Modelling Domain Knowledge of Speech and Language Therapy with an OWL Ontology and OpenEHR Archetypes

Vladimir Robles-Bykbaev[1,2], Martín López-Nores[2], José Pazos-Arias[2], Jorge García-Duque[2] and Juan Ochoa-Zambrano[1]

[1]*GI-IATa, CIDII, Universidad Politécnica Salesiana, Calle Vieja 12-30, Cuenca, Ecuador*

[2]*AtlantTIC Research Center for Information and Communication Technologies, Department of Telematics Engineering, University of Vigo, Vigo, Spain*

Keywords: Health Informatics, Speech and Language Therapy, Ontology, Archetypes.

Abstract: Researchers in the area of health informatics have made significant progress in the standardization of ICT tools to support the management, storage, retrieval and exchange of health-related data. However, the adoption of these advances is largely uneven across different areas. We present a comprehensive knowledge model for the realm of Speech and Language Therapy (SLT), based on an OWL ontology, normalized vocabularies and OpenEHR constructs. This model, validated by several collaborating institutions, is being used as the cornerstone to build a comprehensive framework with supporting tools for the different people involved in SLT, including therapists, patients and their relatives, and students.

1 INTRODUCTION

With estimations of nearly one billion people in the world living with some form a disability, adequate access to healthcare and rehabilitation services has become a major issue in global politics (WHO, 2011). The area of speech and language disorders is a small and often overlooked part of this picture, even though the development of proper communication skills is an important mainstay in a person's life, given that it allows him/her to express his/her feelings, needs and opinions as an active member of society. Nowadays there are no reliable and representative statistics of people suffering from communication disorders in the world, but the existing data show a complex outlook, due to the number of people affected (e.g. around 15 million people suffer from stutter in the world, whereas about 6 millions in the United States have other language impairments) and the lack of systematic approaches to develop and provide services of Speech and Language Therapy (SLT) (WHO, 2011; NIDCD, 2014).

SLT is a complex discipline that relates to many aspects of the patients' health condition, social environment and cognitive development. Accordingly, there are many people involved in the SLT processes (patients and their relatives, doctors, teachers, speech and language pathologists, …) and the flows of information (among homes, hospitals, schools, clinics, …) are very complex. On these grounds, it is important to develop a knowledge model of the SLT domain using standardized formal artifacts, in order to facilitate the exchange of information and thus foster the creation of convenient ICT tools, with which to support the initial speech-language diagnosis, the design of personalized therapy plans, the treatment of the wide spectrum of disorders, the monitoring of the patients' progress from the different points of view, and many other activities.

Most of research conducted hitherto in ICTs applied to the SLT area have focused on developing expert systems to support diagnosis and treatments of specific disorders, like dysphagia (Sharma et al., 2013), dysarthria and dyslalia (Schipor et al., 2012) or swallowing difficulties (Ward, Burns, Theodoros and Russell, 2014). These works did not pay any attention to the issues of sharing, maintaining and porting the clinical data over different platforms; rather, they relied on ad hoc languages, data structures and procedures. Only a few recent studies have used conceptualizations and classifications — in the form of an ontology— to support the operation of an expert system aimed at the initial

Robles-Bykbaev V., Pazos-Arias J., López-Nores M., García-Duque J. and Ochoa-Zambrano J..
Modelling Domain Knowledge of Speech and Language Therapy with an OWL Ontology and OpenEHR Archetypes .
DOI: 10.5220/0005279405850591
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2015), pages 585-591
ISBN: 978-989-758-068-0

diagnosis of language disorders (Martín Ruiz et al., 2014) but, to the best of our knowledge, there have been no approaches to support SLT within a fully integrative framework for clinicians and students, pathologists, patients, relatives and other potential users. In response to that, we hereby present a knowledge model for SLT that provides the foundations to build a comprehensive set of supporting tools for the following activities, among others:

- Accessing, sharing and querying the information according to specialized taxonomies of SLT concepts and user types.
- Automating statistical procedures to analyse the patients' evolution, the effectiveness of the applied therapies, common SLT patterns, behavioural patterns, etc.
- Automating the adaptation of contents to put in therapy plans or learning courses, according to SLT taxonomies and patient/student profiles.
- Integrating assistive technologies to provide support during the therapy sessions: robot assistants, mobile applications, remote software-monitoring, etc.
- Developing inference mechanisms for recommender and decision-support systems to assist in the preparation of therapy plans, the evaluation of exercise results, the generation of case studies, etc.
- Porting the data-structures through different architectures and systems.

