



HAL
open science

A Dictionary Database of Contemporary English: when the Tool Meets the Needs The Particular Case of Word Stress Assignment in Disyllabic Verbs

Marjolaine Sophie Martin, Pierre Fournier, Sophie Vanhoutte, Élodie Descloux

► To cite this version:

Marjolaine Sophie Martin, Pierre Fournier, Sophie Vanhoutte, Élodie Descloux. A Dictionary Database of Contemporary English: when the Tool Meets the Needs The Particular Case of Word Stress Assignment in Disyllabic Verbs. MFM19 Manchester Phonology Meeting 2011, May 2011, Manchester, United Kingdom. hal-04608585

HAL Id: hal-04608585

<https://hal.science/hal-04608585>

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: When the Tool Meets the Needs

The Particular Case of Word Stress Assignment in Disyllabic Verbs

Elodie Descloux, Pierre Fournier, Marjolaine Martin, Sophie Vanhoutte
Laboratoire Ligérien de Linguistique (E.A. 3850), Université de Tours

CONTEXT AND AIMS

- Use an up-to-date and extensive database of the pronunciation of lexical units in Southern British English, General American English and Australian English. It includes for each dictionary entry its spelling, its pronunciations and stress patterns in the three varieties (including variants), its category/ies, its morphology, its meaning, its lexical status.
- Background : Lionel Guierre's work on D. Jones's *English Pronouncing Dictionary* (12th edition, 1964) in the 70's (graphophonological rules and stress rules) and his followers: A. Deschamps, J.-L. Duchet, J.-M. Fournier, I. Trevian working on J. C. Wells' *Longman Pronunciation Dictionary* (1st edition, 1990).
- Check the validity of established rules for stress placement – Uncover rules for unaccounted parts of the vocabulary.

EXAMPLE OF A STUDY: STRESS ASSIGNMENT IN DISYLLABIC VERBS

Specialized works on English stress based on syllable weight (Roach 1983, Giegerich 1992 or Carr 1999) hold that disyllabic verbs are late stressed. We led a study using the database to check whether it is the case. This specific tool enabled us to base our analysis on a large corpus and to ascertain that 53% of disyllabic verbs are stressed in /-1/ when 47% are stressed in /1-/. We then tested a morphological hypothesis which we found more efficient. We also took into account the usage frequency of these verbs using the COCAE. Our study is based on primary variants only for the three varieties mentioned.

Data cleaning :

First extraction: 3 230 words

✂ Other categories: ± 190 words, mainly nouns from noun/verb pairs

✂ Inflections: ppl, pa, ppl... ± 160 words

⇒ Missing words in 1 or 2 sources

(EPD/LPD/MCQ): ± 140 words

✂ Phonological versus phonetic syllables:

syllabic /l/ and /r/: ± 240 words

✂ Compression: ± 50 words

⇒ Final corpus: **2550 disyllabic verbs**

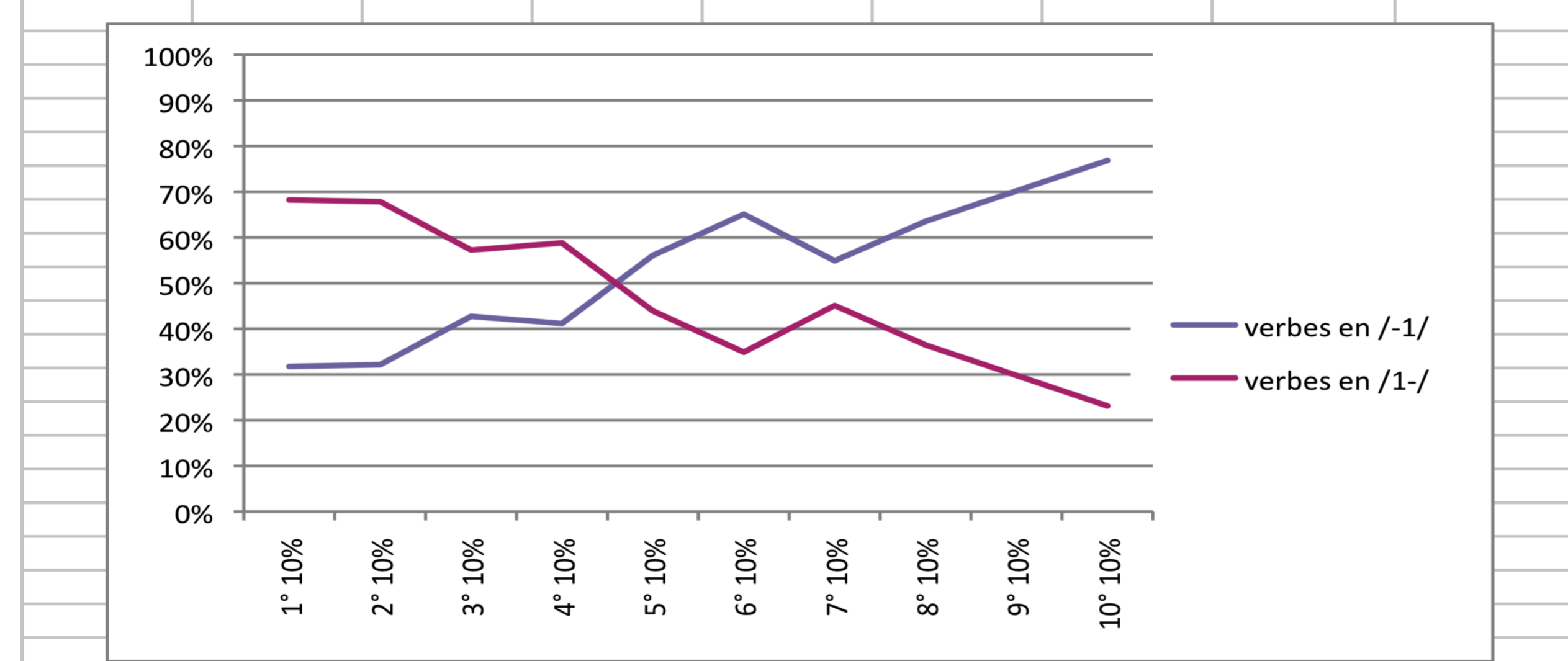
with an **outstanding data reliability**

