "Hercules" permet de générer les messages alphanumériques présentés aux sujets. Ces messages peuvent être des consignes ou des propositions de réponses. Pour utiliser la carte monochrome, il est nécessaire d'écrire à une adresse mémoire protégée dans l'environnement WINDOWS, que nous utilisons. Aussi avons-nous réalisé un programme résident qui permet d'accéder indirectement aux adresses mémoires utiles (Brown et Kyle, 1991). Cet affichage alphanumérique monochrome est géré indépendamment du moniteur couleur classique de type SVGA.

2.1.3. Le recueil des réponses

Nous avons choisi de pouvoir tester jusqu'à seize sujets simultanément avec cinq possibilités de réponse; quatre vingts voies sont donc nécessaires. Pour les contrôler (Gupta & Toong, 1985), nous avons choisi une carte de National Instruments PC-DIO-96, qui comprend 96 entrées/sorties numériques parallèles conçue autour de 4 interfaces périphériques programmables (PPI) de type AMD8255A. A chaque entrée de la carte est connecté un bouton poussoir situé sur un clavier de réponse. En scrutant régulièrement les entrées de la carte PC-DIO96, il est donc possible de détecter les boutons qui sont actionnés sur les claviers.

2.2. L'interface de liaison PC - postes de test

D'une manière générale, toute expérience pilotée par un ordinateur nécessite un interface entre le domaine physique à appréhender et l'environnement numérique de l'ordinateur. Pour notre application, nous avons réalisé un interface qui se connecte d'une part au PC de pilotage, d'autre part à chacun des postes de test (Figure 1). Il assure les fonctions suivantes :

- le conditionnement audio

L'amplification et le contrôle du niveau du signal issu de la sortie ligne de la carte Sound Blaster sont réalisés à travers un amplificateur de puissance capable d'alimenter 16 casques. Celui-ci est caractérisé par un niveau constant à + ou - 0,2 dB quel que soit le nombre de casques qui lui sont connectés. Il dispose d'un atténuateur calibré permettant d'ajuster les niveaux d'écoute.

- le conditionnement vidéo

Dans une utilisation classique, la liaison vidéo entre la carte alphanumérique "Hercules" et le moniteur monochrome ne peut pas être d'une longueur supérieure à quelques mètres. De plus, la sortie ne supporte qu'un seul écran. Pour notre utilisation spécifique, nous avons donc intégré, dans l'interface, 16 amplificateurs de ligne vidéo pour contrôler directement chaque sortie vers les écrans d'affichage et ceci, sans limitation stricte de la longueur des câbles de liaisons.

- la saisie de l'état des boutons poussoir des claviers de réponses des sujets

Cette opération ne présente pas de problème majeur. La seule précaution particulière consiste à traiter les rebonds à cause des grandes longueurs des câbles de liaison, ainsi que les interférences électromagnétiques susceptibles de provoquer des déclenchements erronés.

2.3. Les postes de test

Chaque poste de test se compose d'un clavier de réponses, d'un casque audiophonique et d'un moniteur monochrome. Des câbles de connexion assurent la liaison bidirectionnelle de chaque poste à l'interface décrit précédemment. Par ces connexions transitent d'une part les signaux sonores et vidéo de l'interface vers les postes, d'autre part les réponses des sujets dans le sens inverse. Grâce aux conditionnements électroniques effectués par l'interface, les câbles peuvent atteindre une longueur de plusieurs dizaines de mètres, ce qui permet de disposer les postes de test dans une grande salle.

Les casques audiophoniques utilisés sont des BEYER type BT 100 électrodynamiques. Nous les avons choisi pour leur bonne sensibilité et leur bonne réponse en fréquence, leur faible dispersion d'impédance qui dispense de les calibrer unitairement et pour leur grande robustesse.

Pour la diffusion des informations alphanumériques destinées aux sujets, nous avons choisi un affichage sur un écran vidéo. Compte tenu de la définition de l'écran nécessaire pour les messages affichés (Galitz, 1987) et des contraintes de prix de revient, notre choix s'est porté sur un moniteur monochrome. (Philips modèle BM7)

Les claviers de test comportent cinq boutons poussoirs qui permettent de saisir les réponses des sujets. Chaque clavier est contenu dans un boîtier qui sert également de relais à la connexion du casque et du moniteur alphanumérique monochrome. Les boutons poussoir sont de type micro-rupteur à bascule (switch), caractérisés par une faible course, un contact très franc et un temps de commutation rapide bien adapté à la mesure des temps de réaction.

