Tempo : towards a conceptualisation of pathology in speech and language therapy

Frédérique Brin-Henry, Rute Costa, Sylvie Després
(1) ATILF UMR 7118 CNRS-Université de Lorraine, 44 avenue de la libération, 54000 Nancy, France
(2) NOVA CLUNL (Centro de Linguística da Universidade NOVA de Lisboa), Lisbon, Portugal
(3) LIMICS (laboratoire d'informatique médicale et d'ingénierie des connaissances en e-Santé) INSERM, Paris, France
frederique.henry@atilf.fr, rute.costa@fcsh.unl.pt, sylvie.despres@univ-paris13.fr

Abstract

Tempo (TEMporalité, Pathologie orthophonique, Ontologie) is an attempt to transform a draft conceptual model created for the French orthophonie/speech and language therapy (SLT) community into an ontological resource focusing on language and communication pathology. It is interdisciplinary (SLT, linguistics, terminology and descriptive logic) and focuses on temporality as an entry point. It is part of a 3 step project aiming at harmonising European SLT terminology in order to facilitate communication about language pathology. The first step was to extract some characteristics using semantic and syntactic analyses, the second one is the creation and validation of the ontology (Tempo) and it will be followed by a larger multilingual terminological project at a European level.

Keywords: ontology, terminology, temporality, pathology in speech and language therapy.

1 Introduction

Tempo (TEMporalité, Pathologie orthophonique, Ontologie) is an attempt to transform a draft conceptual model created for the French orthophonie/speech and language therapy (SLT) community into an ontological resource focusing on language and communication pathology. We opted for an original viewpoint in pursuing temporality throughout the concepts of pathology. The aim is to work towards a definition of the concept of PathologieOrthophonique that is specific to SLT and holds clinical, ontological, terminological and linguistic relevance, and contributes to harmonious and precise SLT terminology. The scope of the model is for now voluntarily limited to French SLT; however, part of the project will include other SLT communities using other European languages (English and Portuguese texts and diagnostic terms).

This project represents a fundamental standpoint on how we can work on the specialists’ concepts of pathology from a tripartite perspective. Its researchers belong to three complementary...
disciplines. On the one hand, the expertise of an experienced SLT - who is also a linguist - ensures clinical and terminological relevance. On the other hand, the views of an expert in model and descriptive logic ensures that the built ontological resource is consistent and in relation to the conceptual model. A linguist and terminologist questions and guarantees the proper correspondence of a formal definition of the concept of PathologieOrthophonique and sub-concepts into natural language.

2 Context

This project is part of a larger attempt to clarify terminology used by speech and language therapists (SLTs). Once completed and validated, it supports the idea that SLTs have a specific conceptual representation of language pathology. We consider parameters more usually present in the Humanities than in the medical world, and especially temporality. This ontological resource (OR), called TemPO, focuses mainly on diagnostic labels, but we aim to extend the project in a near future to other conceptual nodes (Vandaele, 2001), like assessment or therapy. The OR along with a corresponding terminology could be used by SLTs to support the diagnostic phase of their work.

Speech and Language therapists use diagnostic labels to name all of their patients’ difficulties in an official, and presumably, consensual manner (Brin-Henry et al, 2018a). All health professionals encounter many obstacles in this process because in order to give a label to the patients’ difficulties, they need to observe and assess their patients, and find a term that corresponds to their representation of the pathology. This may be considered as an onomasiological process, by which a concept is linked to a term. This process is particularly difficult in SLT because the focus is on communication, language, and other motor functions related to the face and larynx. The challenge is to find true evidence for choosing a label in relation to three areas: the patients’ condition, the representation of the disease that the SLT has, and the available terminology. These three aspects are detailed hereafter:

- The reality of the patients’ difficulties, the strengths (Braun et al., 2016) and the weaknesses are identified through the process of assessing the disorders and hearing what the patient tells the SLT. For instance, a patient may explain that her speech is impaired at a constant level since her stroke and quotes: “it is as if I had been drinking alcohol”. The SLT proceeds to evaluate the difficulties and their severity.
- The representation that the health professionals (HPs) have of the disease (including SLTs), may evolve and vary in time and place, depending on each professional’s career path and on recent scientific progress. In this example, the SLT will retrieve the knowledge he/she has of the possible consequences of stroke, and the way it correlates to speech problems.
- Finally, a corresponding term has to be available, in a clear and relevant classification, in order for the HP to find it and decide it is the most suitable term for this particular situation. In our example, the SLT may choose diagnostic labels to name the disease, e.g. a single term "dysarthria", or a complex term "motor speech disorder".

