

# Comprendre et agir : les tests pragmatiques de compréhension de la parole et EloKanz\*

Lionel FONTAN<sup>1</sup>, Pascal GAILLARD<sup>1</sup>, Virginie WOISARD<sup>1,2</sup>  
U.R.I. Octogone (EA4156)<sup>1</sup>, C.H.U. Toulouse-Rangueil<sup>2</sup>

## 1. Introduction

Dans toute situation de communication verbale, l'objectif principal du locuteur est de transmettre des informations à l'auditeur – que celles-ci concernent des idées, des événements, des sentiments ou des sensations. Dans ce cadre la compréhension de l'auditeur peut être considérée comme l'aboutissement de l'acte de communication verbale, et la performance du locuteur peut être jugée en évaluant l'aptitude de l'auditeur à comprendre, c'est-à-dire à saisir et à interpréter les paroles qu'il a prononcées.

Dans tous les domaines où l'on s'intéresse aux performances de locuteurs dans une situation de communication, il est donc crucial de disposer de méthodes fiables, valides et aisément praticables pour évaluer la compréhension de la parole. Nous pensons en particulier à l'évaluation des performances de locuteurs qui ne sont pas en mesure de satisfaire les contraintes de prononciation de la langue, soit parce qu'ils n'en ont pas encore acquis les représentations ou les mécanismes de production (cas des apprenants de langue étrangère par exemple), soit parce que des déficiences pathologiques les en empêchent – que celles-ci se situent sur le plan neurologique (par exemple dans les cas de dysarthrie) ou bien sur le plan structurel (par exemple dans le cas de malformations de l'appareil bucco-phonatoire ou de conséquences d'opérations chirurgicales).

Dans le domaine des troubles pathologiques de production de la parole (désormais TPPP), la performance de communication des locuteurs est habituellement évaluée grâce à des tests d'intelligibilité (Özsancak 2001) ; ces méthodes, au départ, ont été élaborées pour juger un autre aspect de la performance verbale : l'articulation des phonèmes. Or l'étude de Hustad (2008) a montré que les résultats obtenus par ces tests ne sont pas nécessairement représentatifs de la performance communicative des locuteurs car, en particulier, ils n'entretiennent pas de relation significative avec la compréhension des auditeurs. Pour évaluer la performance communicative de locuteurs pathologiques, il est donc primordial de développer des méthodes spécialement conçues pour évaluer cette dimension particulière qu'est la compréhension de la parole en situation de communication.

Dans cet article nous proposons un protocole qui s'inscrit dans cette perspective. Plutôt que d'évaluer la compréhension de la parole *via* des tests basés sur la mesure du transfert de code linguistique, nous avons choisi de nous inspirer de tests pragmatiques dédiés à l'évaluation de la compréhension de la parole *chez*

---

\* <http://petra.univ-tlse2.fr/elokanz>

*l'auditeur*, notamment chez des sujets dont les compétences de compréhension orale sont lésées ou bien sont encore en cours d'acquisition. Notre protocole place le locuteur que l'on désire évaluer et l'auditeur dans un jeu de communication où le premier est instructeur et donne des commandes verbales au second. La compréhension de l'auditeur est évaluée à travers sa capacité à agir en adéquation avec les instructions reçues. Nous avons implémenté ce protocole sous la forme d'un logiciel dénommé « EloKanz ».

