Carcinologic Speech Severity Index Project: A Database of Speech Disorder Productions to Assess Quality of Life Related to Speech After Cancer

Corine Astésano[1], Mathieu Balaguer[2], Jérôme Farinas[1], Corinne Fredouille[3], Pascal Gaillard[1], Alain Ghio[4], Laurence Giusti[4], Imed Laaridh[3], Muriel Lalain[4], Benoît Lepage[1], Julie Mauclair[1,5], Olivier Nocaudie[1], Julien Pinquier[1], Oriol Pont[1], Gilles Pouchoulin[4], Michèle Puech[2], Danièle Robert[4], Etienne Sicard[1], Virginie Woisard[2]

[1]Toulouse University, [2]Toulouse Hospital, [3]University of Avignon, LIA,[4]Aix-Marseille Univ, CNRS, LPL, [5]Paris Descartes University

Toulouse France, Toulouse France, Avignon France, Aix-en-Provence France, Paris France

c2si@irit.fr

Abstract

Within the framework of the Carcinologic Speech Severity Index (C2SI) InCA Project, we collected a large database of French speech recordings aiming at validating Disorder Severity Indexes. Such a database will be useful for measuring the impact of oral and pharyngeal cavity cancer on speech production. That will permit to assess patients' Quality of Life after treatment. The database is composed of audio recordings from 135 speakers and associated metadata. Several intelligibility and comprehensibility levels of speech functions have been evaluated. Acoustics and Prosody have been assessed. Perceptual evaluation rates from both naive and expert juries are being produced. Automatic analyzes are being carried out. That will provide to speech therapists objective tools to take into account the intelligibility and comprehensibility of patients which received cancer treatment (surgery and/or radiotherapy and/or chemotherapy). The aim of this paper is to justify the need of this corpus and his data collection. This corpus will be available to the scientific community through the GIS Parolotheque.

Keywords: speech intelligibility and comprehensibility, quality of life assessment, speech corpus, pathological speech

1. Introduction

The decreasing mortality in cancerology highlights the importance to reduce the impact on the Quality of Life (QoL) after cancer. That particularly concerns Head and Neck Cancers (HNC), because their treatment can be mutilating and disabling.

However, the usual tools for assessing QoL are not relevant for measuring the impact of the treatment on the main functions involved by the sequelae. And, there is a clear lack of uniform methods for assessing functional outcomes.

Measuring the impact on one or several of the most altered functions after the therapeutic care of a given tumoral localization, would allow:

1. to complete the expression of the therapeutic outcomes by functional forecast index,
2. to adjust the treatment for reducing their functional consequences.

For the HNC, it is mainly about impacts of (oral) communication and feeding (swallowing) (Mlynarek AM et al 2008). QoL research has, at times, failed to provide health care professionals with clinically relevant and interpretable information that can guide treatment decisions. This has led researchers to attempt to make commonly used research tools more accessible to the clinicians. Health-related Quality of Life (HRQoL) reflecting the effect of disease and disease treatment on general well being (Cardol M et al 1999) evolved to the creation of handicap questionnaires and specific related QoL questionnaires for numerous chronic diseases with a rise in importance of specific and symptom modules. But validated tools to measure the functional outcomes of carcinologic treatment are still missing, in particular for speech disorders. Some assessments are available for voice disorders in laryngeal cancer but they are based on very poor tools for oral and pharyngeal cancers involving more the articulation of speech than voice. Because the usual tools to assess QoL are not relevant to measure the impact of the treatment on the main functions involved by the sequelae, and because it is acknowledged that an unbiased and objective assessment of the communication deficiency caused by a speech disorder calls for automatic speech processing tool, we proposed to develop a severity index of speech disorders describing the outcomes of therapeutic protocols completing the survival rates. The principle is to perform an audio recording of the patient's speech and to compute the intelligibility of the utterances produced in the aim to get a score. Middag in 2012 presented a new method that predicts running speech intelligibility in a robust way against changes in the text and against differences in the accent of the Dutch speakers applicable to patients treated for HNC.

Therefore, our hypothesis is that an automatic assessment technique can measure the impact of the speech disorders on the communication abilities giving a severity index of speech in patients treated for HNC and particularly for oral and pharyngeal cancers. We will name this index the Carcinologic Speech Severity Index (C2SI). Speech intelligibility is the usual way to quantify the severity of neurologic speech disorders. But this measure is not valid in clinical practice because of several difficulties as the familiarity effect of this kind of speech and the poor inter-judge reproducibility. Moreover, the scores do not accurately reflect listener comprehension.

