Automatic Detection and Recognition of Swallowing Sounds

Hajer Khlaifi[1], Atta Badii[2], Dan Istrate[1] and Jacques Demongeot[3]

[1] *University of Technology of Compiegne, UTC University, Compiegne, France*

[2] *University of Reading, Department of Computer Science, School of Mathematical, Physical and Computational Sciences, Reading, U.K.*

[3] *Laboratory AGEIS EA 7407, University Grenoble Alpes, Faculty of Medicine, Grenoble, France*

Keywords: Swallowing Sounds, Automatic Detection, Classification, Non-invasive Dysphagia Clinical Assessment Support.

Abstract: This paper proposes a non-invasive, acoustic-based method to i) automatically detect sounds through a neck-worn microphone providing a stream of acoustic input comprising of a) swallowing-related, b) speech and c) other ambient sounds (noise); ii) classify and detect swallowing-related sounds, speech or ambient noise within the acoustic stream. The above three types of acoustic signals were recorded from subjects, without any clinical symptoms of dysphagia, with a microphone attached to the neck at a pre-studied position midway between the Laryngeal Prominence and the Jugular Notch. Frequency-based analysis detection algorithms were developed to distinguish the above three types of acoustic signals with an accuracy of 86.09%. Integrated automatic detection algorithms with classification based on Gaussian Mixture Model (GMM) using the Expectation Maximisation algorithm (EM), achieved an overall validated recognition rate of 87.60% which increased to 88.87 recognition accuracy if the validated false alarm classifications were also to be included. The proposed approach thus enables the recovery from ambient signals, detection and time-stamping of the acoustic footprints of the swallowing process chain and thus further analytics to characterise the swallowing process in terms of consistency, normality and possibly risk-assessing and localising the level of any swallowing abnormality i.e. the dysphagia. As such this helps reduce the need for invasive techniques for the examination and evaluation of patient's swallowing process and enables diagnostic clinical evaluation based only on acoustic data analytics and non-invasive clinical observations.

1 INTRODUCTION

The act of swallowing or deglutition constitutes a complex process in humans. It is a critical enabler for eating and drinking; ensuring the safe transport of nourishment from the oral cavity to the stomach through the pharynx and oesophagus while keeping the epiglottis in the closed position to protect the trachea and thus the airway security. A mouthful of food, a bolus, or a gulp of water, once ingested will then go through the oral, pharyngeal and oesophageal stages of transport as phases of the swallowing process. Any impairment along the above stages can manifest as dysphagia or difficulty in swallowing leading to laboured aspiration and/or coughing to avoid chocking by attempts to expel any elements of the food that may have been accidentally ingested into the airway or misplaced or stuck somewhere along the pharyngeal to stomach pathway.. Thus. Dysphagia can be caused by a wide variety of complications that can alter the functionality of different parts of the throat beginning at the upper level, the buccal cavity, through the pharynx and then at the oesophagus. This can present as a pathology larynx, the oesophagus, and finally at the sphincter or a neuro-motor impairment causing difficulty in the orchestration of coordinative structures that support the transport of food from mouth to the stomach. The prevalence of dysphagia increases with age, from 9% for subjects aged 65 to 74, to 28% after age 85 (Singh and Hamdy, 2006). Statistics also show that 40% of people over 75 have experienced swal-lowing disorders and up to 66% of those affected are residents in a social/medical institution (Kawashima et al., 2004). Dysphagia is also common in the post-stroke population with up to 76% of this group being affected. The mortality rate arising from suffocation due to a swallowing disorder is however relatively low in the general population; for example in France it stands as 5.99 cases per 100000 in men and 6.1 cases per 100000

Khlaifi, H., Badii, A., Istrate, D. and Demongeot, J.
Automatic Detection and Recognition of Swallowing Sounds.
DOI: 10.5220/0007310802210229
In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2019)*, pages 221-229
ISBN: 978-989-758-353-7

in women according to the National Institute for Demographic Studies (INED) (INED, 2013). However, the majority of dysphagic cases often remain undiagnosed (Kawashima et al., 2004) and are therefore not treated thus leading to otherwise preventable emergencies and even deaths caused by some form of undiagnosed dysphagia. In any event once this disease is diagnosed then a treatment involving physiotherapy would be helpful but personal daily care by trained physiotherapists is impossible to provide given their lack of availability. However ensuring independent living capability is crucial due to the demographic trend. According to a recent report by INSEE, dating from August 2017 in 2013, 21% of men and 48% of women aged 75 or over lived alone. It is essential to be supported the safe ingestion of food by those living alone and with swallowing difficulties. Responding to this challenge, this PhD research has designed, implemented and evaluation a system for the automation of dysphagia therapy support in a way that should be cost-effective, non-intrusive, and scalable for long-term therapy management. Accordingly, the objective of this research is to develop and validate a system for an adaptive Swallowing Events Recogniser and Choking Risk Assessor Adaptation (SER-CORA) through a non-invasive means of measuring the acoustic signature of the coordinative actuation of the muscles involved in safe swallowing.

