



HAL
open science

A Dictionary Database of Contemporary English

Véronique Abasq, Jean-Michel Fournier, Pierre Fournier, Marjolaine Sophie
Martin

► **To cite this version:**

Véronique Abasq, Jean-Michel Fournier, Pierre Fournier, Marjolaine Sophie Martin. A Dictionary Database of Contemporary English. PAC workshop 2009, Sep 2009, Aix en Provence, France. halshs-04608641

HAL Id: halshs-04608641

<https://shs.hal.science/halshs-04608641>

Submitted on 11 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Dictionary Database of Contemporary English



Véronique Abasq, Jean-Michel Fournier, Pierre Fournier, Marjolaine Martin
Laboratoire Ligérien de Linguistique (E.A. 3850), Université de Tours



CONTEXT AND AIMS

- Lionel Guierre's work on D. Jones's *English Pronouncing Dictionary* (12th edition, 1964) in the 70's: graphophonological rules and stress rules (morphophonology).
 - Work on J. C. Wells' *Longman Pronunciation Dictionary* (1st edition, 1990) by Guierre's followers: A. Deschamps, J.-L. Duchet, J.-M. Fournier, I. Trevian
- ➔ Define the parameters that underlie the morphophonological behaviour of English speakers and analyze their workings, with and without reference to the written form
- ➔ Build an up-to-date and extensive database of the pronunciation of lexical units in 3 reference varieties with all necessary information

DATABASES

Samples:

Guierre's original corpus: UNINTERRUPTED ;4N/,in/T6;/R4P/TID Q KT 21020

Guierre's 1990 corpus: uninterrupted fQKT "Vn°°Int#@rVpt...IdY Ç-@dÇ !-°In«}»@-πi ~ly Πlimi ~ness Πn@s βnlsβπ5.23010\detpurretninu

LPD 2008 (xml): <p:Entry id="p067-000415924"><p:Head id="p067-000834624"><p:HWD id="p067-000834625">uninterrupted</p:HWD></p:Head><p:PronCodes id="p067-000834626"><p:PRON id="p067-000834627"> ,ʌn ,ɪnt ə 'rʌpt id</p:PRON><p:VARPRON id="p067-000834628">-əd</p:VARPRON><p:AMEPRON id="p067-000834629"><p:TEXT id="p067-000834630">- ,ɪn</p:TEXT><p:i id="p067-000834631">ʔ</p:i><p:TEXT id="p067-000834632"> ə</p:TEXT></p:AMEPRON></p:PronCodes><p:Derivative id="p067-000834633"><p:DERIV id="p067-000834634">~ly</p:DERIV><p:PronCodes id="p067-000834635"><p:PRON id="p067-000834636">li</p:PRON></p:PronCodes></p:Derivative><p:Derivative id="p067-000834637"><p:DERIV id="p067-000834638">~ness</p:DERIV><p:PronCodes id="p067-000834639"><p:PRON id="p067-000834640">nəs</p:PRON><p:VARPRON id="p067-000834641">nɪs</p:VARPRON></p:PronCodes></p:Derivative></p:Entry>

Databases content:

| | Spelling | British pronunciation (EPD/LPD) | Stress pattern (EPD/LPD) | Variants (EPD/LPD) | American pronunciation (EPD/LPD) | Australian pronunciation (Macquarie) | Syntactical category | Lexical status | Morphology | Meaning | Frequency | Oral corpora check |
|--------------|----------------|---------------------------------|--------------------------|--------------------------|--------------------------------------------------------------|--------------------------------------------------------------|--------------------------------------------------------|----------------------------------------------|---------------------------------------------------------------|-----------------------------------------------|--------------------------|--------------------------|
| Guierre 70's | uninterrupted | ;4N/,in/T6;/R4P/TID | 21020 | Variant type (coded): KT | | | Q | | | | | |
| Guierre 1990 | uninterrupted | "Vn°°Int#@rVpt...IdY | 23010 | Ç-@dÇ | !-°In«}»@- | | Q | | | | | |
| LLL corpus | un interrupted | ,ʌn ,ɪnt ə 'rʌpt id | 22010 | Expanded variants | ,ʌn ,ɪnt ə 'rʌpt id (+ stress pattern and expanded variants) | ,ʌn ,ɪnt ə 'rʌpt əd (+ stress pattern and expanded variants) | Adj. ('multicategorical' when diverging pronunciation) | interrupted ('deriving form' where relevant) | un: prefix inter: prefix rupt: bound root ed: suffix | learned / current (definition where relevant) | (from available corpora) | (from available corpora) |

METHODOLOGY

- Agreements with C.U.P., Pearson and Macquarie to make use of their xml databases of *Cambridge English Pronouncing Dictionary* (17th ed.), *Longman Pronunciation Dictionary* (3rd ed.) and *Macquarie Dictionary* (4th ed.)
- Expansion of their data with morphological and lexical analyses, data from frequency corpora, and data from oral corpora when motivated by a high degree of variation (2 years)
- Systematic analyses based on the resulting database (probably no less than 3 years)

A FEW EXAMPLES OF PLANNED RESEARCH

- Check the validity of established rules for stress placement – uncover rules for unaccounted parts of the vocabulary.
- Systematic analysis of prefixes, roots and suffixes with relevance on vowels pronunciation, stress placement, and morphological structure of syntactical categories.
- In partnership with the FORELL-A laboratory of the University of Poitiers: phonological evolutions since the 17th century.
- In cases such as structures in *un-* where dictionaries mention a number of differing stress patterns, and whose morphology implies the variation might *not* be lexical, resort to oral sources to try and identify possible discourse parameters.

REFERENCES

- Deschamps, A., Duchet, J.-L., Fournier, J.-M., O'Neil, M., 2000, *English Phonology and Graphophonemics*, Paris : Ophrys.
- Duchet, J.-L., 1994, *Code de l'anglais oral*, 2^{ème} édition, Paris : Ophrys.
- Fournier, J.-M., 2009, *Manuel d'anglais oral*, Paris : Ophrys.
- Guierre, L., 1984, *Drills in English Stress-Patterns*, 4th ed., London: A. Colin - Longman.
- Trevian, I., 2003, *Morphoaccentologie et processus d'affixation en anglais*, Bern : Peter Lang.