Our knowledge model, preliminarily validated by SLPs from several collaborating institutions of speech and language therapy of Azuay - Ecuador, is based on an ontology that integrates concepts from standardized vocabularies from the American Speech-Language-Hearing Association (ASHA, 2014) and constructs from OpenEHR, an international standard to model healthcare information (www.openehr.org). Ontologies have been previously used in the e-health domain to model clinical data repositories (Rubi et al., 2014), whereas our research contribution has to do with using an ontology as an enabling tool for a set of ICT-based healthcare services in a very specific area.

The paper is organized as follows. The core ideas relating to the construction of the ontology are presented in Section 2, whereas Section 3 provides details about the methodology followed to populate it with instances of disorders, case studies, diagnosis information, exercises, etc. Section 4 contains an overview of a group of ICT tools we are developing on top of the knowledge model to support different aspects of SLT, including an expert system to automate the generation of therapy plans, a web/mobile portal to deliver training courses to students of phonoaudiology and a robotic assistant to support SLT sessions.

2 AN ONTOLOGY FOR SLT

Next, we will describe the main structures and elements of our proposed model. In the same way, we present two main diagrams to facilitate the comprehension of the model developed and how it is integrated in the research context for a comprehensive solution supporting SLT.

In order to provide a formal representation of the main health care concepts related with SLT and obtain the domain knowledge contained in the ontology, a team of engineers, SLPs and doctors of several collaborating institutions of special education have selected some of the most representative disorders, speech-language areas, and therapy-evaluation strategies. These were:

- *Disorders* (according to the classification provided in ASHA, 2014): dysarthria, expressive language disorder, dysphasia, dysphonia, speech and language developmental delay due to hearing loss, problems with swallowing and mastication, fluency disorder, moderate intellectual disabilities, severe intellectual disabilities, profound intellectual disabilities, infantile cerebral palsy (with the aim to offer SLT to children), and epilepsy and recurrent seizures.
- *Language and speech areas*: expressive language, articulation, receptive language, oral structure and function, hearing, and linguistic formulation.
- *Therapy strategies*: the ontology allows establishing several semantic relations between the therapy, educational contents, rehabilitation concepts, the patient's profile and the SL skills. Thereby, a speech-language skill must be able to adapt to patient's profile with which is related. For example, for a patient that suffers from cerebral palsy and severe athetosis and cannot produce speech, the SL skill representing communication through voice must automatically change to represent an alternative communication way (signs, gestures, etc.). Likewise, a given therapy plan could contain or not all SL areas before mentioned, under that a patient can only suffer a functional dyslalia and