## 2. Cadre théorique

La question de l'évaluation de la compréhension de la parole est extrêmement complexe, car elle renvoie à des processus mentaux non observables directement. Depuis les années 1950, de nombreuses études se sont attachées à mettre au jour les représentations et les mécanismes impliqués dans la compréhension de la parole, depuis la perception du signal de parole jusqu'à l'interprétation des messages dans un contexte de communication particulier. Typiquement, ces recherches ont été menées à l'aide de matériels linguistiques très contrôlés, et en employant des tests fondés sur l'observation des temps de réaction des auditeurs dans des tâches d'ordre métalinguistique (Tanenhaus et Brown-Schmidt 2008). En dépit de ces efforts de recherche qui ont donné lieu à des courants de recherche entiers, de nombreuses inconnues subsistent encore aujourd'hui. C'est par exemple le cas concernant la façon dont l'être humain apparie les signaux de parole aux représentations linguistiques dont il dispose en mémoire (voir par exemple le débat entre les modèles à représentations phonologiques abstraites et les modèles à exemplaires – *cf.* Nguyen *et al.*, 2009), ou encore concernant la manière dont nous prenons en compte les informations liées au contexte de communication lors de la compréhension d'un message parlé (voir par exemple le débat entre les modèles d'interprétation en bloc et fragmentée – *cf.* Lotze *et al.* 2011).

Le manque de connaissances actuel sur les représentations et les mécanismes engagés dans les processus de compréhension de la parole fait qu'il paraît difficile d'inférer, à partir de l'observation du traitement cognitif d'éléments linguistiques particuliers par un auditeur (par ex. les phonèmes entendus), dans quelle mesure ce dernier sera en mesure de comprendre des messages pris dans leur ensemble et dans un contexte particulier. Ceci pourrait expliquer que, dans le cas de l'évaluation de la compréhensibilité de locuteurs pathologiques, les tests se concentrant sur les aspects de perception phonémique — comme les tâches de retranscription orthographique — ne permettent pas de prédire l'aptitude d'auditeurs à comprendre le sens des propos prononcés par les patients (Hustad 2008).

Pour dépasser cette question des représentations et des processus cognitifs impliqués dans la compréhension de la parole, une solution alternative est offerte par les tests pragmatiques, reposant sur l'observation des réactions comportementales des auditeurs aux énoncés entendus.

### *3. Les tests pragmatiques d'évaluation de la compréhension de la parole*

Nous qualifions de *pragmatiques* les tests de compréhension de la parole qui impliquent de *véritables* interactions de la part du locuteur et de l'auditeur. La particularité de ces tests est de juger de la compréhension de l'auditeur en vérifiant l'adéquation de ses actions avec les énoncés qu'il a entendus. Alors que les tests d'intelligibilité de la parole impliquent chez les sujets locuteurs et auditeurs des tâches de lecture et de retranscription, les tests pragmatiques consistent à comparer les intentions de communication du locuteur et les réactions de l'auditeur.

Ainsi, en aphasiologie, des batteries de test comme le MT-86 (Nespoulous *et al.* 1992) proposent d'évaluer la compréhension orale de patients aphasiques en leur donnant des ordres verbaux tels que « fermez les yeux » ou « mettez la fourchette dans le verre ». Si le patient effectue l'action correspondante, alors sa compréhension de l'énoncé est jugée correcte. Dans ce type de test la compréhension de la parole est donc évaluée à travers une interaction verbale directe entre les interlocuteurs, c'est-à-dire sans passer par des tâches métalinguistiques comme la lecture ou la retranscription d'énoncés. Ceci explique que les tests pragmatiques de compréhension de la parole ont été utilisés pour investiguer les performances de sujets qui ne disposent pas de ces compétences de lecture ou d'écriture, comme les enfants en bas âge ou les animaux (cf. par exemple les études menées par Herman *et al.* 1984 ou Savage-Rumbaugh *et al.* 1993).

Notons que, contrairement à la problématique qui nous occupe, les tests pragmatiques de compréhension de la parole ne sont généralement pas utilisés pour évaluer la performance communicative de locuteurs, mais plutôt pour juger directement de la compétence de compréhension orale d'auditeurs sains ou pathologiques. Néanmoins, dans le cadre de l'évaluation des performances de locuteurs – comme par exemple dans le cas des TPPP, ce paradigme de test peut se révéler particulièrement pertinent, en apportant un éclairage différent de celui proposé par les tests d'intelligibilité de la parole. C'est pourquoi nous avons adopté une approche pragmatique pour l'élaboration du protocole que nous proposons.