This paper is focused on the automatic recovery of swallowing sounds from the captured ambient acoustic environment mix of sounds (speech plus noise) i.e. detecting and recognising the swallowing sounds to pave the way for the subsequent diagnostic analysis of such acoustic footprints of the swallowing process chain as illustrated in Figure 1 below

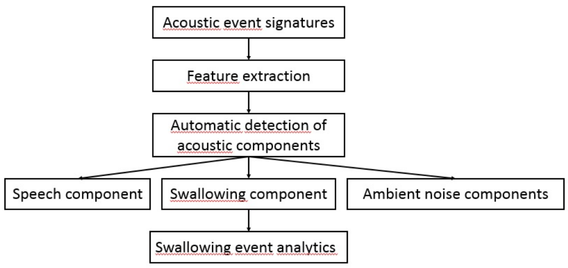

Figure 1: Analysis procedure of the proposed system.

2 ETHICAL DESIGN AND DEPLOYMENT

In routine use the system will detect and recognise swallowing events and can support the characterisation of the level of risk, sensing any choking precursors and alerting of actual chocking. Such alerts can be securely transmitted direct to the patient electronic record via Bluetooth so as to be available for periodic inspection and analysis by their clinician and physiotherapist to inform diagnostic and remedial protocols to be prescribed for the patient to support ongoing therapy management.

3 PHYSIOLOGY OF THE NORMAL SWALLOWING PROCESS

As illustrated in Figure 2 below, the swallowing process chain comprises of a set of coordinated muscle movements controlled by the cranial nerves, transports food from the mouth to the stomach passing by the pharynx while ensuring the safety of the respiratory tract as the paths of air and food cross in the pharynx. The physiology of normal swallowing was originally described in three sequential phases which are i) oral phase, ii) pharyngeal phase, and, iii) oesophageal phase for drinking and swallowing foods (solid/ liquid).

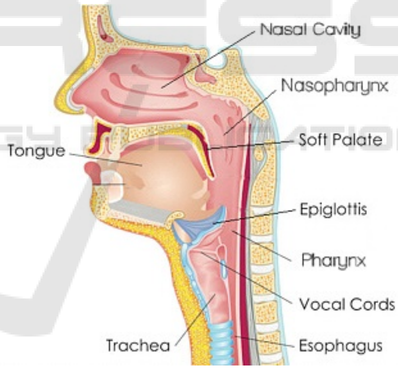

Figure 2: The anatomy of the human swallowing related organs.

An initial phase can be added to this model which is the preparatory phase for eating and swallowing solid foods. Once the bolus is formed and is ready to be expelled to the pharynx, the tip of the tongue is raised up and applied against the alveolar ridge of the upper incisors and the tongue takes the form of a spoon where the bolus slips and forms a single mass. The bolus is then moved backwards by the action of the tongue as gradually applied to the palate from front to back. At this stage, the soft palate ensures the closure of the oropharynx and prevents the penetration of the bolus into the pharynx; however, the larynx is still open. Simultaneously, the velum rises upwards to close the nasal fossae. By the time the bolus has reached the throat isthmus, the oral phase (the voluntary phase of the swallowing process) is over. The back of the tongue then moves forward and forms an inclination allowing the bolus to move towards the oropharyngeal cavity. In terms of control, the pharyngeal and oesophageal phases which constitute the swallowing reflex, without voluntary control, called the swallowing reflex, has thus been triggered as the food leaves the oral cavity and is moved on to the pharyngeal stage. Precisely at this mo-ment, the passage of food into the trachea is also pre-vented. The larynx opens during chewing to allow breathing and is closed as soon as the bolus arrives at the base of the tongue. Simultaneously, the vocal cords close ensuring airway closure, the moving cartilages of the larynx (arytenoids) swing forward in the laryngeal vestibule, covered by the rocking movement of the epiglottis. The larynx is pulled up and down by the hyoid muscles, which places it under the protection of the base of the tongue. Thus, the pharyngeal phase is triggered by the contact of the bolus with the sensory receptors of the throat isthmus and of the oropharynx. At this stage, breathing is interrupted and at the same time the last stage of swallowing begins with the bolus entering the oesophagus and being moved on through the oesophageal peristaltic waves (muscle contractions) to-wards the stomach. The pharyngeal process is a continuous phenomenon in time, considered as a reflex accompanied simultaneously by the velo-pharyngeal closure by the velum, by the laryngeal occlusion assured by the elevation of the larynx, and, by the retreat of the tongue base, the movement at the bottom of the epiglottis, the pharyngeal peristalsis and finally the opening of the upper sphincter of the oesophagus allowing the passage of the food bolus into the oesophagus. This phase lasts less than one second. The opening of the upper sphincter of the oesophagus is initiated by the onset of the pharyngeal peristalsis and food passage through the oesophagus is ensured by the ongoing coordinated peristalsis of pharyngeal and oesophageal stages.