2 In this paper, “diagnostic labels” refer to the end-result of a clinical and terminological process by which SLTs name the patients’ difficulties.
Naming pathology is therefore a complex process, requiring both expertise and knowledge. The professionals need to use a coherent and structured terminology, which corresponds to their everyday clinical needs. One of the strongest obstacles in the naming process is the fact that the classifications SLTs use, are mainly based on a biomedical conceptualisation of language and communication disorders (DSM-5, CIM-11) and in particular medical activities and symptoms. Some medical ontologies or vocabularies, such as Unified Medical Language System (Bodenreider, 2004) attempts to group several vocabularies, and is used to create operational terminologies for coding (SNOMED-CT), which induces payment of hospital activities (CCAM). Medical libraries also use thesaurus for indexing (MeSH).

We do not wish to reject the objective organic part of language pathology, nor the necessity to examine the etiological grounding or the anatomo-clinical correlations, which are frequently at the foundation of the classifications. However, our experience (20yrs+) of practice as a SLT, as well as some previous terminological studies and corpus linguistic analyses in the field of Speech Therapy, show that SLTs create a great number of different complex terms and collocations to replace the term they cannot find in the current classifications. Moreover, similar terms in different languages do not always designate the same concept, showing that the way SLTs translate the reality of the patient is different from one country to another. Lastly, we know that some characteristics of pathology, such as temporality do not appear in current classifications. Pathologies may evolve in time, and recent scientific progress may have an effect on therapies, and on patients.

We will give the example of aphasia. The description of aphasia (acquired language disorder) occurring in stroke is nowadays very different to the way Trousseau and Broca originally considered it in the mid 19th century (Ardila, 2010). Indeed, some factors have widely influenced in the way language disorders occur. Early intervention of SLTs in acute stroke units has proven useful in avoiding severe consequences in the patients’ communication and language skills. For instance stereotypical utterances have clearly been encountered less since early rehabilitation enables patient to recover earlier the ability to speak again (what French SLTs call “démutisation”). In addition, SLTs have an influence on the level of knowledge families and other HPs may have on aphasia and its consequences, thus helping the patients in his/her recovery (Grawburg et al., 2013). Secondly, recent medical breakthrough in the treatment of cerebral vascular accidents reduce the complications due to artery blocage (also called ischemia). Early thrombolysis is a treatment in which, in some cases, the clot blocking the brain artery may dissolve. Therefore, the brain necrosis can be reduced and the symptoms related to ischemia may vary from their usual combination (Jacquin et al., 2014), or even disappear. Time is relevant in all these aspects.

---

3 For example, the French term "dysphasie" is not the equivalent of the English term “dysphasia”, and it doesn’t either correspond to the normally used equivalent “specific language impairment”.

4 In this paper, the term “characteristics” relates to an abstraction of a property of an object or of a set of objects used for describing concepts (ISO 1087 standard).

5 "stéréotypie" is an irrepressible replacement of all words by the same one. Some well-known stroke victims, such as the French poet Charles Baudelaire have unfortunately been stricken by this extremely persistent disorder. He was famously known, after the age of 40, to be only able to repeat “cré nom”, a simplified form of the swear word “sacré nom de Dieu” (Dieguez, Bogousslavsky 2007).
In conclusion, we know SLTs need terms to describe the reality of their patients’s life with a particular condition. The current classifications do not offer enough relevant terms, hence their terminological creativity. Temporality is worth examining and we would like to clarify SLT terminology to improve communication about language and communication pathology, amongst SLTs, other HP and as a consequence between SLTs and patients.

