

INDEXAÇÃO VÍDEO E ANOTAÇÃO LINGUÍSTICA DE CRIAÇÕES COREOGRÁFICAS CONTEMPORÂNEAS: PARA UMA ONTOLOGIA MULTIMÉDIA

CARLA MONTEZ FERNANDES
(Instituto Politécnico de Leiria e
Centro de Linguística da Universidade Nova de Lisboa)

RUTE COSTA
(Centro de Linguística da Universidade Nova de Lisboa)

ABSTRACT: This article describes the project *VILA_CCC_Mont*, which purpose is the video indexation and the linguistic annotation of original videos from contemporary dance productions shot on site during their rehearsal phases, towards the creation of multimodal corpora for the study and description of the specific movements, styles and vocabularies that correspond to recent productions of international contemporary choreographers.

The project has started with the analysis of the latest dance pieces by Rui Horta, but it is to be extended to other internationally recognised choreographers we wish to closely accompany and analyse.

It is an interdisciplinary project, for which the initial motivation has been to study the relations between verbal and non-verbal language in mental processes involving creativity. It will be based on a strongly terminological/ontological approach, where the final aim is to create a multimedia knowledge-base and an ontology based upon the linguistic theory of Frames.

KEYWORDS: video indexation; linguistic annotation; contemporary dance productions; multimedia corpora and ontologies; ANVIL software;

1. Introdução

Este artigo resume um projecto que se encontra na sua fase inicial¹. O objectivo central do projecto *VILA_CCC_Mont* é o de indexar e anotar linguisticamente videos originais de ensaios de dança contemporânea, com vista à criação de *corpora* multimodais para o estudo das diferentes gramáti-

¹ À data da publicação deste artigo, o projecto já não se encontra em fase inicial, mas em pleno desenvolvimento da segunda fase, que será objecto de novo artigo a publicar em breve.

cas e vocabulários específicos, correspondentes a recentes produções contemporâneas de coreógrafos portugueses ou estrangeiros.

Trata-se de um projecto interdisciplinar, onde a motivação inicial é a de analisar as relações entre a linguagem verbal e a não-verbal em processos mentais que envolvem criatividade e o objectivo final a criação de uma base de conhecimento multimédia e uma ontologia assente na teoria linguística das *Frames*.

Pretende-se criar uma base de conhecimentos multimédia que possa contribuir para o avanço da investigação em ontologias terminológicas multimédia, partindo do caso da linguagem corporal artística.

A indexação e anotação do movimento cobrem até ao momento actual do projecto duas situações diferenciadas: a comunicação dos coreógrafos com os bailarinos/intérpretes durante a criação de uma peça coreográfica; e a observação da peça acabada por parte dos investigadores, analisando-se os movimentos dos bailarinos em palco (passos, sequências, gestos, etc.), tanto em silêncio como nos momentos em que fazem uso da linguagem verbal.

Para as fases de indexação e anotação dos vídeos, faz-se uso do programa computacional ANVIL², sendo que a base de conhecimento multimédia a criar será desenvolvida através de um editor baseado em *frames*, para inserção de conteúdos e posteriores buscas pelo utilizador final, a implementar de acordo com o modelo teórico original *Octopus*³.

2. Motivação inicial

Não existem ainda hoje em dia, na Europa ou EUA, sistemas ontológicos computacionais de apoio à investigação em linguística e dança que façam uso das tecnologias multimédia para organizar as diferentes linguagens, vocabulários ou estilos dos inúmeros coreógrafos de dança, nem dos discursos críticos proferidos acerca dos mesmos. Por outro lado, não tem havido suficiente interesse por parte dos linguistas em estudar as relações entre a linguagem verbal e a não-verbal no que diz respeito à inseparabilidade entre capacidades corporais como o movimento e a cognição humana.

Os sistemas de notação de dança mais reconhecidos (*LabanNotation*, *Benesh Movement Notation*) não são fáceis de utilizar nem permitem a constituição de ontologias ou estruturas que possam contribuir para um registo acessível e organizado das diferentes linguagens já criadas em dança contemporânea em todo o mundo.

Em Portugal, se a linguística já se afirmou como disciplina incontornável e portão de acesso para a compreensão do cérebro e inteligência humanos, não podemos ainda afirmar que o interesse pelos estudos sobre a lin-

² <http://www.dki.de/~kipp/anvil/description/htm>.