4 DYSPHAGIA

Swallowing disorders or dysphagia can result from a wide variety of structural and/or functional deficits of the oral cavity, pharynx or oesophagus. Several abnormalities can cause a poorly-functioning oral cavity. These include: i) Cleft lip and palate and (ortho) dental pathologies which can inhibit mastication and impair the swallowing process. ii) Tumours affecting the head or the neck have the potential to cause oropharyngeal disorders by damaging the cranial nerves. iii) Xerostomia or "dry mouth" can disrupt the insalivation of the bolus and thus cause the accidental passage of granular, or inadequately chewed food, into the respiratory tract. iv) Infections can cause inflammation and ulceration thus reducing the masticatory performance. v) Chronic Gastroesophageal reflux disease (GERD) can weaken the oesophagus wall, affecting in some the normal peristalsis contractions for propulsion of food, and the functioning of the oesophagus sphincter. vi) Zenker's diverticulum or pharyngo-oesophageal diverticulum can occur in the pharynx or oesophagus and cause regurgitated into the pharynx, which may result in coughing or aspiration. vii) Dysphagia can also arise from iatrogenic dysfunction due to medication side-effects or surgical complications.

5 TEXTURE OF FOOD

Depending on the state of the "swallowing mechanism", some foods may be hard to swallow and a physiotherapist can suggest softer foods and beverages to meet nutritional needs. In this context, foods of varying texture, on a hard to easy-to-swallow spectrum are considered including solid foods ranging from the hardest-to-the-softest and smoothest to roughest/most-granular and liquids of varying stickiness and viscosity. Penman and Thomson's review of dysphagia diets (Penman and Thomson, 1998), maintains that thickened fluid can be classified, graduating from thin fluid such as water and all juices thinner than pineapple juice, to thick fluids such as milk and most fruit juices, and liquids thickened with starch to pureed consistency like yoghurt and compote. Solid foods can range from soft to hardest-to-swallow such as: fruits like melon and tomato, vegetables like carrots, potatoes and dry bread which is considered difficult to chew.