3. Le fonctionnement logiciel de la station

D'un point de vue logiciel, la station de test a été développée dans l'environnement WINDOWS de Microsoft (Petzolt, 1990 ; Norton et al. 1992) pour sa facilité d'emploi, de compatibilité avec d'autres logiciels et sa forte diffusion. La longue expérience du laboratoire dans la mesure de l'intelligibilité de la parole et de la réalisation ces dernières années de campagne de mesure à grande échelle (Cavé et al., 1992, 1994), nous ont permis de concevoir un système permettant un maximum de souplesse par ses possibilités de paramétrisation et d'efficacité par l'automatisation des différentes opérations.

3.1. La vérification de l'installation matérielle

La station de test, dans son ensemble, implique l'emploi d'une grande variété de matériels: écrans, casques, boîtiers, câbles, interface, cartes spécialisées, ordinateur, alimentation électrique. Malgré la robustesse de l'équipement choisi, la panne d'une partie de ce matériel ne demeure pas improbable. De même, un mauvais branchement reste possible. Nous avons donc prévu des utilitaires permettant de vérifier l'ensemble du matériel avant chaque expérimentation. Ceux-ci autorisent le test de l'installation sonore, de l'ensemble de distribution vidéo et de chaque bouton des boîtiers de réponses. Cette aide au diagnostic s'est révélée, à l'usage, très utile..

3.2. La paramétrisation des conditions de l'expérimentation

Lors de la réalisation logicielle, nous avons cherché à rendre le plus paramétrable possible les fonctions de la station de test. Il existe pour cela, une phase préliminaire à chaque expérimentation, qui consiste à configurer le dispositif: définition du nombre de postes utilisés, plan d'occupation de ces postes, configuration des boutons, conditions de test, informations sur les sujets... Il est aussi possible de préciser les répertoires où sont stockés les fichiers d'échantillons de parole, d'indiquer où doivent être enregistrées les réponses... Des utilitaires permettent à l'opérateur de choisir une configuration standard ou personnalisée. Il peut, par exemple, paramétrer le dispositif pour plusieurs groupes de sujets ou encore autoriser seulement les boutons 1 et 5 des boîtiers lors d'un test à réponses binaires.

3.3. La notion de "fichiers pilotes"

Le déroulement d'un test est contrôlé par un fichier "script". Cette façon de procéder s'inspire du dispositif S.O.A.P. (rapport SAM, 1992). En fait, le coeur logiciel de la station de test est construit comme un analyseur lexico-syntaxique qui lit et interprète le fichier "script" qu'on lui envoie. Il adapte alors le

déroulement de l'expérience en fonction des commandes qu'il reçoit. Ce type de fonctionnement permet une grande souplesse dans la variété des tests. A chaque ligne (Figure 2), on trouve un mot clé suivi d'un argument qui spécifient l'opération à effectuer.

<u>Commandes préparatoires</u>		
CONDITIONS	helium.180m.tourelle.repetition1	=> conditions du test
CONSIGNES	C:\perceval\bin\consigne.txt	=> fichier de consignes
TYPE_TEST	1	=> codage du test ¹
RÉSULTATS	C:\perceval\reponses\he180T1	=> stockage des réponses ²
DELAI	2000	=> délai d'attente autorisé après écoute (en ms)
PAUSE	1000	=> pause entre deux propositions (en ms)
DEBUT		=> début du balayage
#----		
<u>Balayage</u>		
STIMULUS	C:\PERCEVAL\STIMULI\pile.wav	=> fichier sonore à faire écouter
AFFICHAGE	1) bile 2) pile	=> chaîne de caractère qui apparaît sur les écrans
REPONSE	5 ³ 2 ⁴ pile.bile ⁵	=> codage des résultats
CORRECT	2	=> n° de la réponse attendue
CODAGE	+Vois -Vois	=> codage pour les erreurs (Voisé/non voisé)
CONTEXTE	_i	=> invariant (contexte /i/)

STIMULUS	C:\PERCEVAL\STIMULI\vingt.wav	
AFFICHAGE	1) vingt 2) bain	
REPONSE	11 2 vingt.bain	
CORRECT	1	
CODAGE	-Inter +Inter	=> codage pour les erreurs (Non interrompu/interrrompu)
CONTEXTE	__~E	=> invariant (contexte /E/)

...		