3 A conceptual model based on temporality

Identification and definition of concepts and their designation (to be used adequately in discourse activities) are fundamental steps for this terminological project aiming at the organisation of knowledge in the area of Speech and Language Therapy. According to the standard (ISO 704:2000, p. V), "Objects, concepts, designations and definitions are fundamental to terminology work [...]. Objects are perceived or conceived and abstracted into concepts which, in special languages, are represented by designations and/or definitions. The set of designations belonging to one special language constitutes the terminology of a specific subject field.". We adopted this viewpoint, and consider concepts, terms and definitions to be central in addressing the concept of PathologieOrthophonique. An ontology is defined as "a formal, explicit specification of a shared conceptualization" (Studer et al., 1998), which is an important tool for modeling, sharing and reuse of domain knowledge. It allows domain knowledge to be represented explicitly through concepts and relations between them and hence to manipulate it automatically. Moreover, a terminology, according to ISO 1087:2000, is defined as “a set of designations belonging to one special language” and in this framework, term can be considered a “verbal designation of a general concept in a specific subject field”. From a morphosyntactic point of view, a term assumes the linguistic form of a simple or a multiword term (compound or complex term).

3.1 Building the model

Traditionally, two methods coexist in terminological work. The first one is the onomasiological approach, in which the concept (given by experts) is the starting point (Wüster, 1985). The aim is to give a definition of the concept and to discriminate it from all the other concepts within the same system. Finding the lexical unit – the term – to describe it comes later. In the semasiological approach, the starting point is the term, i.e. the designation of the concept, which is central in the building of the discourse, and which indicates the existence of a concept shared between the members of a group of experts, who also form a communicating group. In this second approach, the aim is to use linguistic analyses to identify the terms used by the group who shares common knowledge. Our work has used both approaches as complementary methods, at different times. In TemPO, the starting point was our expert knowledge of the discipline (regarding both concepts and used terms), combined with the results of some previous work done on SLT terminology using semi-automatic methods on 2 corpora. The first corpus (Brin-Henry, 2011) gathered more than 460 SLT reports (in French), showing an example of the clinical use of diagnostic labels. The
second corpus⁶ (Brin-Henry et al., 2018b) is made out of 957 articles (4 million words) written between 1997-2014 in the French scientific journal Rééducation Orthophonique. We used the conclusions of some syntactic and semantic analyses⁷ on these two types of written professional and scientific texts, at a phraseological level, exploring motives and routines (Née et al, 2014), human nominalization (Barque, 2015), or right and left contexts (Anscombe, 1995). We also considered the lexical unit level, examining collocations and frequencies. (for more info, see Brin-Henry 2015, 2016, 2017, 2018).

The semantic and syntactic analyses were developed in 3 directions (Brin-Henry, forthcoming): linguistic specificities (format and semantic categories), frequent and central terms and their terminological status (e.g. “difficulté(s)”, “trouble(s)”), and the relationship between SLT and their patients (study of introductory verbs and naming humans). These analyses along with the clinical experience we mentioned earlier led us to list some characteristics (see TABLE 1) that were usually expressed in the written texts. We are now trying to define the concepts that reveal what a SLT pathology may be. The main focus is temporality.

### 3.2 The importance of temporality

The objective of TemPO is to produce an ontology with a new perspective on the concept of PathologieOrthophonique according to the parameter of temporality. We consider temporality crucial because it holds very important information for the patients and the therapists about the diagnosis and prognosis of pathology. First, this concept relates to the onset-time, i.e. the moment when the pathology starts or happens. It relates to the way the pathology evolves (or not), thus showing whether the symptoms should decrease and/or disappear, or increase, or even be stable. Finally, we consider that the duration of the pathology is also relevant. A patient recently gave us an example of the importance of the temporality. He recalled how he realised that some symptoms were prodromal of what had later been diagnosed as Parkinson’s disease. What could have been caused by an acute problem (i.e. a stroke) was later considered due to the neurodegenerative disease. These facts are seen through the lense of a singular and personal story, but could modify the way we look at pathology. After having identified a certain amount of essential characteristics, we drew a hierarchical model, which helped us to examine the importance of each concept. We split pathology into progressive/non progressive, and worked our way down each characteristic. The following binary couples form the base of our reasoning.