In order to develop and evaluate this C2SI, a project has been contracted with French National Cancer Institute (Grant InCA SHS n°2014-135) in 2014 and CHU Toulouse, LPL Aix-En-Provence laboratory, PETRA MSH Toulouse, Octogone-Lordat Toulouse, LIA University of Avignon, Paul Sabatier Toulouse University and IRIT Toulouse laboratory, which form an interdisciplinary team. This C2SI project aims to create a speech corpus in order to validate the assumptions of the speech severity index. The corpus is presented in this paper. The structure and the list of tasks performed by each speaker are presented in section 2. Section 3 presents the available material, and some statistics on this corpus are reported in section 4.

2. Method description

The corpus associates audio recordings and QoL questionnaires. The content was chosen to assess a broad spectrum of intelligibility linked to quality of life: from acoustic to understandability.

2.1 Self Assessment questionnaires

Self assessment questionnaires are used in practice to evaluate QoL in its several dimensions.

The main generic quality of life questionnaire is the MOS-SF36 (Wade and Sherbourne 1992). It is validated in all kind of illnesses and explores physical and mental health. In the case of Head and Neck Cancer, the cancer specific HRQoL questionnaire used frequently in Europe is the European Organization for Research into the Treatment of Cancer (EORTC QLQ-C30) with its complementary module assessing HNC specific problems and symptoms respectively QLQ-H&N35. These questionnaires, as generic or specific, give independent information.

Self-questionnaire of Handicap was proposed for various functions of the upper aerodigestive tract (UADT). The Speech Handicap Index (SHI) for speech (Rinkel 2008) is validated for HNC. The Phonation Handicap Index (PHI) is a French similar tool but validated for all kind of speech production disorders (Fichaux-Bourin et al 2009).

The relations between the questionnaires of QoL and the questionnaires of handicap were often analyzed, the quality of life being of use to the validation of contents of the questionnaires of handicap. Strong correlations (0.7 to 0.9) were computed between the SHI and the speech domain of QoL questionnaire. This correlation is lower, if not absent, according to the other domains (Borggreven et al 2007, Dwivedi et al 2011, Thomas et al 2009).

Because the use of a handicap questionnaire targeting a function is well correlated to the domains of the questionnaires of QoL in relation with, we selected the generic QoL questionnaire (SF36) and the specific speech related Handicap questionnaire (SHI and PHI) to integrate the communication dimension.

2.2 Intelligibility assessment

To cover the several aspects of intelligibility, different tasks were performed by the speakers (both controls and patients):

Sustained vowel /a/ (AAA): A sustained vowel gives information about the voice level, phonation time, stability, harmonics contents, noise, unvoiced segments, etc. Despite the weak correlation between voice production and intelligibility of speech, the capacity to hold a vowel more than 5 seconds is a minimal condition for speech production. Recording this production is a global measure of the acoustic/aerodynamic balance for speech and may contribute to acoustics analysis.

Acoustico Phonetic Decoding (DAP): The limitations of intelligibility tests performed on speakers with speech production disorders lie in the ability of listeners to restore distorted sequences. This effect is emphasized when the auditors have a strong knowledge of the words used in the test and if these words are unambiguous and therefore strongly predictable (Enderby, 1983, 2008). This is generally the case for speech therapists who can make such an extensive use of these lists that they eventually know them by heart. The bias associated with this knowledge and therefore with the strong influence of the top-down perceptual mechanisms results in an overvalued intelligibility score because the phonemic restoration of the listener makes opaque the distortions of production (Warren et al., 1970; Samuel, 1981)

The solution we have adopted consists in using pseudo-words, complying with the frequent phonotactic structures in French, in large quantities so as to completely neutralize the effects of lexicality, familiarization and learning of the items by the listeners (Ghio et al, 2016).

Image Description (DES) and Spontaneous Speech (SPO): In real life, the top-down effect is present. This is why the spontaneous speech remains often used for assessing intelligibility (Woisard et al., 2013). In order to reduce the predictability of the speech produced, we recorded patients/controls describing a picture and telling their comments about a text which they read before.

Reading a Short Text (LEC): Using the same text is in complete opposition with the previous tasks but is interesting for the comparison of acoustic analysis and the automatic intelligibility scoring. This makes it possible to produce automatic phonetic alignments, even if the speech production is very altered. Speech rate, prosody, consonant and vowel precision, pauses and other speech features may be easily extracted and compared between the normal and patient groups.