FIN		=> fin du test

Figure 2: Exemple de fichier de pilotage

¹ type 1 = tests en N-uplets

² à chaque sujet correspond un fichier de réponses du type he180T1.REZ pour le sujet dont les initiales sont EZ, he180T1.RCA pour le sujet dont les initiales sont CA...

³ numéro de la paire

⁴ nombre de candidats (2 = paire, 3 = triplet, 4 = quadruplet...)

⁵ liste des candidats

Là encore, la paramétrisation permet une grande souplesse dans le déroulement des tests :

- le choix du fichier de consignes autorise, par exemple, une utilisation multilingue.
- le nom des fichiers où seront inscrites automatiquement les réponses des sujets (cf. §. 3.7) est paramétrable.
- le test peut être de type N-uplets, préférences ou échelles (cf. § 3.4).
- le délai de réponse autorisé est ajustable.
- la pause entre chaque présentation sonore est paramétrable.
- la mesure du temps de réponse peut se faire en début ou en fin de présentation du stimulus.
- des modules de brassage permettent une présentation aléatoire de l'ordre des stimulus et des propositions à l'écran.

3.4. Les types de test

Compte tenu de la passation collective, on ne peut pas utiliser de procédures adaptatives. A part cela, il n'existe pas de contrainte sur le type de test réalisable ou sur la façon de mettre en oeuvre un test particulier. On peut toujours écrire un fichier pilote spécifique pour contrôler la procédure que l'on veut réaliser. Actuellement trois types de tests sont disponibles.

- test en N-uplets

Plusieurs réponses possibles sont proposées à l'écran, puis le sujet reçoit un stimulus sonore. Par exemple, on fait écouter le mot "baffe" et on propose 1-basse 2-baffe 3-bache. La réponse attendue est le deuxième mot. Parmi ce type de tests, figure le test de paires minimales (Peckels & Rossi, 1973)

- test de préférences

Le sujet entend plusieurs stimulus. Par exemple, on fait écouter la phrase "demain, il fera beau" selon 3 versions (Synthèse A, B ou C). Le sujet indique celle qu'il préfère en pressant le bouton correspondant.

- test d'échelle

Le sujet reçoit un stimulus sonore. Plusieurs évaluations de ce stimulus sont proposées à l'écran. Par exemple, le sujet doit choisir une note sur une échelle de 1 à 5, ou bien choisir entre faible, moyen ou fort.

3.5. Le codage des erreurs

Dans le cas d'un test en N-uplets, où une réponse correcte est attendue, il est

possible d'obtenir un codage des erreurs commises par les sujets. Cela facilite l'analyse des résultats en classant les erreurs par types, par contexte, etc... Il faut, pour cela, fournir au dispositif une liste complète contenant non seulement les mots mis en compétition mais aussi leur codage. Le principe de codage est le suivant :

Mot1	mot2
code1	code2
contexte	

Exemple pour un test de paires minimales :

pile	bile	
-voix	+voix	=> Voisé/non voisé
_i		=> contexte /i/
vingt	bain	
-inter	+inter	=> Non interrompu/interrupted.
_~E		=> contexte /E/

3.6. La mesure du temps de réponse

Le temps de réponse peut être un paramètre crucial pour l'évaluation de la difficulté de la tâche et pour l'interprétation des résultats (Talbot, 1987).

Pour assurer la précision de la mesure des temps de réponse, nous avons utilisé les outils 'Multimédia' de WINDOWS (Microsoft, 1991). La bibliothèque 'mmsystem' autorise le réglage de la période d'accès à l'horloge interne de l'ordinateur. Ainsi, il est possible d'obtenir des résolutions inférieures à 50 ms, qui est la période de base. En fait, le microprocesseur agit sur les mécanismes de l'horloge par le biais d'une interruption dont le rôle consiste à interrompre une action en cours pour réaliser un traitement prioritaire (sonder les boîtiers de réponses dans notre cas). L'utilisation d'une interruption nous permet d'être certain que le traitement que nous voulons effectuer se réalisera à intervalles réguliers et fixes. La mesure du temps de réponse est donc fiable à la milliseconde près.