6 STATE OF THE ART IN SWALLOWING ANALYSIS

Neurological disorders associated with sensory and neuro-motor impairments arising from degenerative diseases and/or cardio-vascular incidents such a stroke may lead to disruption of the neuro-motor orchestration required to accomplish the swallowing process chain i.e. causing dysphagia. The clinical evaluation of swallowing disorders involves a number of tools. The Video-Fluoroscopic Swallowing Study (VFSS) which enables the real-time visualisation of bolus flow versus structural motion in the upper aero-digestive tract, is considered as the preferred instrument for dysphagia assessment by the clinicians. The VFSS helps identify the symptoms of swallowing disorders and localise the physiological causes triggering the typical repeated attempts at aspiration associated with dysphagia conditions (Dodds et al., 1990; Logemann, 1999). This enables the clinicians to make an accurate judgment about the patient's level of sensory or neuro-motor impairment and thus evaluate the extent of swallowing deficiency to be treated (Martin-Harris and Jones, 2008). Furthermore, ultrasound imaging has proven to be an excellent non-invasive means of studying the dynamics of the oropharyngeal system and the function of the muscles and other soft tissues of the oropharynx during swallowing (Jones, 2012). This tool is used to examine movements of the lateral pharyngeal wall (Miller and Watkin, 1997), and to visualise and examine tongue movement during swallowing (Shawker et al., 1983; Stone and Shawker, 1986). It has also been used to track the motion of the hyoid bone and thereby evaluate the extent of normal or abnormal swallowing (Sonies et al., 1996). Ultrasound imaging has also been used to assess the laryngeal phonation function by identifying the morphology of vocal folds and quantifying the horizontal displacement velocity of the tissue in the vibrating portion of vocal folds (Hsiao et al., 2001). Based on the assumption that it is advisable to lean forward whilst eating, a system comprising of an Inertial Measurement Unit (IMU) and an Electro-myographic sensor (EMG), has been presented (Imtiaz et al., 2014) to measure the head-neck posture and submental muscle activity during swallowing. Another study (Huckabee et al., 2005) has concerned the comparative evaluation of the influence of two swallowing manoeuvres (effortful versus normal swallowing) on an anterior supra-hyoid surface through EMG measurement and pharyngeal manometric pressure. They found significant change in both supra-hyoid surface EMG and pharyngeal pressures during effortful swallowing compared with normal swallowing. The same laboratory has conducted further research to prove that the tongue-to-palate emphasis during execution of effortful swallowing increases the amplitudes of submental surface EMG, the oro-lingual pressure, and the upper pharyngeal pressure to a greater degree than a strategy of inhibiting tongue-to-palate emphasis. The assessment of the functioning of the swallowing process is also supported by acoustic measurement and analysis. In 1990, Vice et al. (Vice et al., 1990) described the detailed pattern of throat sounds in newly born infants during suckle feeding. In the same way as the normal swallowing process can be divided into three stages (oral, pharyngeal and oesophageal phases), the associated swallowing sounds can be divided into three different segments (Vice et al., 1990): Initial Discrete Sound (IDS), Bolus Transit Sound (BTS) and Final Discrete Sound (FDS). The opening of the cricoi-pharynx in the pharyngeal phase contributes to the IDS and the bolus transition into the oesophagus in pharyngeal phase contributes to the BTS. The oesophageal phase then contributes to the FDS. By applying a Hidden Markov Model (HMM, researchers have also confirmed that sounds arising from the swallowing process chain comprise of three stages (Moussavi, 2005). An automated method has been set up to separate swallowing sounds from breathing sounds using multilayer feedforward neural networks (Aboofazeli and Moussavi, 2004). The algorithm was able to detect the acoustic footprints of 91.7% of swallowing events correctly. Lazareck and Moussavi (Lazareck and Moussavi, 2004a; Lazareck and Moussavi, 2004b) proposed a non-invasive, acoustic-based method to differentiate between individuals with and without dysphagia. A wearable swallowing monitor system was developed by Dond and Biswas, (Dong and Biswas, 2012) in order to assess the overall food and drink intake habits as well as other swallowing-related disorders. A piezo-respiratory belt able to be used for identifying non-intake swallows, solid intake swallows, and drinking swallows is reported to be under development. A study of dietary eating activity has been conducted using various sensors (Amft and Troster, 2006; Amft and Tröster, 2008). To monitor arm movements, they used an inertial sensor containing an accelerometer, gyro-scope and compass sensor. For chewing cycle recognition, they used an ear microphone and finally for swallowing process recognition, they used both the stethoscope and EMG electrodes attached at the infra-hyoid throat position. They found higher recognition accuracy for body movements and chewing sounds than for swallowing sounds which needed more investigation. Typical recognition rates achieved were around 79% for movements, 86% for chewing and 70% for swallowing. The recognition rate for swallowing was 68% using the fusion approach of EMG and sound versus 75% if using EMG alone and 66% using sound alone. Another study deployed a non-invasive wearable sensing system using a bone conduction microphone for capturing swallowing process sounds, counting the number of mastications and classifying the food swallowed according to its texture by means of frequency spectrum analysis (Shuzo et al., 2009). The large number of dysphagic patients has continued to inspire research on automated methods of monitoring and evaluation and classification of dysphagia to support functional rehabilitation of patients' swallowing disorders. In the context of the e-Swallhome project as part of the ANR Research Programme, our study involved the monitoring of the patients suffering from swallowing disorders. Accordingly our approach was motivated to support the automated monitoring of the swallowing process chain through a non-invasive and minimalist intervention strategy using only one wearable neck-attached microphone. Thus we have developed an acoustic-based system to detect and classify swallowing events to be used to identify cases of distress during food intake such as laboured aspiration responsive to swallowing distress, chocking which may or may not lead to asphyxiation and/or falling.