## *4. EloKanx : test informatisé pour l'évaluation pragmatique de la compréhension de la parole*

### *4.1. Description du protocole*

Notre protocole implique une communication verbale entre un sujet locuteur et un sujet auditeur. Cette communication se déroule en différé : dans un premier temps les paroles du locuteur sont enregistrées, puis dans un second temps les enregistrements sont diffusés à l'auditeur. L'enregistrement audio des énoncés permet de tester la fiabilité des résultats auprès de différents auditeurs, et également de comparer les scores de compréhension obtenus avec d'autres mesures portant sur les mêmes stimuli.

Le locuteur comme l'auditeur sont placés devant un même jeu d'images représentant divers objets. Le locuteur a pour instruction de donner des consignes orales afin de faire exécuter des actions de déplacement à l'auditeur. Tout comme dans les tests pragmatiques, la compréhension de l'auditeur est considérée correcte si ce dernier effectue les actions correspondant aux consignes. Dit autrement, si le locuteur demande à l'auditeur de positionner une banane à droite d'une souris, alors sa compréhension est jugée correcte s'il a sélectionné l'image correspondant à la banane et qu'il l'a positionnée à droite de l'image représentant une souris.

La mise en œuvre de ce protocole a nécessité, en amont, de mesurer l'accord entre sujets sur la dénomination des images. Le protocole a ensuite été implémenté dans un logiciel qui permet à la fois d'enregistrer le locuteur, de tester la compréhension de l'auditeur et de récupérer les données correspondantes.

#### *4.2. Pré-test : accord sur le nom des images*

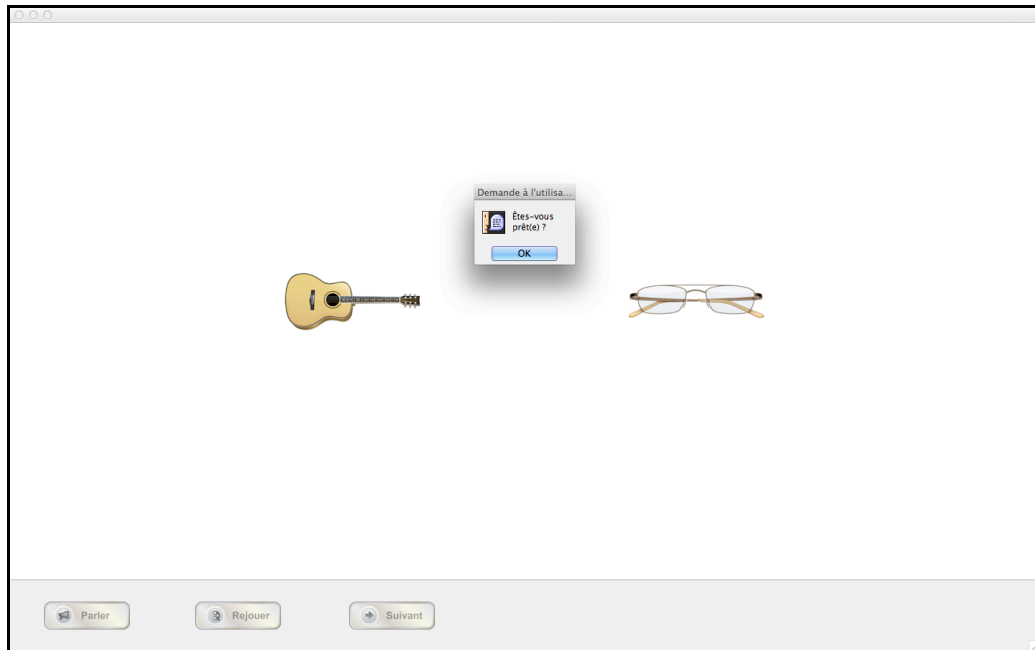
Notre protocole prévoit, pour le locuteur et l'auditeur respectivement, de produire et de résoudre des références langagières concernant des objets représentés par des images. Puisque l'objet du test est de rendre compte de la compréhensibilité du locuteur, il est nécessaire de contrôler les obstacles à la compréhension des énoncés qui ne relèvent pas de la performance du locuteur. Un tel obstacle pourrait être un désaccord entre le locuteur et l'auditeur à propos des concepts représentés par les images, et/ou de la dénomination à attribuer à celles-ci.