2.3 Prosody assessment

Prosody helps structuring different levels of linguistic information, be it lexical, syntactic, semantic or pragmatic. The patients we focus on in this project have undergone treatment at the supraglottic level of their anatomy (glossectomy, mandibulectomy for example). Hence, the source was not affected. However, we hypothesize that compensatory mechanisms at the segmental level will impact prosodic characteristics, particularly affecting prosodic fluency. Our prosodic tasks are designed to evaluate which structural functions of prosody are most affected by these types of cancer.

Modal Prosody Function (MOD): Classically, clinical investigations of speech pathologies involve assessing the modal and emotional prosody functions, although speech intelligibility/comprehensibility is more related to structural functions, which are never tested.

Focus (FOC) and Disambiguation Syntactic (SYN) Prosody: these tasks are taken from (Aura, 2012), who adapted (Magne et al, 2005) and (Astésano et al, 2007) for clinical use. The modality task consists in producing ten identical sentences with 3 different modalities: assertion, question and injunction (*Tu manges les pâtes* ?/ . /! *eg. You eat pastas* ?/ ./!). In the focus task (Aura, 2012), speakers had to resolve a paradigmatic opposition (contrastive focus) between two words given in an auditorily presented sentence so as to prosodically highlight the relevant word ("*Tu as vu un canard ou un cochon dans le jardin?*" *eg. "You saw a duck or a pig in the garden?"* with the written answer: "*j'ai vu un CANARD dans le jardin*" *eg. "I saw a DUCK in the garden"*). The syntactic task (Aura, 2012; Astésano *et al*, 2007) consists of similar written scripts that only prosody

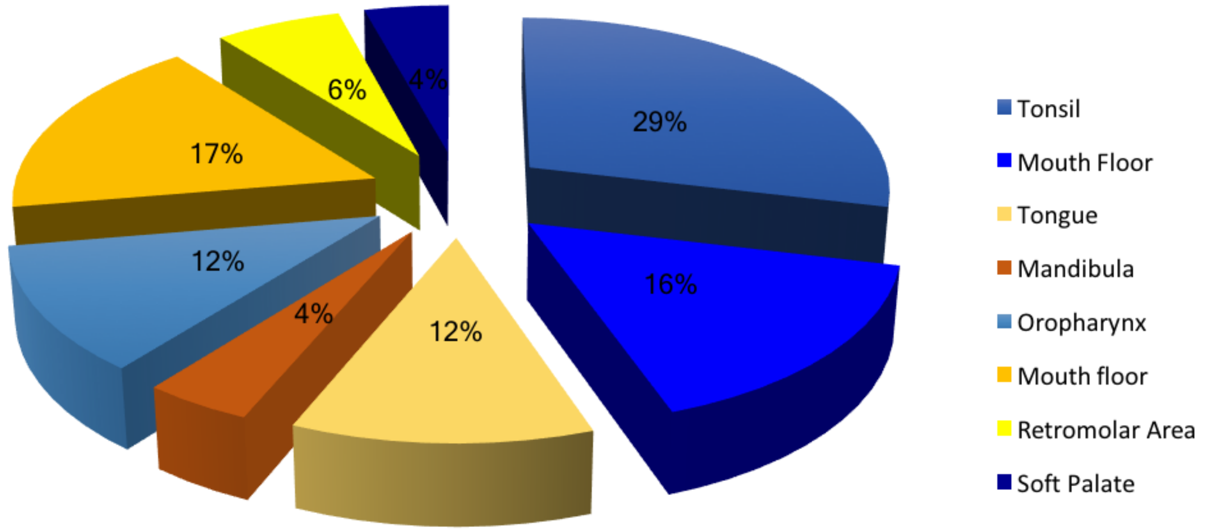

Figure 1: Tumor localisation distribution.

can disambiguate. For example, in the sentence "*les chevaux et les poneys blancs*"(eg. *"White horses and poneys"* but note that the adjective in French is at the end of the sentence), the adjective "*blancs" (eg. white)* can either apply to the second noun only (narrow scope) or to the two nouns (broad scope): prosodic cues such as final lengthening, pause and *f0* excursions can give the proper syntactic parsing (either *les chevaux// et les poneys blancs* or *les chevaux et les poneys // blancs)*.
The capacity of the speakers to properly use prosodic cues in these different tasks is then intended to be evaluated through perception tests on naive, healthy listeners (Nocaudie et al., 2017)

2.4 Comprehensibility assessment

In order to evaluate the comprehension of speech, it is important to go beyond the simple tests on isolated words.