(differences between EPD/LPD = 32 verbs out of 2550)

Spelling	EPD SBE	EPD GAm	LPD SBE	LPD GAm	MCQ SAusE	Overall Frequency	Spoken Frequency	Deriving Form	Morphology	Category	Meaning
divorce	01	01	01	01	01	6,304	1,345	no	préfixé	VN	divorcer
follow	10	10	10	10	10	124,780	15,915	no	base	V	suivre
hand-pick	21	21	21	21	01	241	34	no	composé	V	cueillir
injure	10	10	10	10	10	7,864	2,452	no	préfixé	V	blessier
salute	01	01	01	01	01	1,837	325	no	base	V	saluer

FREQUENCY

Freq.	/-1/	/1-1/	total	/-1/	/1-1/	COCA min	COCA max
1° 10%	81	174	255	32%	68%	0	0
2° 10%	82	173	255	32%	68%	0	3
3° 10%	109	146	255	43%	57%	3	35
4° 10%	105	150	255	41%	59%	35	134
5° 10%	143	112	255	56%	44%	134	346
6° 10%	166	89	255	65%	35%	346	701
7° 10%	140	115	255	55%	45%	702	1 358
8° 10%	162	93	255	64%	36%	1 364	2 928
9° 10%	179	76	255	70%	30%	2 938	9 557
10° 10%	196	59	255	77%	23%	9 563	269 749



- **47% stressed /10/ Vs 53% stressed /01/**
- a tendency towards /01/ (70%Vs 30%) can be observed for the 20% most frequent verbs.

MORPHOLOGY

	/1-1/		/-1/		Total
	nbr	%	nbr	%	
Suffixed	177	74%	63	26%	240
<i>derivatives</i>	79	95%	4	5%	83
<i>not derived</i>	98	62%	59	38%	157
Compounds	245	85%	44	15%	289
Prefixed	92	7%	1171	93%	1263
Bases	673	89%	85	11%	758
Total	1187	47%	1363	53%	2550

- non prefixed verbs are mainly in /10/ : suffixed verbs (74%), compounds (85%) and bases (89%)
- the vast majority of **late stressed verbs** are **prefixed verbs** (86%)
- the morphological analysis is the key factor of stress assignment here : **prefixed in /01/ 93%, non-prefixed in /10/ (92%)**

REFERENCES

- Carr, P. (1999), *English Phonetics and Phonology: an Introduction*, London: Blackwell.
- Chomsky, N., Halle, M., (1968), *The Sound Pattern of English*, New-York: Harper & Row.
- Descloux E., Fournier, J.-M., Fournier, P., Girard, I., Martin, M., (2010) "Structure, Variations, Usage & Corpora: the Case of Word Stress Assignment in Disyllabic Verbs", PAC Workshop 2010 : Structure, Variation, Usage and Corpora, Université Paul Valéry – Montpellier III.
- Giegerich, H.J. (1992), *English Phonology. An Introduction*, Cambridge: CUP.
- Guierre, L., (1984), *Drills in English Stress-Patterns*, 4th ed., Londres: A. Colin - Longman.
- Roach, P., (1983), *English Phonetics and Phonology, a Practical Course*, Cambridge: CUP.
- Jones, D., (2006), *Cambridge English Pronouncing Dictionary*, 17th edition, edited by P. Roach, J. Setter & J. Hartman, Cambridge: CUP.
- Macquarie Dictionary*, (2005), 4th edition, Sydney: Macquarie Dictionary Publishers Pty Ltd.
- Wells, J.C., (2008), *Pronunciation Dictionary*, 3rd edition, London: Longman.
- Corpus Of Contemporary American English, Mark Davies: <http://www.americancorpus.org>.