3.7. Les résultats

Pour chaque sujet on crée un fichier de réponses, qui contient les informations nécessaires aux traitements des données. Dans le cas d'un test en quadruplets par exemple, on trouve les informations suivantes: le nom du script, le codage des conditions expérimentales, l'identification du sujet, le numéro du quadruplet, le nombre de candidats, les candidats, la réponse attendue, la réponse

donnée par le sujet, le codage de la réponse (non réponse = 0, réponse correcte = 1, réponse incorrecte = -1), le temps de réponse du sujet, et éventuellement le codage de l'erreur et son contexte (Figure 3).

Fichier script	Conditions	Sujet	n° N-uplet	Candidats	Réponse attendue		Réponse donnée	
								Temps de réaction
CAPERCEVALSCRIPTS\HE180.SCR	helium.180m	AA	2 4	huile pile mille bile	2	0	0	X
CAPERCEVALSCRIPTS\HE180.SCR	helium.180m	AA	4 4	sain.vingt.faim.rein	2	2	1	1017 X
CAPERCEVALSCRIPTS\HE180.SCR	helium.180m	AA	3 4	baffe.base.bave.basse	4	3	-1	1720 v/s ba_
CAPERCEVALSCRIPTS\HE180.SCR	helium.180m	AA	5 4	rang.Zan.dans.gens	3	1	-1	1781 r/d -a
CAPERCEVALSCRIPTS\HE180.SCR	helium.180m	AA	1 4	doigts.loi.bois.poids	2	2	1	1245 X

Nb de candidats

Codage réponse
 1=correct
 -1=incorrect
 0=non réponse

Contexte
 Codage erreur

Figure 3: Exemple de fichier de réponses

Les fichiers sont remplis automatiquement en cours de test. Ils sont construits en format "tabulé", ce qui permet leur importation immédiate dans un tableur tel que QUATTRO PRO, EXCEL (Figure 4) ou dans un logiciel d'analyse statistique tel que STATWIEW, SYSTAT etc...

	A	B	C	D	E	F	G	H	I	J	K	L
	CAPERCEVALSCRIPTS\HE180.SCR	helium	YY	2	4	huile pile mille bile	2	3	-1	1024	rn/p	il
1	CAPERCEVALSCRIPTS\HE180.SCR	helium	YY	4	4	sain vingt faim rein	2	2	1	1099	X	
2	CAPERCEVALSCRIPTS\HE180.SCR	helium	YY	3	4	baffe base bave basse	4	1	-1	1591	/s	ba_
3	CAPERCEVALSCRIPTS\HE180.SCR	helium	YY	5	4	rang Zan dans gens	3	3	1	2077	X	
4	CAPERCEVALSCRIPTS\HE180.SCR	helium	YY	1	4	doigts loi bois poids	2	1	-1	1500	d/	wa_

Figure 4 : Extrait d'un fichier de réponses importé dans le logiciel Excel

Il est alors possible d'effectuer tous les traitements courants ainsi que des analyses détaillées. On peut ainsi, par exemple, dénombrer les erreurs ayant porté sur la proposition n°10 pour tous les sujets dans telle et telle condition de test. Le codage des erreurs permet, par exemple, d'obtenir le nombre d'erreurs ayant porté sur le trait de voisement en contexte [i] dans une condition de test donnée.

4. L'utilisation de la station de test

Le fonctionnement logiciel de la station de test est semblable à celui de n'importe quel autre produit Windows. Il existe un icône pour le lancement, une fenêtre principale, des fenêtres d'informations, des menus déroulants et des boîtes de dialogue. La fenêtre principale fournit différents utilitaires pour le bon fonctionnement de l'expérimentation: vérification de l'installation matérielle, configuration de la session de test, conception des fichiers "script" (Figure 5), déroulement du test, dépouillement des résultats.