7 METHODOLOGY

Based on the observations as presented above, our analysis distinguished the sources of the factors affecting the swallowing process as being of either intrinsic (e.g. neuro-physiologically mediated) or extrinsic (e.g. food-texture-related) influence. Consistent with the above etiological analysis of causal and co-existing conditions, it was decided to recruit a structured sample of participants for swallowing event data acquisition including a representative number of participants to cover the above types of factors. This was to help enhance the generalisability of the adaptive algorithms to be designed to support Swallowing Events Recognition and Choking Risk Assessment (SERCORA). Accordingly, 27 participants were identified and a number of increasingly hard-to-swallow food types were selected for swallowing experiments to monitor the swallowing process. An ethical consent confirming process was completed as a pre-requisite to registration of the participants who were also to complete a questionnaire to declare any dysphagia condition or food allergies. Subsequently, the first phase of the research study was implemented as follows: In the initial stage of the study, 27 participants took part to enable their swallowing event data to be captured using a throat-mounted microphone while swallowing the following substances to establish a baseline profile for their swallowing as follows: i) Saliva; ii) Water; iii) Compote. Additionally, data was also captured from each participant when just breathing in order to establish the acoustic signature baseline associated with only the diaphragm movement. This protocol was carried out with two groups of volunteers: a control population of healthy persons with no dysphagia history and persons with swallowing disorders.

7.1 The Experimental Environment

In this project, non-invasive equipment was used to simultaneously record sounds associated with the swallowing process and also respiratory habits during food intake and other ambient sounds. These were recorded using a discrete module of miniature omni-directional microphone capsule (Sennheiser-ME 102) with IMG Stageline MPR1 microphone pre-amplifier placed on the neck at a pre-studied position midway between the Jugular Notch and the Laryngeal Prominence optimised for the most effective acquisition of the swallowing-elated signals, marked as P8 in Figure 3 presented below. Accordingly swallowing-related sounds were recorded at a 44.1 KHz sampling rate. The Signal was re-sampled at 16 KHz for the processing as required. Participants were seated comfortably on a chair. They were told that the equipment would record swallowing-related sound. The baseline function of each participant's swallowing process for the most fluid-like foods was established by capturing the data from participants during the swallowing of water and saliva as was also the case when fed with compote in a teaspoon. Additionally, they were asked to read aloud phonologically balanced sentences and paragraphs. The sounds of their coughing and yawning were also recorded. There were 39 recordings acquired from 27 people over about 11 hours. The sounds were manually annotated and a corpus of approximately 45 minutes of swallowing-related sounds, 90 minutes of speech and 35 minutes of other ambient sounds were extracted and archived. The table below shows the composition of the sounds.

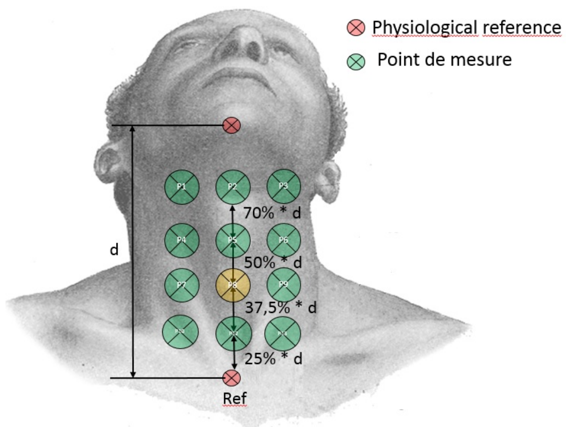

Figure 3: Best microphone position.

Table 1: Sound database.

Classes	Nomber of files			Duration
Swallowing	Compote	1943	565	45 minutes
	Water		644	
	Saliva		734	
Speech	705			1h 31 minutes
Sounds	839			35 minutes