We introduce **Sentence Verification Tasks (SVT)** in order to assess the global comprehension of running speech. In this task, speakers read a set of sentences. The semantic content of each sentence can be true (ex: "january is a winter month") or false (ex: "january is a summer month"). In the perception evaluation, participants are presented a variety of utterances across several knowledge domains and have to decide as fast as possible if these statements are true or false (Pisoni et al., 1987). The accuracy score and the response time are used as a couple of indicators of the comprehension process. Indeed, when auditors need to understand the linguistic content of a message and perform an appropriate response [True or False], the quality of the acoustic-phonetic information of the speech signal plays an important role both in the speed and accuracy of the answer provided.

A simple recording of spontaneous speech can also be interesting to assess the comprehensibility of a text. But the evaluation of an index based on these recordings is not easy as the semantic sense may be very varied. But this could be analysed in order to confirm the other indexes during perceptual analyzes.

3. Corpus description

3.1 Population

We expect a correlation between an automatic index and the perceived index given by the jury to be as high as 0.86 correlation that was achieved in University of Ghent work (Middag et al, 2008, Middag et al, 2009 and Middag, 2012). The size of the sample influences the precision of this estimation, a bigger sample bringing a bigger precision (characterized by a narrower reliable interval). To obtain a reliable interval in 95% the width of which is not superior to 0.15 around a coefficient of 0.8, it is necessary to recruit 94 patients. In september 2017, we have recorded 94 patients and 41 control speakers. That is superior to the corpus used in (Middag, 2012), which contained recordings and perceptual evaluations of 55 patients with advanced Head and Neck Cancer who were treated with concomitant chemoradiotherapy.

The 94 patients are recruited in the three main departments of Toulouse managing patients with HNC (ENT department of the Universitary Hospital, Cancerology department of the Institut Claudius Regaud (surgery and radiotherapy), Maxillofacial surgery department of the Universitary Hospital of Toulouse).

They are selected from the lists of carcinologic follow-up consultations of these 3 departments. These departments are participating within the University Institute of cancer in Toulouse (IUC-T) and will be associated with the unit of Onco-réhabilitation which is located at the IUC-T Oncopole.

3.2 Questionnaires

The SHI and PHI questionnaires presented in 2.1 are given to the patients just before the audio recordings.

3.3 Recordings

The speakers were settled in a comfortable way in an anechoic room in front of a computer. This computer was used to visually display instructions and corpus. For some tasks, the instructions were also produced with an auditory modality (ex: pseudo-words in DAP task). The recordings were made with a Neumann TLM 102 Cardioid Condenser Microphone connected to a FOSTEX digital recorder. The sampling rate was 48 kHz, which facilitates the downsampling to 16 kHz, usually used in automatic speech processing.

The corpus is composed of subpart collections described below. The passation order is: AAA, LEC, DES, SPO. And then the prosodic tasks (MOD, FOC, SYN) or the intelligibility and comprehension ones (DAP, SVT).

3.3.1 AAA

This recording consists in the production of sustained /a/ held at 3 occasions. A lot of analyses are done by speech therapist with this kind of recordings so it was important for us to include them. Indeed, the analysis of vowel /a/ can bring important cues on stability of formants and how the person deals with the breath.

3.3.2 LEC

The reading of the 1st paragraph of “La chèvre de M. Seguin”, a tale by Alphonse Daudet, is performed by the speaker. This text has been chosen because it is long enough and it includes all the French phonemes. It is also well known and widespread in clinical phonetics in France (Ghio et al., 2012).
Here is the full plain text: “*Monsieur Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes de la même façon. Un beau matin, elles cassaient leur corde, s'en allaient dans la montagne, et là-haut le loup les mangeait. Ni les caresses de leur maître ni la peur du loup rien ne les retenait. C'était paraît-il des chèvres indépendantes voulant à tout prix le grand air et la liberté.*”

3.3.3 DES

The subject was asked to choose one among several pictures that represent the same field (sea with boats). Each subject had to describe the picture to the examiner so that the latter can redraw it just on the basis of the oral explanations.

3.3.4 SPO

The patient must give his/her opinion on the questionnaire that he/she has to fill out before the recording session. He/she must speak for at least 3 minutes. This task permits to collect spontaneous speech recordings with no constraint on the sentences.