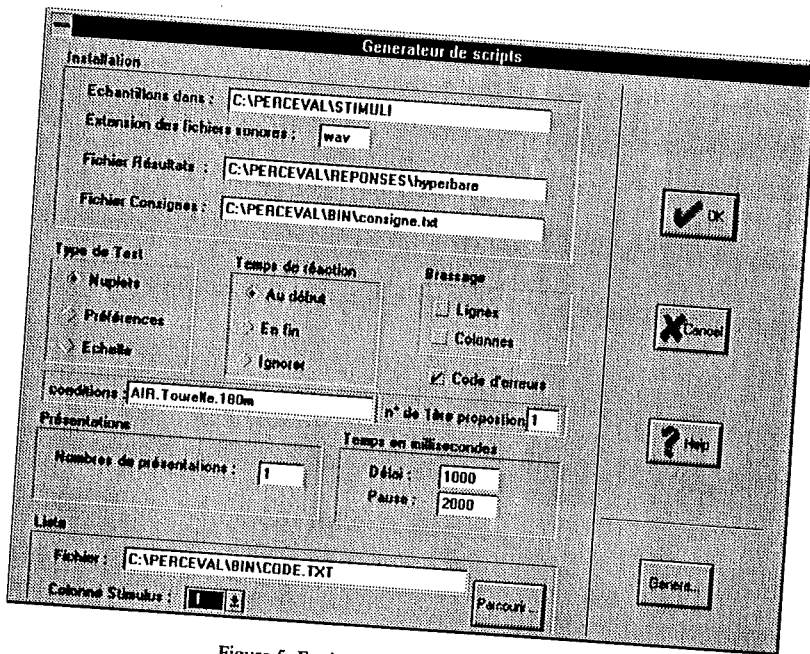


Figure 5: Fenêtre de génération de "script"

- Pendant le déroulement de l'expérimentation, l'opérateur dispose d'un écran de surveillance (Figure 6) qui fournit plusieurs informations :
- un vumètre indiquant l'écoulement du temps,
 - une zone centrale d'affichage reproduisant le message envoyé aux écrans des sujets
 - huit cases représentant les postes de réponse des sujets. Dans chaque case sont indiquées l'identité du sujet, les réponses qu'il donne, les pourcentages de bonne/mauvaise/non réponses et le temps de réponse.

Ce type d'écran permet de détecter d'éventuelles déficiences de la part du matériel ou de possibles "manques d'attention" de la part des sujets. Dans la configuration à seize postes, il est possible de basculer l'affichage entre les postes un à huit et les postes neuf à seize.

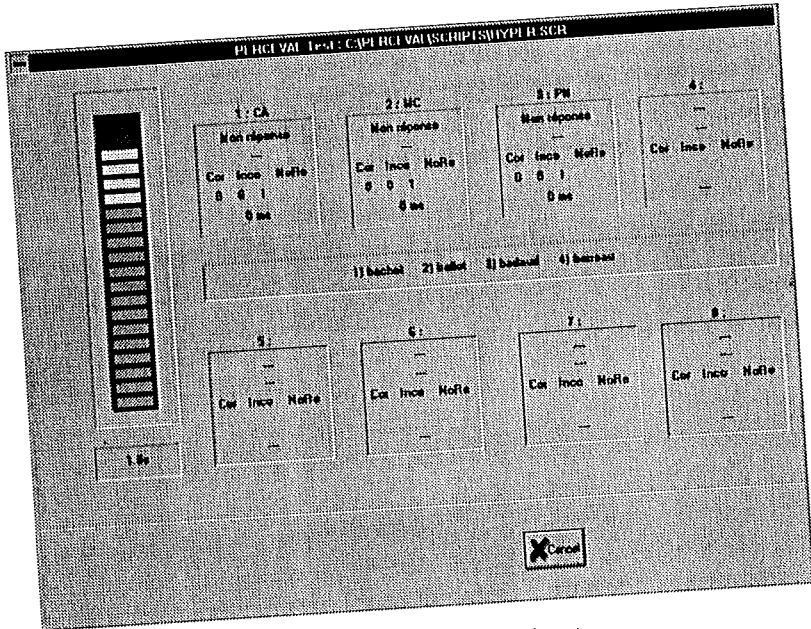


Figure 6 : Fenêtre de contrôle du test

En fin de session de test, l'opérateur peut demander un dépouillement brut des résultats, c'est-à-dire le pourcentage global de bonnes et mauvaises réponses, ainsi que le taux de non-réponses. Il pourra ensuite poursuivre l'analyse en utilisant un logiciel de traitement de données (cf. § 3.7.). Il n'est pas nécessaire de disposer de tout le matériel (interface, cartes spécialisées, postes de test) pour que la partie logicielle de la station fonctionne. Ainsi, la génération de scripts ou le dépouillement des résultats peut s'effectuer sur un autre poste.