Pour traiter cette question nous avons étudié l'accord sur le nom attribué aux différentes images en testant leur dénomination auprès de 68 sujets adultes francophones, dont 30 patients souffrant de TPPP d'origines diverses (Fontan 2012). L'accord sur le nom a été mesuré en amenant les sujets à dénommer les images dans le cadre même du protocole que nous proposons ici, c'est-à-dire en donnant des commandes orales de déplacement. Cette étude a permis de constituer un jeu d'images pour lesquelles un accord sur le nom de 100% a été obtenu.

#### *4.3. Implémentation logicielle*

Nous avons implémenté notre protocole dans un logiciel dénommé EloKanz. Il est prévu que ce logiciel soit manipulé aussi bien par les sujets locuteurs et auditeurs que par l'expérimentateur. Ce dernier a en effet la possibilité, depuis une interface graphique, de choisir les paramètres correspondant aux sessions d'enregistrement ou aux tests de compréhension.

Ainsi, lors de l'enregistrement d'un locuteur, l'expérimentateur commence par choisir le nombre de phrases qu'il désire lui faire enregistrer, ainsi que le nombre et la taille des images à présenter à l'écran. Il lance ensuite la session d'enregistrement, et laisse le locuteur manipuler le logiciel *via* l'interface présentée en figure 1.



**Figure 1 : interface pour le locuteur**

Pour chaque phrase à enregistrer, le logiciel diffuse une animation : une image représentant une main vient saisir une image et la déplace à gauche, à droite, en dessous ou au-dessus d'une autre image (figure 2).



**Figure 2 : animation de déplacement**

Des boutons en bas de l'interface graphique permettent au locuteur d'enregistrer un énoncé verbal, *via* le microphone du système sur lequel le logiciel est installé. L'idée est que le locuteur enregistre une phrase destinée à faire exécuter l'animation qu'il a visionnée à son interlocuteur.

Le logiciel EloKanz permet ensuite d'évaluer la compréhension, par un auditeur, des énoncés enregistrés par le locuteur. Pour ce faire l'expérimentateur sélectionne les phrases à diffuser à l'auditeur, et choisit le nombre et la taille des images présentées à l'écran. L'expérimentateur a également la possibilité de sélectionner des phrases qui ont été prononcées par plusieurs locuteurs ; dans ce cas il peut

choisir entre un ordre de présentation linéaire (les phrases prononcées par chacun des locuteurs sont diffusées consécutivement) ou bien pseudo-aléatoire (pas plus de deux phrases prononcées par le même locuteur ne peuvent être diffusées consécutivement).

Pour répondre à chacune des consignes orales qui lui sont diffusées, l'auditeur dispose d'une interface similaire à celle du locuteur, et qui lui permet de déplacer les images présentées à l'écran. Des commandes lui permettent éventuellement d'annuler ou de recommencer un déplacement, puis de valider son action pour passer à la consigne suivante.

À la fin de la passation de l'auditeur, le logiciel EloKanz compare les animations qui ont été jouées au locuteur et les actions effectuées par l'auditeur. Les données résultant de ce traitement apparaissent directement dans l'interface graphique prévue pour l'expérimentateur (figure 3).