---

⁶ This corpus has been made available to the scientific community through the OrthoCorpus Project, which received funding from the Lorraine Region, FNO, Ch Bar le Duc, and ATILF. It may be downloaded from www.ortolang.fr.

⁷ Performed semi-automatically using the concordance AntConc (Anthony, 2014) and the textometry platform TXM (Heiden et al, 2010).
### Table 1: Characteristics identified to build the conceptual model

<table>
<thead>
<tr>
<th>Property</th>
<th>Characteristics of the pathology in SLT</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evolution</td>
<td>progressive or</td>
<td>the patient’s condition will get better or worse</td>
</tr>
<tr>
<td></td>
<td>non progressive</td>
<td>the patient’s condition is stable</td>
</tr>
<tr>
<td>Start</td>
<td>onset time</td>
<td>the time when the pathology starts or happens</td>
</tr>
<tr>
<td>Structure</td>
<td>impairs the structure/or the use of the structure</td>
<td>when particular system, such as the visual system, is impaired in itself,</td>
</tr>
<tr>
<td></td>
<td>impairs the use of the structure</td>
<td>when it is the way the system is used which is impaired</td>
</tr>
<tr>
<td>Type of structure</td>
<td>motor</td>
<td>affecting motor abilities, e.g. the movement of the tongue</td>
</tr>
<tr>
<td></td>
<td>or neuropsychological</td>
<td>affecting neuropsychological abilities, e.g. our memory</td>
</tr>
<tr>
<td>Language-driven</td>
<td>impairs language</td>
<td>affects linguistic skills, e.g. the way a person pronounces words,</td>
</tr>
<tr>
<td></td>
<td>impairs non-linguistic abilities</td>
<td>affects skills not directly involving linguistic abilities, e.g. the ability to concentrate on a particular task</td>
</tr>
<tr>
<td>Modality</td>
<td>written language</td>
<td>when pathology involves oral language</td>
</tr>
<tr>
<td></td>
<td>oral language</td>
<td>when pathology involves written language</td>
</tr>
</tbody>
</table>

We think this method of describing the pathology in SLT is innovative and gives a better account of how SLTs consider pathology and their patients. It shows a change of paradigm from the usual biomedical conception of pathology, and relates strongly to clinical practice. In that way it shows some relation to narrative medicine, however we still think the focus on the SLT representation of things adds value to a simple patient-centered approach. The association of a team specialised in linguistics, terminology, SLT, and ontology in this project provides a thorough examination of the subject.

## 4 Construction of the ontology from the draft conceptual model

The ontology construction is based on the Neon methodology scenario 8, named "From specification to implementation". In this scenario, ontology developers should specify first the requirements that the ontology should fulfil, by means of the ontology requirements specification

---

activity. Then, the ontology developers assigned to the ontology project should carry out (1) the ontology conceptualization activity, in which knowledge is organized and structured into meaningful models at the knowledge level; (2) the ontology formalization activity, in which the conceptual model is transformed into a semi-computable model; and (3) the ontology implementation activity, in which a computable model (implemented in an ontology language) is generated (Suárez-Figueroa et al., 2015). The scope of the TemPO ontology has been determined by scenarios of use (Table 2) and by competency questions (Table 3) (Uschold and Gruninger, 1996) which were discussed throughout the transformation of the conceptual model into an ontology.