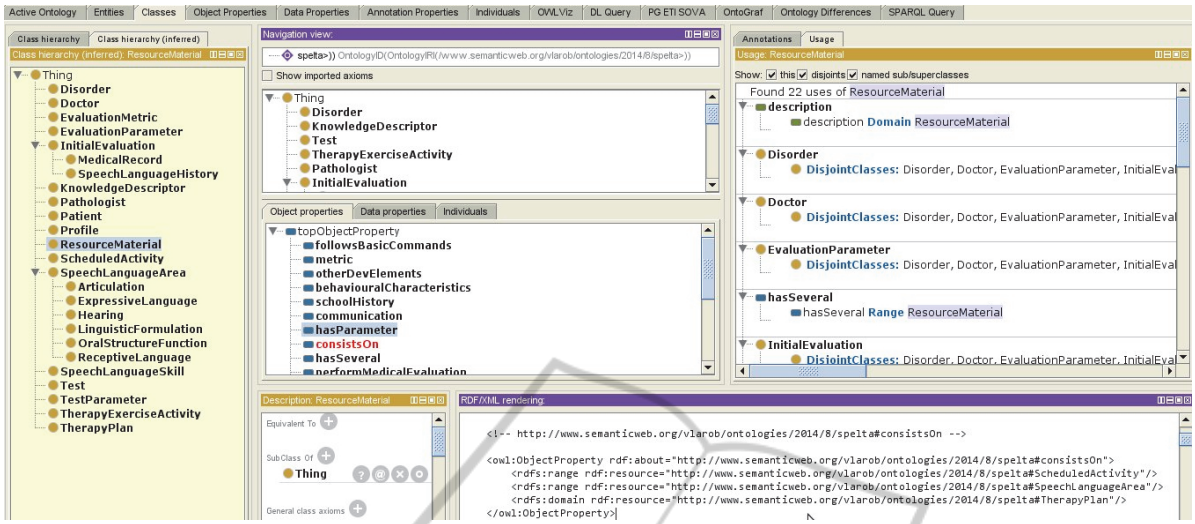


Figure 1: A tool snapshot showing some of the main elements and relations of the proposed ontology.

solely needs to do articulation reinforcement activities.

In Figure 1 we can see a screen capture of the class view menu in Protege (Riaño et al., 2012), depicting the hierarchy of classes, some of the object properties defined, some relations between *ResourceMaterial* and other classes, and the RDF (Resource Description Framework) source.

Figure 2 depicts a summarized view of the static structure (class diagram) of the knowledge model to support the SLT. The presented elements show the generalization and associations concepts as well as some of the main attributes of some classes. The key concepts, relations and information structures are as follows:

- *Initial patient evaluation*. In order to determine the best alternative to design a general therapy plan (the general guidelines to conduct therapy and stimulation exercises and activities), it is necessary to conduct an initial patient diagnostic. This diagnostic allows to determine the general medical condition (suffered diseases, serious surgeries or injuries, prescribed medications, …), initial condition, development history (i.e. data about fundamental aspects of the patient's childhood, such as the age at which he/she was able to sit up or walk) and basic skills of speech and language (communication skills, behavioural characteristics, hearing condition and school record). The initial patient evaluation is represented by *MedicalRecord* and *SpeechLanguageHistory* classes, and is conducted by doctors and speech-language pathologists (*Doctor* and *Pathologist* classes).
- *Patient's Profile.* A patient's profile is characterized by the initial evaluation, his/her personal data (name, date of birth, genre, ethnic group, etc.) and his/her speech-language screening (cognitive age, receptive language age, expressive language age, education level and suffered disorders). In order to prepare the patient's profile, it is necessary to evaluate him/her using special tests that belong to specific areas of Speech and Language (SL). The tests are represented by *Test* class and consist of several evaluation parameters (*EvaluationParameter*) that can represent a test question, a SL screening concept or a medical evaluation/screening parameter. Likewise, each test is related semantically with the disorders (class *Disorder*) that it helps to diagnose.
- *Therapy.* With the aim of developing a therapy plan, the SLPs must know the patient's skills in the several areas (SL, physical and cognitive). The SL skills are represented by *SpeechLanguageSkill* class and are characterized by the area to which they belong (*SpeechLanguageArea*). Similarly, an SL skill has several attributes that describe the knowledge to learn in the school context or in the language pillars (knowledge area, name, cognitive level, skill's description, etc.). With the goal to help patients to develop certain skills, a SLP needs to conduct several therapy exercises and activities. These exercises are represented by the *TherapyActivityExercise* class, have attributes that describe how to be performed (duration, repetitions, complexity, etc.), and consist of several material or resources (class *ResourceMaterial*).