3.3.5 MOD

Each speaker recorded 10 different scripts uttered with 3 modalities: assertion, question and injunction. Each script was presented on a computer screen, with the expected prosodic modality indicated by either of the 3 punctuation marks (‘.’ ; ‘ ?’ ; ‘ !’).

3.3.6 FOC

Each speaker recorded the same set of 20 sentences, for which they had to produce the proper focus on the scripted sentence, following the audio presentation of a question. For example, after listening to the question ‘*What did you see in the garden? a duck or a pig?*’, they had to read the following sentence *‘I saw a duck in the garden’*, with contrastive focus on DUCK.

3.3.7 SYN

Each speaker recorded 13 scripts with two syntactic conditions (narrow vs. broad scope of adjective). The sentences were written on a computer screen, with the expected syntactic grouping indicated visually by vertical bars.

3.3.8 DAP

After two training trials, each speaker had to pronounce 50 pseudo-words. The pseudo-words have the phonotactic structure of the C(C)1V1C(C)2V2 type

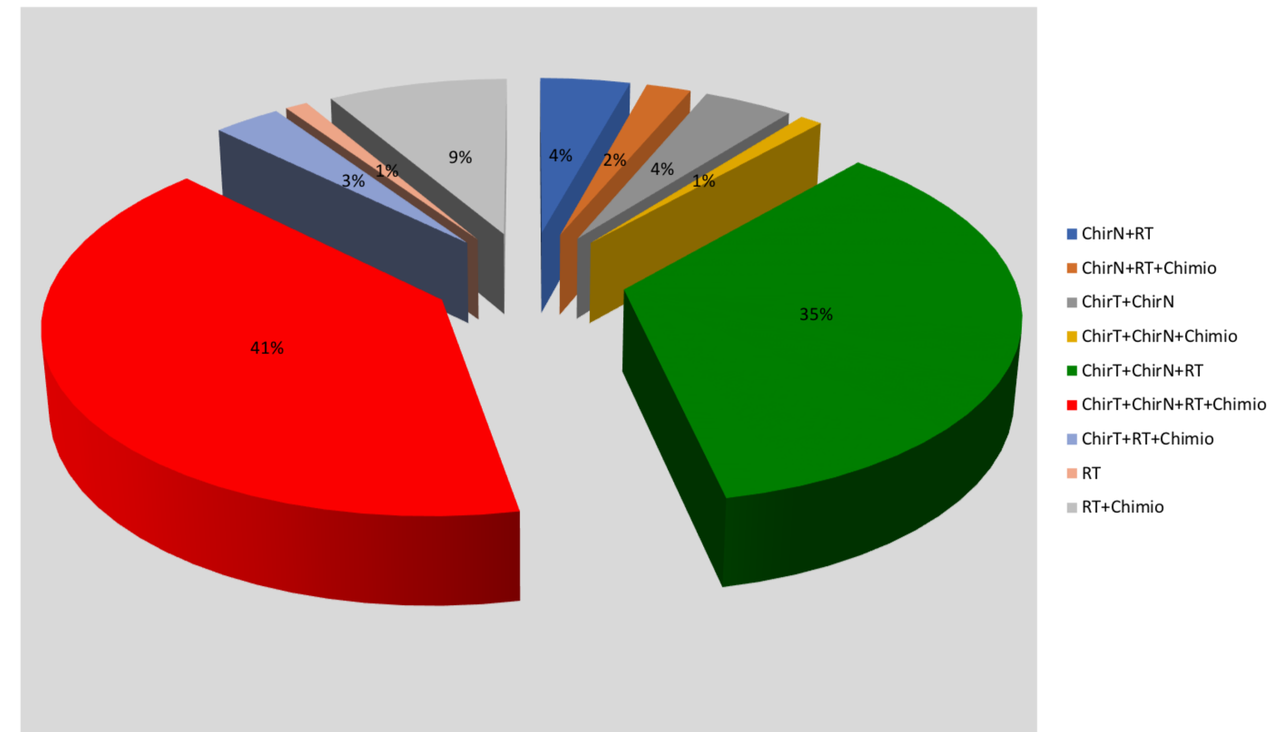

Figure 2: Patients treatment distribution.

where C(C)i is an isolated consonant or a consonant group. Such a combinatorial method makes it possible to generate about 90000 pseudo-words. Each list contains the same amount of phonemes in C1, V1, C2 and V2 position.

3.3.9 SVT

A set of 50 sentences selected from a list of 300 sentences was produced by each speaker. These sentences present a fact that can be correct or incorrect (for example: *Paris is the capital of the United Kingdom*).