8 ALGORITHMS AND RESULTS

8.1 Automatic Detection Algorithm

Automatic feature detection algorithms have been exploited in different domains (Boyer et al., 2016; Baptista et al., 2018; Brune et al., 2017). Recently, the analysis of swallowing sounds has received particular attention (Lazareck and Moussavi, 2002; Aboofazeli and Moussavi, 2004; Shuzo et al., 2009; Amft and Troster, 2006) whereby swallowing sounds have been recorded using microphones and accelerometers. The algorithms developed in our research enabled the automatic detection of swallowing-related sounds from a mixed stream of acoustic input as acquired through the neck-worm microphone positioned as described above. Frequency domain analysis enabled the determination of the frequency band associated with swallowing-related process chain; within this band relative prominence of frequencies varied according to the texture of the food being ingested. Our analysis established that swallowing of liquid-like foods such as compote was associated with a frequency signature with an upper range of 3617 Hz; whereas swallowing water was associated with a frequency signature below 2300 Hz and for saliva the corresponding maximum frequency component remained below a maximum of 200 Hz. Wavelet decomposition was then achieved using symlet wavelets at level 5. After comparison of different combinations of wavelet detail coefficients, details 5, 6 and 7 were selected for optimal decomposition and resolution of the swallowing-related sounds corresponding respectively to the frequency bands 500-1000 Hz, 250-500 Hz and 125-250 Hz; accordingly a new signal was reconstituted as the linear combination of the swallowing-related sounds thus recovered. This resulting signal was analysed using a sliding window running along the signal with an overlap of 50%. For each position of the current sliding window, simultaneously, an associated energy criterion was initially computed as the average energy of the ten preceding windows and accordingly a threshold for the start and end points of swallowing-related sounds was established. This threshold was calculated as a function of the average energy as follows:

$$Threshold = \varepsilon * APW_i + \alpha \tag{1}$$

where APW_i corresponds to an average of the ten preceding windows of each position (position i). Thus, following this framework, the starting point of a swallowing-related signal would be detected where the value of its associated energy as calculated according to the above method, exceeded the threshold and its end-point would be detected when the associated energy fell below the threshold level of the energy. However experiments showed the application of the above associated energy criterion to result in significant start-end detection errors as the extreme peaks of the signal still affected the start-end points detection disproportionately. Accordingly it was proposed not to stop detection of swallowing related sounds at the position where the associated energy falls below the threshold computed as described above but to continue detecting and validating through a two-pass process whereby validated start-end points were ultimately established accordingly to a reference which results in greater than 80%, recovery of swallowing events. Further-more, partial detection was characterised when the associated energy falls below 80% of the above reference and a false alarm was the detection when the associated energy did not correspond to that any annotated event as described above. The overall validated detection rate was of 86.09%. The validated reference rate was 22.19% and the partial reference rate was 63.90%. The rate of missed references was 13.91% and the false alarm rate was 24.97%. Figure 4 below displays the typical result of the automatic detection of sounds in our experiment whereby the green panel corresponds to the validated detections and the light grey corresponds to partial automatic detections

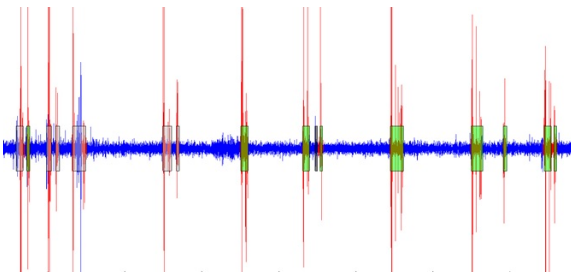

Figure 4: Example of validated and partial detections of swallowing-related sounds.

The overall automatic detection algorithm results are set out in table 2 as presented below.

Table 2: Automatic detection results.

Number of references to be detected = 3501			
Events		**References**	
Validated events	2365	**Validated references**	22.19%
Partial events	1052	**Partial references**	63.90%
False alarms	24.97%	**Missed references**	13.91%

8.2 Classification

The Gaussian Mixture Model (GMM) have been established as one of the most statistically mature methods for clustering (Reynolds and Rose, 1995). GMM have been deployed widely in acoustic signal processing such as speech recognition and music classification. Accordingly GMM are used to classify signals without a-priori information about the generation process; this was conducted in two steps as required: a training step and a recognition step. The training step with the Expectation Maximisation algorithm was used for classification of the three types of sounds in the input stream mix as set out previously: the swallowing-related, the speech and the ambient noise signal components. Parameters computed on sounds were the Cepstre at the outlet of a Mel scale filter bank, the Mel Frequency Cepstral Coefficients (MFCC) and in the linear scale, the Linear Frequency Cepstral Coefficients (LFCC), and their delta and delta-delta, differences between coefficients obtained from one analysis window to another, to include the signal temporal variations. In order to take into consideration the swallowing-related frequency band, only 14 coefficients were chosen through 24 filters. The parameters yielding the best results were the 14 LFCCs coefficients with delta-delta. Results with manually annotated segments showed an overall recognition rate of 95.49% as shown in table 3:

Table 3: Confusion matrix of manually annotated references.