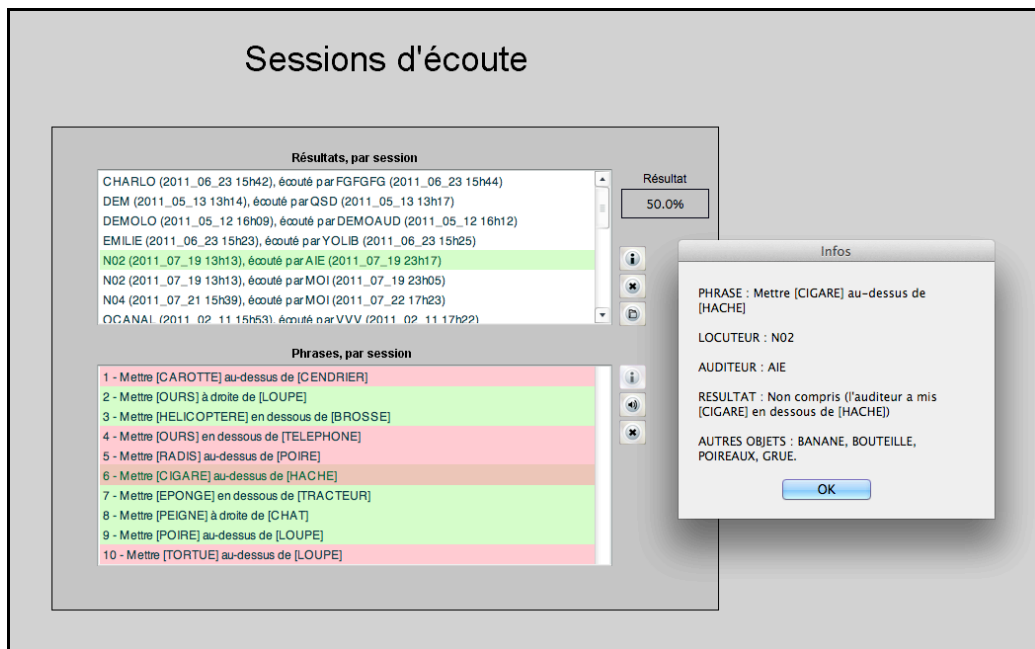


Figure 3 : fenêtre de récupération des données

Dans cette interface graphique apparaît la liste de phrases qui ont été diffusées à l'auditeur, avec le pourcentage de phrases qui ont été comprises par celui-ci (cadre « Résultat » sur la figure 3). Les phrases qui ont été comprises apparaissent en vert et celles pour lesquelles un élément au moins n'a pas été compris apparaissent en rouge. Il est également possible de vérifier les informations relatives à une phrase particulière (fenêtre « Infos » sur la figure 3). Ces données sont également enregistrées dans un fichier texte, qui contient d'autres informations sur la compréhension de chaque phrase par l'auditeur (figure 4).

OBJ1	POS	OBJ2	OBJ1 DONE	POS DONE	OBJ2 DONE	OBJ1 OK	POS OK	OBJ2 OK	RESULTAT	DELAI FIRST ACTION	DELAI VALIDER	NB OBJETS	NB ACTIONS	CONTEXTE
Lit	droite	Scie	Lit	droite	Scie	1	1	1	1	7192	94095	6	1	#Cigarette#Tournevis#Bouteille#Tondeuse#
Canne	haut	Brosse	Canne	bas	Brosse	1	0	1	0	7793	12469	6	1	#Lunettes#Telephone#Poire#Cheval#
Boxeur	bas	Entonnoir	Boxeur	bas	Entonnoir	1	1	1	1	7966	14968	6	2	#Fusée#Tracteur#Bonnet#Fleur#
Avocat	gauche	Avion	Avocat	gauche	Avion	1	1	1	1	7120	23804	6	2	#Zebre#Poisson#Champignon#Moulin#
Tournevis	gauche	Boxeur	Tournevis	gauche	Boxeur	1	1	1	1	8846	11872	6	1	#Plume#Cigare#Telephone#Poire#
Marteau	bas	Pipe	Marteau	bas	Pipe	1	1	1	1	5671	10412	6	1	#Poussin#Rateau#Accordeon#Poisson#
Boutons	haut	Boxeur	Boutons	haut	Boxeur	1	1	1	1	7672	10642	6	1	#Skieur#Elephant#Fleur#Fraises#
Fleur	bas	Tente	Fleur	bas	Tente	1	1	1	1	3778	6682	6	1	#Autruche#Poussin#Pipe#Cuillere#
Oignon	gauche	Fourchette	Oignon	droite	Brouette	1	0	0	0	6958	15256	6	2	#Croissant#Quille#Tracteur#Pipe#
Entonnoir	bas	Cuillere	Entonnoir	bas	Cuillere	1	1	1	1	4772	6842	6	1	#Coccinelle#Cerises#Bouteille#Gomme#