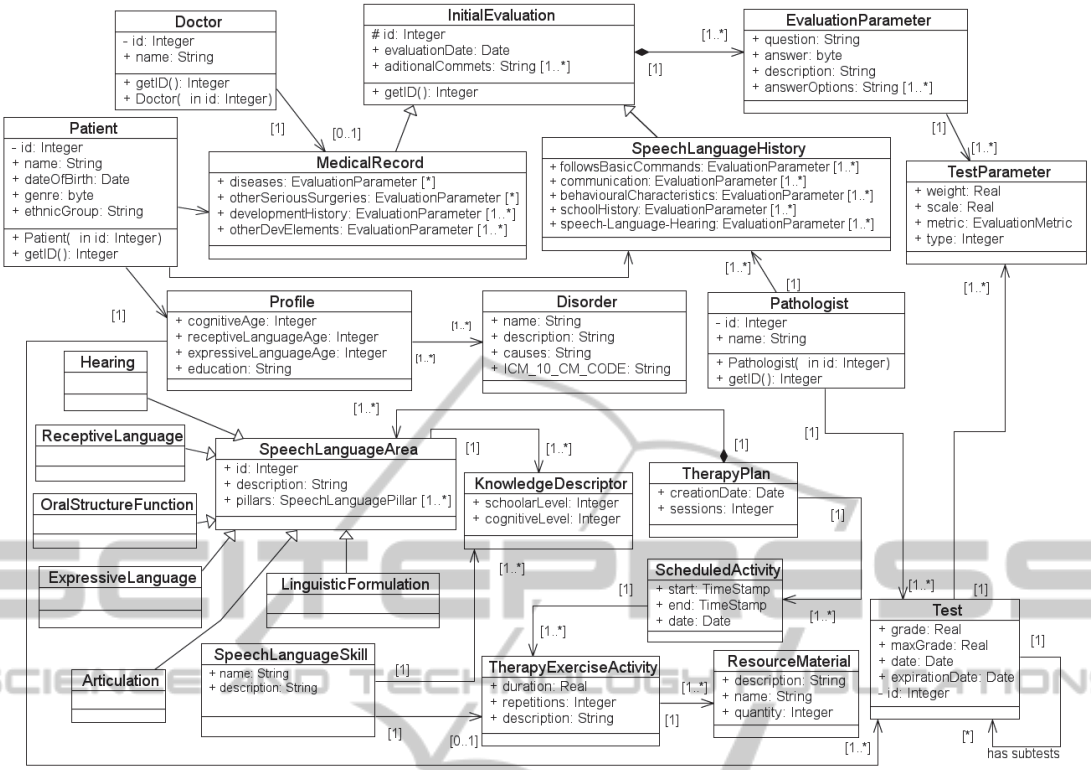


Figure 2: Partial view of the class diagram of the developed knowledge base (the monitoring concept is not shown).

- *Monitoring* (not shown in figures). Given that an SL evaluation is commonly conducted three times during the therapy period (at beginning, in the middle and at end), we propose to store the results of each of these evaluations, with the aim of enabling data mining activities and statistical analysis.

3 POPULATING THE ONTOLOGY WITH INSTANCES OF OPENEHR ARCHETYPES

The ability to use clinical concepts supported by standardized vocabularies and international medical guidelines, just like the ability to share and port the clinical information (semantic interoperability), is one of the most important features needed to support information systems for the healthcare area, regardless of nature and final objective (expert system, decision support system, data mining, ...). In this line, clinical archetypes appear as formal definitions of specific clinical concepts, defined as specializations of a generic reference model, that provide a mechanism to express data structures in a shared and interoperable way (Lezcano et al., 2011).

The OpenEHR specifications proposed using clinical archetypes to represent and manage information, in ways that have been proven to achieve good results in several healthcare areas like inter-institutional health data integration (Vieira-Marques et al., 2014), primary care attention in public institutions (Bacelar-Silva et al., 2013) and the development of health information systems (Kallel et al., 2011). On these grounds, we propose to use archetypes to represent the SLT formal definitions using vocabularies standardized by American Speech-Language-Hearing Association (ASHA, 2014). With the proposed model we can use ontologies to describe all SLT concepts using different levels of granularity, while the archetypes allow us to specify the information that should be captured in order to conduct each stage of SLT.

Figure 3 depicts an example of the main archetype (type *composite*), that has the following slots (sections): medical baseline screening, speech and language screening, therapy plan, and patient's profile.

Each one of these sections is defined in a separate archetype; for example, the speech and language screening section has several entries (evaluation date, additional comments, ...),

evaluation entries (if patient is able to follow basic commands, ...), and so on.

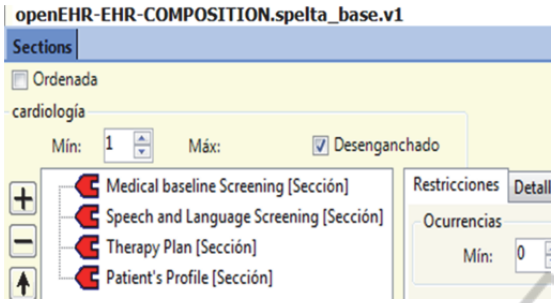


Figure 3: Screen capture of the main archetype (*composition*) of the proposed knowledge model.