4. C2SI indexes and corpus statistics

94 patients and 41 control speakers are now included in the corpus. 87 subjects and 26 controls were finally analyzed because of missing data. Among patients, 51 (59%) were men, and the mean age was 65.8 y.o. (range 36 - 87). 9 controls (35 %) were men, and the controls' mean age was different from the patient group (56.9 y.o., range 35-79, p=0.003 Mann-Whitney). The inclusion criteria were balanced regarding tumor localisation (see figure 1): 39% of oral cavity cancer (Floor of mouth, Tongue, Retromolar Area and Mandibula), and 61% of oropharyngeal cancer (Tonsil, Root of tongue, Soft Palate and when there is a larger extension "OroPharynx").

Figure 2 presents the treatment distribution of patients. The most frequent treatment related to the size of the tumors is surgery (84%). The resection of the tumor (ChirT) is associated with the node resection (ChirN) followed in 40% by a chemoradiotherapy (RT-chimio) and in 37% by only a radiotherapy (RT).

The recorded material is processed in order to produce perceptual indexes:

LEC, DES: these speech productions have been analyzed by expert therapists in order to provide an index based on the level of comprehension as follows:

- nothing is understood (not even noticed that it is sea or boats)
- only the context (sea / boats)
- identification of other elements on the picture which makes possible to differentiate it from the thematic series
- detailed descriptions are comprehensible.

Two indexes are produced from the analysis of 6 expert speech therapists leading to individual intelligibility and severity scores. These judgments are then averaged in order to produce intelligibility and a severity scores per speaker.
The distributions of perceptual intelligibility and severity scores on a scale from 0 (low intelligibility) to 10 (high intelligibility) are described in figure 3. The average for the whole population is 7 for severity and 8.3 for intelligibility.

MOD: the recordings were presented to naive listeners, who had to recognize which modality was meant, between assertion, question and injunction. Each recorded sentence was evaluated by 3 naive listeners

FOC: Each sentence previously recorded was thereafter associated with a congruous (Qu'as-tu vu dans le jardin, un cochon ou un canard ? / eg. What did you see in the garden, a pig or a duck?) or incongruous (Où as-tu vu un canard, dans le jardin ou dans la cour ? eg. Where did you see a duck in the garden or in the yard?) question. Listeners had to judge whether the perceived focus was congruous or incongruous in the manipulated dialogues. Each recorded sentence was evaluated by 3 naive listeners

SYN: each recorded sentence was presented to naive listeners who had to choose between two pictures representing either one or the other syntactic reading (narrow *vs* broad scope of adjective). Each recorded sentence was evaluated by 3 naive listeners. For tasks SYN, FOC & MOD, a perception score was calculated for each speaker, corresponding to the mean of each perceptual evaluation obtained during the test. The mean score was associated with the listeners' mean reaction time.

DAP: All of the 50 pseudo-word lists pronounced by all speakers of the database have been transcribed by 3 naive listeners. Listeners were confronted with a task that resembles acoustic-phonetic decoding followed by a written transcription. The mean distance between the transcribed and expected response is considered as a score of (un)intelligibility.