Figure 4 : fichier contenant les données relatives à la compréhension de chaque phrase

Ce fichier texte contient les informations relatives à chaque consigne verbale : <Objet 1>, <Position> et <Objet 2> (par ex. *lit*, *droite* et *scie* pour la consigne « Mettez le lit à droite de la scie »). Sont ensuite données les informations relatives au déplacement effectué par l'auditeur, et la correction (0 ou 1) de chacun de ces éléments.

Outre les informations relatives au contexte imagé accompagnant chaque consigne orale (nombre et identifiants des images présentées à l'écran), EloKanz fournit d'autres variables relatives à la compréhension de l'auditeur, comme le nombre d'actions de déplacement qu'il a effectuées ou ses temps de réaction avant la sélection d'un objet et la validation de son déplacement. Les temps de réaction sont en effet souvent interprétés en psycholinguistique comme des indicateurs des difficultés de traitement de la parole (voir par exemple le *Sentence-By-Sentence Listening Task* élaboré par Ralston *et al.*, cf. notamment Ralston *et al.* 1991). Dans une étude réalisée en milieu clinique auprès de 30 patients souffrant de TPPP, nous avons par ailleurs observé une corrélation très forte et hautement significative entre les temps de réaction des auditeurs et les jugements de compréhensibilité effectués par des médecins habitués à ce type de parole (Bordenave et Ros 2012 ; Fontan 2012).

EloKanz donne aussi la possibilité de pondérer les résultats obtenus (scores de compréhension et temps de réaction des auditeurs) en tenant compte de facteurs liés à aux énoncés linguistiques et aux contextes imagés les accompagnant, comme :

- la proximité phonologique entre les noms des images cibles et les noms correspondant aux autres images présentées à l'écran (distracteurs). Cette valeur est calculée par l'algorithme de Levenshtein, qui permet de mesurer la distance entre deux chaînes de symboles (cf. Beijering *et al.* 2008 pour une application à la distance phonologique entre deux signifiants linguistiques). Nous avons modifié cet algorithme afin qu'il prenne également en compte la distance séparant deux phonèmes, calculée en fonction du nombre de traits acoustiques qu'ils partagent (Fontan 2012) ;
- le nombre d'images du contexte dont le nom appelle un déterminant similaire à celui des images cibles. Plus ce nombre est réduit, plus l'auditeur peut utiliser une stratégie consistant à se baser sur les déterminants entendus et à procéder par élimination pour déduire les noms qui ont été prononcés par le locuteur ;
- la fréquence et la longueur des noms correspondant aux images, telles qu'indiquées dans la base *Lexique 3*, constituée à partir d'un corpus de sous-titres de films (New *et al.* 2007).

## 5. Conclusion

La compréhensibilité de la parole met en jeu des processus cognitifs difficilement observables et donc difficilement quantifiables. Dans cet article nous avons proposé un protocole novateur permettant de juger de la compréhension d'énoncés produits par un locuteur. En nous inspirant des tests pragmatiques conduits en aphasiologie et en psychologie, nous avons proposé d'évaluer la compréhensibilité d'un locuteur à travers sa capacité à commander les actions d'un auditeur.

Notre travail a conduit à l'élaboration d'un nouveau logiciel de compréhensibilité de la parole dénommé « EloKanz ». Ce logiciel propose un environnement ergonomique pour enregistrer des énoncés prononcés par un locuteur et tester leur compréhensibilité auprès d'un ou de plusieurs auditeur(s).