ORR = 95.49%	**Swallowing**	**Speech**	**Sounds**
Swallowing	**100%**	0	0
Speech	0	**91.49%**	8.51%
Sounds	0	5.02%	**94.98%**

Next, a test was applied to automated detections. For validated detections, a satisfactory recognition accuracy of 93.32% for swallowing-related events was achieved and an overall recognition rate of 87.60%. For partial detections, an overall recognition rate of 71.59% was realised and a higher recognition rate of 80.42 was achieved for swallowing-related events using three classes as shown in the tables 4, 5 and 6 below.

Table 4: Confusion matrix of validated detections.

ORR = 87.60%	**Swallowing**	**Speech**	**Sounds**
Swallowing	**93.32%**	0.65%	6.03%
Speech	3.83%	**91.87%**	4.30%
Sounds	20.43%	1.96%	**77.60%**

Table 5: Confusion matrix of partial detections.

ORR = 71.59%	**Swallowing**	**Speech**	**Sounds**
Swallowing	**80.42%**	18.11%	1.47%
Speech	14.19%	**75.31%**	10.49%
Sounds	39.52%	1.43%	**59.05%**

Table 6: Confusion matrix of partial detections.

ORR = 88.87%	**Swallowing**	**Speech**	**Sounds**	**False alarm**
Swallowing	**87.10%**	12.58%	0.32	0%
Speech	3.46%	**95.60%**	0.94	0%
Sounds	25.44%	1.78%	**72.78**	0%
False alarm	0	0	0	**100%**

9 CONCLUSIONS AND OUTLOOK

In this paper, the rationale for a cost-effective non-intrusive automated care therapy management support system for dysphagia patients has been presented. This has enabled non-invasive adaptive Swallowing Event Recognition and Choking Risk Assessment "SER-CORA" using a minimalist intervention methods enabling the recovery and detection of the acoustic foot-prints of the swallowing-related events based on the acoustic stream captured through a single neck-worn microphone. With routine usage this method can con-tribute valuable data intelligence to establish refer-ence baselines to help characterise normal versus deviant swallowing event analytics to support the assessment of dysphagia conditions by clinicians, and the ongoing monitoring of the level of improvement responsive to any interventions or the emergence of any new pathology as determined and managed by the clinicians. Accordingly such capability for non-invasive Swallowing Events Recognition and Choking Risk Assessment Adaptation (SER-CORA) for each particular patient supports the detect of changes and therefore informs the initial screening and diagnosis of potential pathologies affecting the swallowing process chain and the localisation of the causes of dysphagia thus informing clinical investigation and treatment. This paper presented the initial phase of the SERCORA research study and the work ongoing in implementation of the final phase including a large number of participants and a wider variety of food types to realise the objective of providing scalable cost effective remote therapy management and early warning support dysphagia conditions.

ACKNOWLEDGEMENTS

This work was conducted within the eSwallHome (ANR research Programme) and was also funded within the framework of eBioMed chair of BMBI laboratory. This work is also a part of the STIC AMSUD-EMONITOR project. We would also like to thank Open Health Institute for their support for our research study.

REFERENCES

Aboofazeli, M. and Moussavi, Z. (2004). Automated classification of swallowing and breath sounds. *IEEE*, Vol 3:7803–8439.

Amft, O. and Troster, G. (2006). Methods for detection and classification of normal swallowing from muscle activation and sound. In *2006 Pervasive Health Conference and Workshops*, pages 1–10.

Amft, O. and Tröster, G. (2008). Recognition of dietary activity events using on-body sensors. *Artificial Intelligence in Medicine*, 42(2):121 – 136. Wearable Computing and Artificial Intelligence for Healthcare Applications.

Baptista, M., Silva, R., Jardim, H., and Quintal, A. (2018). Perioperative electronic system - a new approach for perioperative nursing performance. pages 505–509.

Boyer, C., Dolamic, L., Ruch, P., and Falquet, G. (2016). Effect of the named entity recognition and sliding window on the honcode automated detection of honcode criteria for mass health online content. In *In Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2016)*, volume 5, pages 151–158.

Brune, T., Brune, B., Eschborn, S., and Brinker, K. (2017). Automatic visual detection of incorrect endoscope adaptions in chemical disinfection devices. In *In Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017)*, pages 305–312.

Dodds, W. J., Logemann, J. A., and Stewart, E. T. (1990). Radiologic assessment of abnormal oral and pharyngeal phases of swallowing. *American Roentgen Ray Society*, Vol 154:965–974.