³ Cf. Fernandes, C. 2004 (dissertação de Doutoramento), "Interactions between Words and Images in Lexicography: towards new multimedia dictionaries": Universidade Nova de Lisboa e *Vrije Universiteit Amsterdam*.

guagem corporal tenha acompanhado aquela evolução e, como tal, a investigação científica sobre dança é ainda insipiente.

Com o avanço das novas ciências cognitivas e das tecnologias da informação como pano de fundo, o trabalho em linguística, lexicologia e terminologia tem vindo naturalmente a ser exposto a influências provenientes daquelas áreas. Consequentemente, é também na perspectiva cognitiva que se inscreve a teoria linguística das *Frames*, que pretendemos aplicar ao desenvolvimento da base de conhecimento multimédia por nós proposta.

A corrente teórica de base, à luz da qual adaptaremos o editor informático para construção da base multimédia em vista, será a *Frame Theory*, actualmente com aplicações nos mais diversos projectos de investigação, desde a inteligência artificial (IA) na criação de ontologias, até à semântica cognitiva e à lexicografia e terminologia, na criação de bases de dados conceptuais, por exemplo.

O conceito de *frame* a que nos referimos tem origem na IA, foi adaptado nos anos 90 à linguística por Charles Fillmore⁴ e adoptado em lexicologia computacional por vários autores, como Wierzbicka ou Willy Martin.

Curiosamente, também Susan Foster, investigadora em dança contemporânea na Universidade da Califórnia, Los Angeles, realizou conferências em 2002, sobre análise de movimento, em que o título de uma destas era “Frame Theory”⁵, o que ilustra a forte influência desta corrente sobre as mais diversas áreas do saber ao longo dos últimos anos.

Entre os avanços da linguística cognitiva relevantes para a nossa investigação, e que envolvem o conceito de *frame*, *script* ou *schemata*, foram decisivas as teorias para análise da metáfora que deram origem ao conceito de *embodiment*, de Lakoff e Johnson⁶, também na Califórnia, Universidade de Berkeley. Segundo aqueles autores, “There is no such fully autonomous faculty of reason separate from and independent of bodily capacities such as perception and movement. The evidence supports, instead, an evolutionary view, in which reason uses and grows out of bodily capacities.” (1999: 17)

3. Objectivos

Como objectivo inicial, pretende-se constituir um primeiro *corpus* multimodal (de imagens, som e léxico) a partir de filmagens de ensaios de criações em dança contemporânea, com vista à elaboração de uma base de conhecimento multimédia para análise das relações entre linguagem, movimento e pensamento criativo e subsequente identificação dos diferentes vocabulários e estilos coreográficos contemporâneos. Para tal, será necessário criar um editor computacional para estruturar a ontologia que nos propomos atingir. Foram assim definidos os seguintes objectivos específicos:

⁴ Cf. Projecto **FrameNet**, Berkeley University, USA.

⁵ <http://ideasandsociety.ucr.edu>.

⁶ Cf. Lakoff & Johnson (1980); Johnson, M. (1987); Lakoff & Johnson (1999).

- desenvolver técnicas de anotação e etiquetagem de filmagens vídeo, através da criação de uma base de conhecimentos multimédia em constante desenvolvimento, contribuindo assim para o avanço da investigação em ontologias terminológicas, partindo do caso da linguagem corporal artística;

- identificar e inventariar os termos designativos de movimentos, gestos ou sequências já existentes na comunicação entre coreógrafo e intérpretes durante os ensaios para criação de uma peça;

- criar vocabulário para nomear os movimentos que ainda não tenham designação, construindo assim uma classificação de termos designativos de movimentos, que proporcione ao bailarino/coreógrafo, ao crítico ou ao espectador comum de dança contemporânea, um vocabulário específico acessível a todos;

- criar uma base de conhecimento multimédia para visualização, indexação e anotação do movimento em duas situações diferenciadas: na comunicação dos coreógrafos com os bailarinos/intérpretes durante a criação de uma peça coreográfica e na observação da peça acabada, analisando-se os movimentos dos bailarinos em palco, tanto em silêncio como nos momentos em que fazem uso da linguagem verbal.