4 RESEARCH CONTEXT FOR THE SLT KNOWLEDGE MODEL

With the aim of evaluating the proposed model, we have collected 130 real cases of children suffering from different kinds of speech-language disorders. Each of these cases was evaluated by doctors and SLPs from collaborating institutions. Likewise, all cases were collected using an online information system and contain the information structures mentioned in Section 2. Using this information, we are performing a stage to verify the consistency of the model.

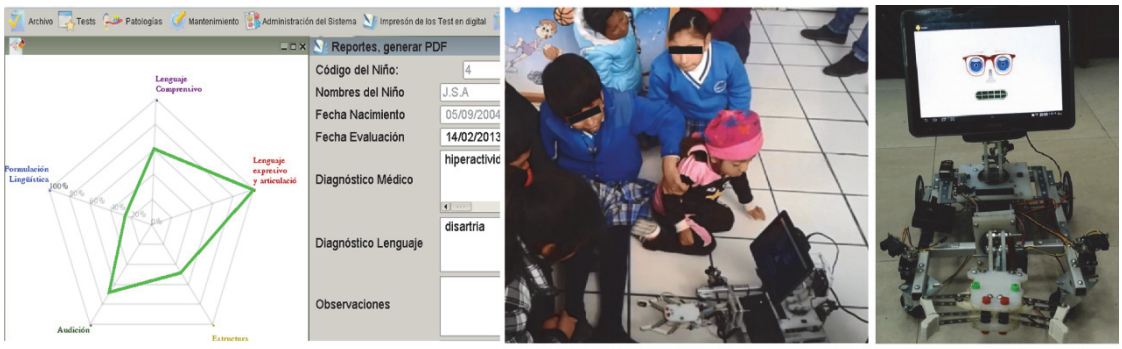
Furthermore, we are developing several ICT tools to support different aspects of SLT, including an expert system to automate the generation of therapy plans, a web/mobile portal to deliver training courses to students of phonoaudiology and a robotic assistant to support SLT sessions. This environment of SLT backing tools consists in the following elements (Figure 4):

- *Database Layer.* All the information represented in the knowledge model is stored in the database. Given that our ICTs tools include a mobile version of the support system for SLT, the information collected using mobile devices is previously stored in his own local base and then will be synchronized with the server database.
- *Knowledge Model.* This layer contains the ontologies, the archetypes and the standardized vocabulary aforementioned.
- *Expert System Layer.* The expert system is able to automatically infer speech-language therapy plans, using those plans manually developed by SLPs. In order to infer a new plan, the system uses the patient's profile information, the results obtained in the screening tests, and the patient's skills/weakness in the speech-language areas mentioned in the section 2. Moreover, the content generation is a process that produces new exercises and therapy activities for patients that suffer slight or moderate speech-language disorders.
- *User Interface Layer.* In this layer are provided software and hardware tools to support the different activities conducted by SLPs or students. The mobile application contains the screening tests to help to reduce time during the SL patient's evaluation. Likewise, it is able to query the information contained in the server database (patient's profile, previous evaluation results, progress reports, ...). The web training system for students is focused in learning activities for future SLPs: asking to develop therapy plans, analysing cases (according to given disorders and language skills), and others. The desktop application has the same functionalities of the mobile version, and is intended for providing portability. The robotic therapy and monitoring assistant consists of two elements: a mobile device (smartphone or tablet) and a displacement electronic platform. In the mobile device is presented an avatar that interacts with patients (especially children), and is able to tell stories, detect some hand gestures (if the user wears a glove of certain color), receive voice orders, and has a remote console to be controlled in "manual mode" by the SLPs. The platform allows robot to move in the floor using two servomotors.

The Figure 4 presents 3 elements: in the left side (a) is shown a screen capture of the desktop version to support SLT where are depicted the patient's progress in the 5 speech-language therapy areas (the expressive language and articulation are presented together). In the centre (b) of the figure we can see an interaction between the robot and children suffering from different speech and language disorders. The robot is telling a story with the aim to relax the children, as previous activity to start the therapy. In the right side of the figure (c) is shown a photo of the robot, where is possible to see the avatar displayed in the mobile device screen as well as the displacement platform.