EloKanz a déjà permis d'étudier la relation entre les mesures d'intelligibilité de la parole, portant sur le transfert de code linguistique, et des tests pragmatiques, davantage tournés vers le succès de la communication verbale (Fontan 2012). Ces études ont confirmé l'intérêt que représentent les tests pragmatiques, notamment pour ce qui concerne l'évaluation clinique de patients souffrant de troubles pathologiques de production de la parole (Hustad 2008). Dans cette optique le logiciel EloKanz peut constituer une solution pertinente, d'autant plus qu'une récente étude de faisabilité clinique a mis au jour des résultats très encourageants pour ce qui est de l'utilisation du logiciel par des patients et par des personnels médicaux (Bordenave et Ros 2012).

## Références bibliographiques

BEIJERING, K., GOOSKENS, C., HEERINGA, W., Modeling intelligibility and perceived linguistic distances by means of the Levenshtein algorithm, in VAN KOPPEN, M., BOTMA, B., *Linguistics in the Netherlands 2008*, Amsterdam, John Benjamins, 2008.

BORDENAVE, J., ROS, A., *Études de faisabilité clinique et de validation d'un test de compréhensibilité de la parole pathologique*, Mémoire pour l'obtention du certificat de capacité d'orthophonie, Toulouse, Faculté de médecine Toulouse-Rangueil, 2012.

FONTAN, L., De la mesure de l'intelligibilité à l'évaluation de la compréhension de la parole pathologique en situation de communication, Thèse de doctorat en Sciences du Langage, Toulouse, Université de Toulouse, 2012.

HERMAN, L. M., RICHARDS, D. G., WOLZ, J. P., Comprehension of sentences by bottlenosed dolphins, *Cognition*, 1984, **16**, 129-219.

HUSTAD, K. C., The Relationship Between Listener Comprehension and Intelligibility Scores for Speakers With Dysarthria, *Journal of Speech, Language and Hearing Research*, 2008, **51**, 562-573.

LOTZE, N., TUNE, S., SCHLESEWSKY, M., BORNKESSEL-SCHLESEWSKY, I., Meaningful physical changes mediate lexical-semantic integration: Top-down and form-based bottom-up information sources interact in the N400, *Neuropsychologia*, 2011, **49**, 3573-3582.



NESPOULOUS, J. L., LECOURS, A. R., LAFOND, D., LEMAY, M. A., PUEL, M., JOANETTE, Y., *et al.*, *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie : MT-86 module standard initial M1b*, Isbergues, Ortho Edition, 1992.

NEW, B., BRYBAERT, M., VERONIS, J., PALLIER, C., The use of film subtitles to estimate word frequencies, *Applied Psycholinguistics*, 2007, **28**, 661-677.

NGUYEN, N., WAUQUIER, S., TULLER, B., The dynamical approach to speech perception: From fine phonetic detail to abstract phonological categories, in PELLEGRINO, F., MARSICO, E. CHITORAN, I., COUPÉ, C., *Approaches to Phonological Complexity*, Berlin, Mouton de Gruyter, 2009.

ÖZSANCAK, C., L'intelligibilité, in AUZOU, P., ÖZSANCAK, C., BRUN, V., *Les dysarthries*, Paris, Masson, 2001.

RALSTON, J., PISONI, D., LIVELY, S., GREENE, B., MULLENIX, J., Comprehension of synthetic speech produced by rule: Word monitoring and sentence-by-sentence listening times, *Human Factors*, 1991, **33**, 471-491.

SAVAGE-RUMBAUGH, E. S., MURPHY, J., SEVCIK, R. A., BRAKKE, K. E., WILLIAMS, S. L., RUMBAUGH, D. M., Language Comprehension in Ape and Child, *Monographs of the Society for Research in Child Development*, 1993, **58**, 1-221.

TANENHAUS, M. K., BROWN-SCHMIDT, S., Language processing in the natural world, in MOORE, B. C. M., TYLER, L. K., MARSLER-WILSON, W. D., *The perception of speech: From sound to meaning*, Oxford, Oxford University Press, 2008.