A metodologia que nos propusemos seguir para a execução deste projecto engloba diferentes tipos de abordagens, sendo necessário o cruzamento de saberes e competências específicas inerentes a formações diversificadas que julgamos complementares. De um ponto de vista teórico e metodológico, assumimos perspectivas transdisciplinares, sendo nosso intento o cruzamento de diferentes discursos: científicos, tecnológicos e estéticos.

4. Descrição detalhada

Numa primeira fase exploratória, que nos serviu de exemplo e teste da metodologia a seguir, tomámos por base filmagens dos ensaios da mais recente criação coreográfica de Rui Horta, “SetUp” (cujos ensaios acompanhámos desde o início e numa base semanal regular), uma vez que o objectivo comum a todas as peças que formos analisando posteriormente será o de indexar linguística e cognitivamente o movimento do corpo humano, analisado a partir de vídeo. Com o resultado desta análise pretendemos conceber uma base de conhecimento multimédia interactiva que funcione como um recurso para a criação coreográfica, dando um contributo para a criação e a fixação de um vocabulário da área de especialidade, essencial na elaboração de discursos críticos e estéticos.

Para a indexação do movimento a partir de vídeo recorreremos ao programa ANVIL (<http://www.dki.de/~kipp/anvil/description/html>). Este programa permite a balizagem do movimento a partir da imagem captada em vídeo, sendo que um movimento corresponde a uma sequência identificada entre dois momentos de pausa.

Antes de proceder à descrição e análise do movimento identificado, é necessário encontrar formas de anotação – *metadados* – que nos permitem

anotar de forma harmonizada “blocos” de movimento, fase em que se encontra actualmente o projecto. Esta anotação terá uma dupla função: a primeira dessas funções é didáctica; a segunda é do foro da engenharia do conhecimento e relaciona-se com a estruturação do conhecimento que possibilita o tratamento automático dos dados: imagens (vídeos e fotos) vocabulários, textos e esquemas.

De seguida, será necessário classificar os movimentos corporais, tendo por base categorias conceptuais que se realizam em discurso por via de vocabulários que exprimam os movimentos do corpo humano e dêem conta das emoções a eles associadas⁷.

Após a identificação das classificações existentes para análise do movimento, pretendemos, a partir da observação dos vídeos, construir um sistema organizativo aberto e dinâmico, em constante possibilidade de actualização, que estará na base de uma ontologia e que estrutura a base de conhecimentos multimédia.

Por outro lado, a existência de termos aceites no seio de uma comunidade profissional é imprescindível para organizar um sistema conceptual sobre o qual possa assentar um sistema computacional como pilar da base de conhecimentos multimédia. Consequentemente, para a implementação de sistemas de tratamento da informação terminológica é necessário proceder à identificação e ao estabelecimento organizado, sistematizado e hierarquizado do conhecimento, recorrendo para o efeito a sistemas de representação previamente formalizados que permitam dar conta das estruturas do conhecimento inerentes a um ou vários domínios. As ontologias, artefactos teóricos e computacionais que explicitam conceptualizações, permitem a representação do conhecimento inerente a um grupo no seio de uma classe profissional. Para Guarino (2006), uma ontologia é uma forma de expressar “the intended meaning of a vocabulary in terms of the nature and structure of the entities he refers to”. Do ponto de vista da engenharia do conhecimento, a conceptualização é entendida como uma estrutura formal “of a piece of reality as perceived and organized by an agent, independently of the vocabulary used and the actual occurrence of a specific situation (Guarino, 2006).

A estrutura do conhecimento hierarquizado está directamente ligada à estruturação do vocabulário necessária para a construção não ambígua de discursos. Desta forma, e em paralelo, iremos conceber uma base de dados terminológica que dará conta do vocabulário de especialidade e funcionará como um módulo no seio da base de conhecimento multimédia. Esta base irá pôr em prática o modelo *Octopus* referido anteriormente, que tem no seu centro *frames*⁸ conceptuais para definição dos vários tipos de movimentos e oito “tentáculos” (galerias de acesso) constituídos pelas diversas modalida-

⁷ A título de exemplo, segundo Susan Foster (1986), devem ser tidos em conta cinco itens para análise de uma peça de dança: *contextos* (existência ou não de narrativa, espaço performativo, etc), *modos de representação*, *estilos*, *vocabulários* e *sintaxes*.