5 CONCLUSIONS

The proposed approach presents an innovative model that relies on ontologies, archetypes and standardized vocabularies, and provides an integrative environment able to include new therapy



(a) (b) (c)

Figure 4: Screen capture of some applications belonging to user's layer.

reinforcement pedagogical elements, applications for monitoring and educational software (robotic assistant, mobile versions of the application, educational web-based system, ...). In the same way, this model supports the semantic portability and is capable to support data mining processes and intelligent content generation (automatic SLT plans generation). Moreover, the researches carried out are focused on finding/providing solutions to treatment of some specific disorders, initial screening of disorders or the use of ontologies to sustain more elaborated processes in the initial SL diagnosis.

As future work in the speech-language therapy domain-modelling, we are starting to extend our model with the aim of covering some deeper areas of knowledge, like the automatic generation of specific therapy plans based on daily activities and considering these elements: the existing levels of granularity of the patient's cognitive development, the incidence of other disorders (cerebral palsy, athetosis ...), etc.

ACKNOWLEDGEMENTS

The authors from the University of Vigo have been supported by the European Regional Development Fund (ERDF) and Xunta de Galicia under project CN 2012/260 "Consolidation of Research Units: AtlantTIC", and by the Ministerio de Educación y Ciencia (Gobierno de España) research project TIN2013-42774-R (partly financed with FEDER funds). We would like to thank the support provided by María Augusta Zambrano and the "Jesús para los Niños" foundation of Cañar, Ecuador.

REFERENCES

American Speech-Language-Hearing Association (ASHA), 2014. *International Classification of Diseases, Tenth Revision - Clinical Modification, Related to Speech, Language, and Swallowing Disorders*. (ICD_10_CM).

Bacelar-Silva, G. M., Cesar, H., Braga, P., and Guimaraes, R., 2013. OpenEHR-based pervasive health information system for primary care: First Brazilian Experience for Public Care. In *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, pp. 572-873. IEEE.

Kallel, F., Ellouze, A. S., and Bouaziz, R., 2011. Generating Context-Awarness Interface for Medical Applications. In *Developments in E-systems Engineering (DeSE)*, 2011, pp. 105-110. IEEE.

Lezcano, L., Sicilia, M. A., and Rodríguez-Solano, C., 2011. Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules. In *Journal of biomedical informatics*, vol. 44, no. 2, pp. 343-353.

Martín Ruiz, M. L., Valero Duboy, M. Á., Torcal Loriente, C., and Pau de la Cruz, I., 2014. Evaluating a Web-Based Clinical Decision Support System for Language Disorders Screening in a Nursery School. In *Journal of medical Internet research*, vol. 16.

National Institute of Deafness and Other Communication Disorders (NIDCD), 2014. Statistics on Voice, Speech and Language.

Riaño, David, and et al., 2012. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. In *Journal of biomedical informatics*, vol. 45, no 3.

Rubi, R., Pastor, X. and Lozano, E., 2014. OWLing Clinical Data Repositories With the Ontology Web Language. In *JMIR Medical Informatics*, vol. 2, no 2.

Sharma, S., Ward, E. C., Burns, C., Theodoros, D., and Russell, T., 2013. Assessing dysphagia via telerehabilitation: Patient perceptions and satisfaction. In *International journal of speech-language pathology*, vol. 15, pp. 176-183.

Schipor, O. A., Pentiuc, S. G. and Schipor, M. D., 2012.

Automatic Assessment of Pronunciation Quality of Children within Assisted Speech Therapy. In *Electronics & Electrical Engineering*, pp. 15-18.

Vieira-Marques, P., Patriarca-Almeida, J., Frade, S., Bacelar-Silva, G., Robles, S., and Cruz-Correia, R., 2014. OpenEHR aware multi agent system for inter-institutional health data integration. In *Information Systems and Technologies (CISTI), 2014 9th Iberian Conference on*, pp. 1-6. IEEE.

Ward, E. C., Burns, C. L., Theodoros, D. G., and Russell, T. G., 2014. Impact of dysphagia severity on clinical decision making via telerehabilitation. In *Telemedicine and e-Health*.

World Health Organization (WHO) and World Bank, 2011. World Report on Disability, Malta: WHO Press.