Conclusion

La station de test a déjà été utilisée pour l'évaluation de l'intelligibilité et la qualité de la parole dans le cadre de plusieurs projets. Lors d'une campagne de

mesures de l'intelligibilité de parole codée, près de 750 000 réponses ont été recueillies, ce qui aurait été impensable sans un système automatisé. La station de test a aussi été utilisée pour évaluer l'intelligibilité de la parole en milieu subaquatique et hyperbare (Cavé et al., 1994). Toutes ces expériences nous ont permis d'acquérir un savoir qui a été réinvesti pour améliorer les performances et la fiabilité de la station de test.

Remerciements

Nous tenons à remercier les personnes suivantes :

M. Chérigüène a réalisé le câblage et la connectique de la station de test,
 R. Espesser a contribué au bon fonctionnement de la version précédente.
 C. Meunier, principale utilisatrice de la version précédente,
 L. Tendo et F. Germonneau, étudiants à l'I.U.T. d'informatique d'Aix-en-Provence, qui ont participé à l'élaboration logicielle de la station de test dans le cadre de projets de fin d'étude.

Bibliographie

- Brown, R.; Kyle, J. (1991), *PC Interrupts*, Addison Wesley, Reading, pp 520.
- Cavé, C.; Marchal, A.; Meunier, C. (1992) *Constitution d'un test d'intelligibilité pour la parole hyperbare*. Rapport final du contrat DRET n°90-1201.
- Cavé, C.; Di Cristo, P.; Ghio, A.; Marchal, A.; Meunier, C. (1994) *Intelligibilité de la parole subaquatique et hyperbare*. Rapport final. Contrat FSH/CEP&M 4737/90 et M4737/91.
- Galitz, W.O., (1987), *Handbook of screen format design*, North Holland, Amsterdam, pp 248.
- Gupta, A.; Toong, O., (1985), *Insights into personal computer*, Elsevier, New York, pp 380.
- Levingston, B., (1993), *WINDOWS 3.1 Shareware and Freeware*, IDG Book, San Mateo, pp 644.
- Microsoft, (1991), *Multimedia programmer's workbook*, Microsoft Press, chap. 3 et 4, Redmond, pp 456.
- Microsoft, (1991), *Multimedia programmer's references*, chap. 8, Microsoft Press, Redmond, pp 512.
- Morellini, L., (1994), Etalonnage d'une mesure objective d'intelligibilité par des tests subjectifs, *Actes du 2ème Congrès Français d'Acoustique, Suppl. Journal de Physique III, Vol. 2, CI, 339-342*.
- Norton, P.; Yao P., (1992), *Borland C++ Programming WINDOWS*, Barton Book, Los Angeles, pp 837.
- Petzolt, C., (1990), *Programming WINDOWS*, Microsoft Press, Redmond, pp 1024.

- Peckels, J.P.; Rossi, M. C., (1990), Le test de diagnostic par paires minimales, adaptation au français du "Diagnostic Rhyme Test" de VOIERS, *Revue d'Acoustique*, n°27, 245-262.
- Rapport SAM, 1992, Speech Output Assesment Package (SOAP), V.4.0, Esprit Project n°2589.
- Steeneken, H., J., M., 1987, "Comparisons among three subjective and one objective intelligibility test", *Rapport in IZF 1987-88*, pp 33, TNO, Institute of perception, Soenteeberg.
- Talbot, M., 1987, "Reaction time as a metric for the intelligibility of synthetic speech", in *Speech and Language based interactions with machines*, Waterwoth, J., A. and Talbot, M.(Eds),Ellis Horwood Lt, 66-88.