⁸ Ilustradas na figura 1.

des de representação de conhecimento para ilustração dos conceitos introduzidos.

Será a partir de estruturas semelhantes à ilustrada seguidamente que nos propomos classificar e, logo, categorizar, os diversos tipos e sub-tipos de movimentos, respectivas qualidades, dinâmicas, intenções, etc, que nos permitirão criar um sistema ontológico que funcione como futura base de trabalho *online*, motivando e gerando novos conteúdos a favor de uma abrangente investigação em dança contemporânea e, conseqüentemente, de uma nova forma de registo e documentação da mesma.

SLOTS	FILLERS: item: <i>stumble</i>
Sub-type	MOVE2 (move_in_space)
Agent	human
Means:	
body_part	foot
tool	—
Path	forward + down/over
Aim	—
Manner	hit smth. with foot / put foot down awkwardly / walk unsteadily
Setting	smb. walking and stumbling over an obstacle
Result	(almost) fall / get hurt
Intensity (degree of force/speed/...)	strong
Cause	unnoticed object on the way

Figura 1: uma das *frames* utilizadas para descrever a categoria de “verbos de movimento que implicam deslocação espacial” (aqui designada por MOVE2) em tese de doutoramento dedicada à relação entre palavras e imagens em lexicografia⁹

Na figura 2 apresentamos um esquema simplificado do modelo *Octopus* a implementar neste projecto. É importante notar que se partirá de uma base de dados bilingue no que se refere aos termos de entrada, que poderão ser consultados quer a partir do Português quer do Inglês, sendo naturalmente possível a extensão a outras línguas, se tal vier a demonstrar-se interessante ou necessário por parte de coreógrafos e/ou investigadores que conosco venham a colaborar. Esta estrutura modular é possível, uma vez que os seus diversos componentes permitem buscas independentes a activar de acordo com o interesse e motivação do utilizador.

⁹ In Fernandes, C. 2004. *Interactions between words and images in lexicography: towards new multimedia dictionaries* Vrije Universiteit Amsterdam e FCSH da Universidade Nova de Lisboa.

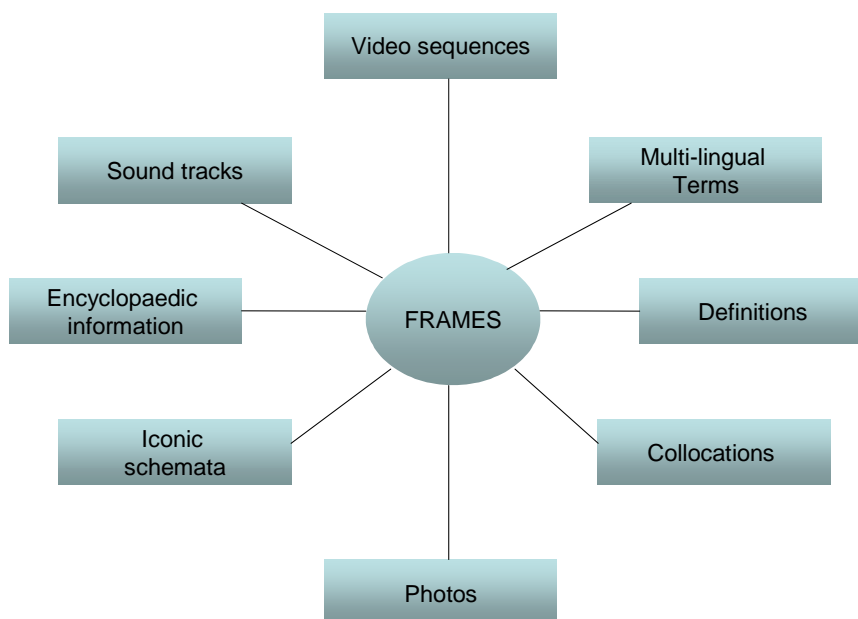


Figura 2: Modelo simplificado do “Octopus Model”. In Fernandes, C. 2004.