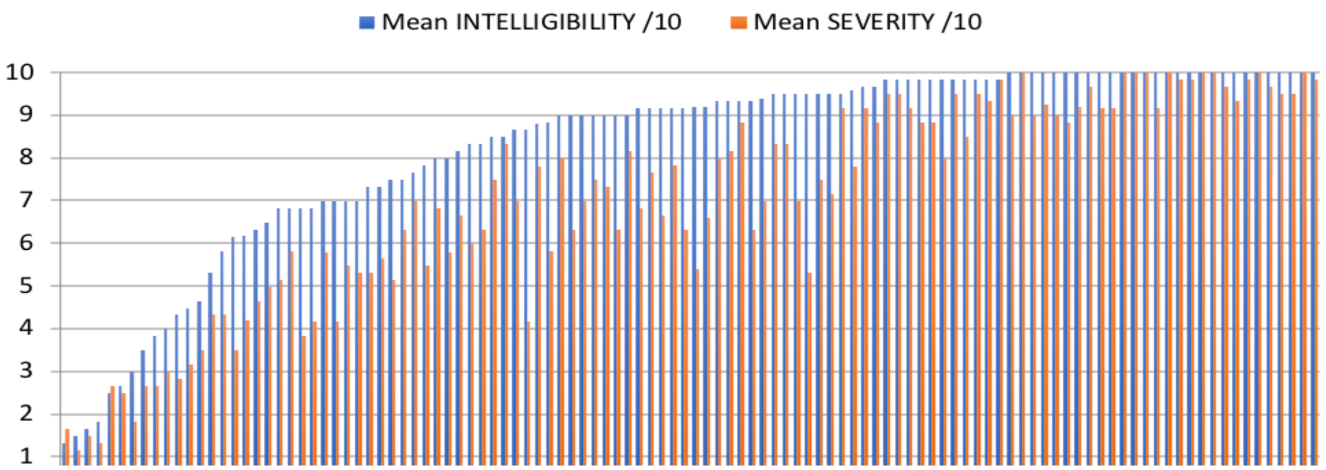

Figure 3: Distribution of intelligibility and severity scores in ordinate and subjects (by increased scores of intelligibility) in abscissa.

SVT: The sentences are evaluated by 3 naive listeners that judge if the sentence presents a correct fact or an incorrect one. This produces an indicator based of the global comprehensibility of the sentence recorded.

5. Conclusion and future work

In this paper, we have presented the design and recording of a corpus of 135 speakers, which allows us to consider the automatic production of indexes with a high level of correlation. During the constitution of the corpus, we faced several issues. Considering DAP task, patients' recordings were initially achieved, using only a visual presentation of the DAP items and the pseudo-word was simultaneously read aloud by the experimenter. But, the phonological construction of the items sometimes permitting different possible pronunciations, this configuration could have modified the speaker's repetition. To cope with this statement, we replace the aloud reading of the experimenter with a recorded synthesized voice for each item to standardize its pronunciation and to limit the potential biases. Furthermore, some tasks were considered as particularly hard to understand and to achieve by the patients (SYN, for example): the impact of these perceived difficulties will have to be checked and studied during the analysis of the results. Perceptual evaluations are in progress in order to complete the usable metadata, and to obtain reliable intelligibility/comprehensibility scores, which will be compared to self-assessed quality of life scores. We are also working now on extracting information from the different recordings in order to analyze them and to produce automatic indexes (Ghio et al. 2017, Sicard et al. 2017, Laaridh et al., 2017). This is our main goal to get objective judgments, which can help speech therapists in clinical practice. Data will be available to the scientific community by the mean of the GIS Parolotheque (https://goo.gl/4NNEZg): a scientific structure whose purpose is to facilitate access and research of pathological speech recordings (like the tumor library "thomorotheque" for access to cancer cell samples).

6. Acknowledgements

Grant n°2014-135 from Institut National pour le CAncer (INCA) in november 2014, "Sciences Humaines et Sociales, Épidémiologie et Santé Publique" call. Lead by Pr Virginie Woisard at University Hospital of Toulouse.

7. Bibliographical References

Astésano C., Bard E. G. & Turk A. (2007). Structural influences on initial accent placement in French. *Language and Speech*, *50*(3), 423-446.

Aura, K. (2012). *Protocole d'évaluation du langage fondé sur le traitement de fonctions prosodiques: étude exploratoire de deux patients atteints de gliomes de bas grade en contexte péri-opératoire* Doctoral dissertation, Université Toulouse le Mirail-Toulouse II.

Borggreven P. A., Verdonck-de Leeuw I. M., Muller M. J., Heiligers M. L., de Bree R., Aaronson N. K., Leemans C. R. (2007) *Quality of life and functional status in patients with cancer of the oral cavity and oropharynx: pretreatment values of a prospective study*. Eur Arch Otorhinolaryngol. June; 264(6):651-7

Cardol M., Brandsma J. W., de Groot I. J., van den Bos G. A., de Haan R. J., de Jong B. A. (1999) *Handicap questionnaires: what do they assess?* Disability Rehabilitation 21:97-105

Dwivedi R. C., St Rose S., Roe J. W., Chisholm E., Elmiyeh B., Nutting C. M., Clarke P. M., Kerawala C. J., Rhys-Evans P .H., Harrington K. J., Kazi R. (2011) *First report on the reliability and validity of speech handicap index in native English-speaking patients with head and neck cancer.* Head Neck. March; 33(3):341-8

Enderby P. (1983) Frenchay Dysarthria Assessment. *1st ed. San Diego: College-Hill Press*.