<table>
<thead>
<tr>
<th>Scenarios of use</th>
<th>Ontology content</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1: The SLT proceeds to evaluate the difficulties of the patient and their severity.</td>
<td>Ontology must allow to represent the concept of process. In this ontology version, the assessment concept appears in the conceptual model but is not detailed.</td>
</tr>
<tr>
<td>S2: The SLT decides whether to quickly implement some compensation strategies</td>
<td>The ontology must make it possible to represent the dynamics of the pathology so that the speech therapist can take into account (for instance, if the pathology is stable or progressive).</td>
</tr>
<tr>
<td>S3: The SLT gives the patient some information on the pathology they are dealing with</td>
<td>Ontology must permit the identification of the impaired target (structure/system, language, memory) by the pathology.</td>
</tr>
</tbody>
</table>

Table 2: scenarios of use and the ontology content

<table>
<thead>
<tr>
<th>Competency questions</th>
<th>Corresponding scenarios of use</th>
</tr>
</thead>
<tbody>
<tr>
<td>The wife of a person with Alzheimer’s disease asks his SLT whether his reading skills will improve or not with therapy.</td>
<td>S3</td>
</tr>
<tr>
<td>A person has only recently been diagnosed with Parkinson’s disease. He thought he had had a stroke because half of his face felt different.</td>
<td>S1</td>
</tr>
<tr>
<td>A parent asks how long it will take for his 10-year-old dyslexic son to be able to read and understand a page</td>
<td>S1, S2, S3</td>
</tr>
</tbody>
</table>

Table 3: Competency questions and corresponding scenarios of use

In the draft conceptual model of pathology in SLT (Figure 1), the concept <Pathologie du langage> is central. The identifier <Pathologie du langage>9 was chosen because it corresponds to the French general designation "pathologie du langage" in use among the professionals of the domain of SLT10. It is subdivided into two sub-concepts <Pathologie du langage évolutive> and <Pathologie du langage non évolutive>. This division highlights the importance of identifying how the pathology will evolve in real life. The next division highlights the moment when a pathology is starting or happening by introducing the notions of "pathologie développementale",

9 For operationalisation questions, we decided in the TempO to rename it into PathologieOrthophonique.
10 In this paper, <pathologie du langage> refers to the language, communication and oromotor difficulties established in a person (« ensemble des difficultés du langage de la communication et des fonctions oromyofaciale mise en évidence chez une personne »).
"pathologie dégénérative", "pathologie congénitale" et "pathologie acquise". Then the property "affecte" links these different pathologies to the system parts that are impaired by the pathology.

From the ontologist’s point of view, the fact that a pathology is stated as being progressive or not is a discriminant qualification. Therefore, the notion of temporality will be qualified by means of a property used for splitting the concept PathologieOrthophonique in two sub-concepts PathologieEvolutive and PathologieNonEvolutive. The object property $a_{DynamiquePathologie}$ created for characterising the temporal aspect of the pathology is defined as follows:

- $a_{DynamiquePathologie}$(domain: PathologieOrthophonique, range: DynamiquePathologie).

The concept DynamiquePathologie = {evolutive, nonEvolutive} is defined as an enumerated type. Then four subconcepts PathologieCongenitale, Pathologie Developpementale, PathologieAcquise, PathologieDegenerative are defined as sub-concepts of PathologieOrthophonique. Each of these can be qualified by means of properties describing the start of the pathology, offering a possibility to differentiate them. A discussion led to the added property of duration, giving yet another aspect of temporality. The object properties created for characterising the duration, the onset time of the pathology are defined as follow:
- **aDureePathologie**:(domain: PathologieOrthophonique, range: DureePathologie. The concept DureePathologie = {durable, temporaire} is defined as an enumerated type.

- **aDebutInstallation**:(domain: PathologieOrthophonique, range: DebutInstallation). The concept DebutInstallation = {accidentVie, debutVie, developpementApprentissage} is defined as an enumerated type.