5. Nota final

A constituição da primeira terminologia estruturada partirá sempre de discussões com os coreógrafos e consultores participantes no projecto quanto à sua aceitabilidade dos termos, tendo por base critérios conceptuais e linguísticos. Para a proposta de criação de termos, seguiremos prioritariamente uma metodologia onomasiológica que nos permitirá identificar o conceito. A selecção e/ou criação de termos adequados à denominação do conceito baseia-se em critérios morfossintácticos.

Prevê-se que o projecto tenha uma duração mínima de 3 anos, sendo que as diversas tarefas serão executadas por fases desde o corte das imagens vídeo – tendo em conta as diversas secções coreográficas, para análise de sequências delimitadas – até à criação de um *site* público em constante actualização para divulgação dos resultados e troca de aprendizagens/experiências. Pretende-se que o público interessado em dança contemporânea, bem como profissionais da área, possam contribuir também com as suas opiniões ou sugestões, naturalmente a validar pela equipa de investigação.

Sendo o nosso objecto inicial de estudo as criações coreográficas em dança contemporânea, tal não significa que a metodologia a desenvolver neste plano de trabalho apenas seja aplicável àquele domínio específico. De facto, esta área artística servir-nos-á como estudo de caso, podendo o modelo teórico desenvolvido vir a ser aplicado, no futuro, a trabalhos de análise do gesto em conversação, por exemplo, ou a qualquer outra área de investigação

onde a criação de sistemas terminológicos estruturados se revele fundamental para uma melhor compreensão do pensamento e cultura do ser humano.

Referências

- Barsalou, L. 1991. Constructing categories to achieve goals. In G. Bower (ed.) *The Psychology of learning and motivation: Advances in research and theory*. vol. 27. San Diego: Academic Press.
- Costa, R. 2005. Terminology, Corpus Linguistics and Ontology, TKE – *7th International Conference on Terminology and Knowledge Engineering. Terminology & Content Development*. Copenhagen. <http://www.GTW.org> [Invited Keynote Speaker].
- Croft, W. & Cruse, D. 2003. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Cunningham, M. 1988. *Le Danseur et la danse*. Paris: Pierre Belfond.
- Damásio, A. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. London: Heinemann.
- Fernandes, C. 2004. *Interactions between Words and Images in Lexicography: towards new multimedia dictionaries*. PhD dissertation, Universidade Nova de Lisboa, FCSH e Vrije Universiteit Amsterdam.
- Frétard, D. 2004. *Danse Contemporaine – Danse et non-Danse: vingt-cinq ans d'histoires*. Paris: Éditions Cercle d'art.
- Foster, S. 1986. *Reading Dancing, Bodies and Subjects in Contemporary American Dance*, Berkeley & London: University of California Press.
- Ginot, I. & Michel, M. 2002. *La Danse au XX siècle*. Paris: Larousse.
- Guarino, N. 2006. Ontology and Terminology. How can formal ontology help concept modelling and terminology. Powerpoint in Workshop «Terminology, Concept Modelling and Ontology. Which approach for which problem?». EAFT – NordTerm. Vaasa: 10th February.
- Johnson, M. 1987. *The body in the mind*. Chicago: Chicago University Press.
- John-Steiner, V. 1997. *Notebooks of the mind*, Oxford / New York: Oxford University Press.
- Laban, R. 1956. *Principles of Dance and Movement Notation*, London.
- Lakoff & Johnson 1980. *Metaphors we live by*. Chicago: The University of Chicago Press.
- Lakoff & Johnson 1999. *Philosophy in the flesh*. Basic Books.
- Lamb, S. 1999. *Pathways of the brain – The Neurocognitive Basis of Language*. Amsterdam: John Benjamins.
- Martin, W. 2001. A Frame-based approach to polysemy”. In *Polysemy in Cognitive Linguistics*. Cuyckens and Zawada (eds), Amsterdam: John Benjamins.
- Minsky, M. 1975. A framework for representing knowledge. In P.H. Winston (ed.) *The Psychology of computer vision*. New York, pp. 211-277.
- Ribeiro, A.P. & Sasportes, J. 1991. *História da Dança*. Lisboa: Imprensa Nacional Casa da Moeda.
- Ribeiro, A.P. 1997. *Corpo a Corpo: Possibilidade e Limites da Crítica*. Lisboa: Edições Cosmos.