Dong, B. and Biswas, S. (2012). Swallow monitoring through apnea detection in breathing signal. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6341–6344.

Hsiao, T. Y., Wang, C. L., Chen, C. N., Hsieh, F. J., and Shau, Y. W. (2001). Noninvasive assessment of laryngeal phonation function using colour doppler ultrasound imaging. *Ultrasound in Med. & Bioll*, Vol 27:1035–1040.

Huckabee, M.-L., Butler, S. G., Barclay, M., and Jit, S. (2005). Submental surface electromyographic measurement and pharyngeal pressures during normal and effortful swallowing. *Arch Phys Med Rehabil*, Vol 86:2144–2149.

Imtiaz, U., Yamamura, K., Kong, W., Sessa, S., Lin, Z., Bartolomeo, L., Ishii, H., Zecca, M., Yamada, Y., and Takanishi, A. (2014). Application of wireless inertial measurement units and emg sensors for studying deglutition preliminary results. *IEEE*, Vol 978:4244–7929.

INED (February 12, 2013). "journées d'information sur les troubles de la déglutition". https://etablissements.fhf.fr/annuaire/hopital-actualite.php?id=998.

Jones, B. (2012). *Normal and abnormal swallowing: Imaging In Diagnosis And Therapyl*. Springer, 2nd edition.

Kawashima, K., Motohashi, P. Y., and Fujishima, P. I. (2004). Prevalence of dysphagia among community-dwelling elderly individuals as estimated using a questionnaire for dysphagia screening. *Dysphagia*, 19:266–271.

Lazareck, L. J. and Moussavi, Z. (2002). Automated algorithm for swallowing sound detection. In *Canadian Med. and Biol. Eng. Conf.*

Lazareck, L. J. and Moussavi, Z. (2004a). Classification of normal and dysphagic swallows by acoustical means. *IEEE Transactions on Biomedical Engineering*, Vol 51:0018–9294.

Lazareck, L. J. and Moussavi, Z. (2004b). Swallowing sound characteristics in healthy and dysphagic individuals. *IEEE*, Vol 3:7803–8439.

Logemann, J. A. (1999). Behavioral management for oropharyngeal dysphagia. *Folia Phoniatr Logop*, Vol 51:199–212.

Martin-Harris, B. and Jones, B. (2008). The videofluorographic swallowing study. *Phys Med Rehabil Clin N Am*, Vol 19:769–785.

Miller, J. L. and Watkin, K. L. (1997). Lateral pharyngeal wall motion during swallowing using real time ultrasound. *Dysphagia*, Vol 12:125–132.

Moussavi, Z. (2005). Assessment of swallowing sounds stages with hidden markov model. *IEEE*, Vol 6:7803–8740.

Penman, J. and Thomson, M. (1998). A review of the textured diets developed for the management of dysphagia. *J Hum Nutr Diet*, Vol 11:51–60.

Reynolds, D. and Rose, R. (1995). Robust text-independent speaker identification using gaussian mixture speaker models. *Speech Commun*, 17:121 – 136. Wearable Computing and Artificial Intelligence for Healthcare Applications.

Shawker, T. H., Sonies, B., Stone, M., and Baum, B. J. (1983). Realtime ultrasound visualization of tongue movement during swallowing. *Dysphagia*, Vol 11:485–490.

Shuzo, M., Lopez, G., Takashima, T., Komori, S., Delaunay, J., Yamada, I., Tatsuta, S., and Yanagimoto, S. (2009). Discrimination of eating habits with a wearable bone conduction sound recorder system. In *SENSORS, 2009 IEEE*, pages 1666–1669.

Singh, S. and Hamdy, S. (2006). Dysphagia in stroke patients. *Postgraduate medical journal*, Vol 82:383–391.

Sonies, B. C., Wang, C., and Sapper, D. J. (1996). Evaluation of normal and abnormal hyoid bone movement during swallowing by use of ultrasound duplex-doppler imaging. *Ultrasound in Med. & Bml*, Vol 22:1169–1175.

Stone, M. and Shawker, T. H. (1986). An ultrasound examination of tongue movement during swallowing. *Dysphagia*, Vol 1:78–83.

Vice, F. L., Heinz, J. M., Giuriati, G., Hood, M., and Bosma, J. F. (1990). Cervical auscultation of suckle feeding in newborn infants. *Developmental Medicine and Child Neurolog*, Vol 32:760–768.