Enderby P., Palmer R. (2008*)* FDA-2: Frenchay Dysarthria Assessment. *2nd ed.Tex.: Pro-Ed.*

Fichaux-Bourin P., Woisard V., Grand S., Puech M., Bodin S. (2009) *Validation of a self assessment for speech disorders (Phonation Handicap Index)*. Rev Laryngol Otol Rhinol (Bord). 130(1):45-51.

Ghio A., Pouchoulin G., Teston B., Pinto S., Fredouille C., De Looze C., Robert D., Viallet F., Giovanni A. (2012), *How to manage* sound, physiological and clinical data of 2500 dysphonic and dysarthric speakers? Speech Communication, vol. 54. 2012, p. 664-679.

Ghio A., Giusti L., Blanc E., Pinto S., Lalain M., Robert D., Fredouille C. and Woisard V. (2016). *Quels tests d'intelligibilité pour évaluer les troubles de production de la parole ?* 31èmes Journées d'Etudes sur la Parole, Paris.

Ghio A., Lalain M., Giusti L., Robert D., Pouchoulin G., Rebourg M., André C., Fredouille C. and Woisard V. (2017). *Du décodage acoustico-phonétique pour mesurer l'intelligibilité de locuteurs atteints de troubles de production de la parole*, 7èmes Journées de Phonétique clinique, Paris, 29-30 june.

Laaridh I., Khader B. W., Fredouille C., Meunier C. (2017). *Automatic Prediction of Speech Evaluation Metrics for Dysarthric Speech*, Interspeech'17, Stockholm, Sweden.

Magne C., Astésano C., Lacheret-Dujour A., Morel M., Alter K., & Besson M. (2005). *On-line processing of "pop-out" words in spoken French dialogues.* Journal of cognitive neuroscience, 17(5), 740-756.

Middag C., Van Nuffelen G., Martens J. P., & De Bodt M. (2008). Objective intelligibility assessment of pathological speakers. Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech'08, 1745–1748.

Middag C., Martens J. P., Van Nuffelen G. et al. (2009). *Automated Intelligibility Assessment of Pathological Speech Using Phonological Features*. EURASIP Journal on advances in Signal Processing. 2009: 629030.

Middag, C. (2012). *Automatic analysis of pathological speech.* Doctoral dissertation, Ghent University, Department of Electronics and information systems, Ghent, Belgium.

Mlynarek A.M., Rieger J.M., Harris J.R., O'Connell D.A., Al-Qahtani K.H., Ansari K., Chau J., Seikaly H. (2008) *Methods of functional outcomes assessment following treatment of oral and oropharyngeal cancer: review of the literature.* Journal of Otolaryngol Head Neck Surg. 2008 Feb;37(1):2-10.

Pisoni D., Dedina M. (1986) "Comprehension of Digitally Encoded Natural Speech using a Sentence Verification Task: a first report" in Research on Speech Perception. Progress Report N°12, Indiana University.

Nocaudie O., Astesano C. and Woisard V. (2017). *Conservation des fonctions prosodiques post traitement des cancers de la cavité buccale et du pharynx,* 7emes Journées de Phonétique Clinique, Paris, 29-30 june.

Rinkel R. N., Leeuw I. M, van Reij E. J, Aaronson N. K, Leemans C. R. (2008) *Speech Handicap Index in patients with oral and pharyngeal cancer: better understanding of patients' complaints.* Head Neck, 30 (7) , pp. 868–874.

Samuel A. (1981) *Phonemic restoration: insights from a new methodology*. J Exp Psychol Gen.;110:474-494.

Sicard E., Mauclair J. and Woisard V. (2017). *Etude de paramètres acoustiques des voix de patients traités pour un cancer ORL dans le cadre du projet C2SI*, 7emes Journées de Phonétique Clinique, Paris, 29-30 june.

Thomas L., Jones T. M., Tandon S., Carding P., Lowe D., Rogers S. (2009) *Speech and voice outcomes in oropharyngeal cancer and evaluation of the University of Washington Quality of Life speech domain.* Clin Otolaryngol. 2009 Feb;34(1):34-42.

Wade J. E., Sherbourne C. D. *The MOS 36-item short-form health survey (SF-36)*. Medical Care (1992); 30; 473–483.

Warren R. M., Warren R. P. (1970), *Auditory illusions and confusions*. Sci. Am.; 223, 30-36.

Woisard V., Espesser R., Ghio A., Duez D. (2013). *De l'intelligibilité à la compréhensibilité de la parole, quelles mesures en pratique clinique ?* Revue de laryngologie, otologie, rhinologie, vol. 1, no. 134. 2013, p. 27-33