Once these concepts are described, it is possible to agree on some defined classes\(^\text{11}\) to allow some reasoning on pathologies in SLT. Four defined classes are constructed:

- **PathologieDurable** ≡ PathologieOrthophonique AND (aDureePathologie value durable)

- **PathologieEvolutive** ≡ PathologieOrthophonique AND (aDynamiquePathologie value evolutif)

- **PathologieNonEvolutive** ≡ PathologieOrthophonique AND (aDynamiquePathologie value nonEvolutif)

- **PathologieTemporaire** ≡ PathologieOrthophonique AND (aDureePathologie value temporaire)

The Protégé OWL editor (Version5.2) was used as a tool for building TemPO in Ontology Web Language (OWL2) format\(^\text{12}\). TemPO was tested with regard to formal consistency and absence of cycles using the Fast Classification of Terminologies FaCT++ description logic reasoned (Tsarkov et al., 2006). From these defined classes, an automated classification enables to deduce that only PathologieDeveloppementale and PathologieDegenerative pathologies are progressive.

During the discussion with the domain expert, the importance of describing the type of structure was considered. The expert gave examples of structures that could be involved, such as auditory, visual, phonatory, linguistic, deglutition. Eventually the expert decided to discard this detail because knowing which structure does not affect the property. If it is necessary to model them later, it will be possible to reuse the existing resources in the domain. We also decided to focus on language as a means to assess and consider communication. Recent legal documentation in France has explicitly included communication as a specific purpose for SLT, however as linguists, we considered here that by using the designation “language”, we were referring to the general skills of verbal communication between humans. While examining the possible ways to qualify the pathology, we added the psychoemotional notion to Motor and Neuropsychological issues. Then the concepts **Nature** and **Cible** used to qualify the pathology in SLT are defined as enumerated types:

- **Nature** = {fonctionnemeMoteur, fonctionnemeNeuropsychologique, fonctionnemePsychoAffectif, fonctionnemeSysteme, structureMotrice, structureNeuropsychologique, structurePsychoAffectif, structureSysteme}

- **Cible** = {communication, fonctionOromyofaciale, langageEcrit, langageOral}

---

\(^{11}\) A defined class has a set of necessary and sufficient restrictions and is defined by equivalent statement in OWL.

\(^{12}\) [http://protege.stanford.edu/](http://protege.stanford.edu/)
The object properties created for characterizing the target Cible, the impact of the pathology Nature are defined as follow:

- \( aPourCible = (\text{domain: PathologieOrthophonique, range: Cible}) \).

  The concept \( \text{Cible} = \{\text{communication, fonctionOromyofaciale, langageEcrit, langageOral}\} \) is defined as an enumerated type.

- \( \text{affecte} = (\text{domain: PathologieOrthophonique, range: Nature}) \).

  The concept \( \text{Nature} = \{\text{fonctionnementMoteur, fonctionnementNeuropsychologique, fonctionnementPsychoAffectif, fonctionnementSysteme, structureMotrice, structureNeuropsychologique, structurePsychoAffectif, structureSysteme}\} \) is defined as an enumerated type.

In short, this method using expert discussion in SLT, ontology and linguistics enables to achieve the transformation of a conceptual model into a structured ontology, which will be used to consider scenarios of use. This ontology is based on clinical practice but the decisions made on concepts and concept designations were motivated by the needs to identify and qualify pathology in SLT. Temporality finds a proper place in this ontology and defines a central attribution of properties. The knowledge provided by this concept will provide a more comprehensive approach to pathology in SLT and will make it possible to reason on terminology at a multilingual level.

5 Perspectives

A corpus of 39 French texts from the ISTEX platform is currently being increased (to the amount of 71). Metadata, automatically and manually extracted, will enable us to align another corpus from the same source but in English. At the same time, we are currently producing a list of 30 to 50 main diagnostic labels (expertly chosen) in French and English, which will also be used to test the ontology. The corpus will help in making sure the terms are aligned using the context and in testing the correspondence between terms and concepts. Further work will contribute to make the ontology more comprehensive. Firstly, we will test its logical validity, and write some SWRL rules in order to increase its potential interoperability. We will also introduce SLTs to the ontology in order to explore its relevance. TemPO will be presented and discussed by a group of experts (mixing SLT researchers and practitioners to NLP researchers and linguists), using a consortium previously gathered in the OrthoCorpus project (2015-2017). Finally, we will write and validate definitions of pathology in SLT in natural language.

Références


