

Ton Broeders and Phil Hyams

THE PRONUNCIATION COMPONENT OF AN ENGLISH-DUTCH DICTIONARY

Introduction

Although this is perhaps not always fully realized, any thinking about the design of the pronunciation component of a dictionary should primarily consider the needs of the user. In the case of our dictionary, the forthcoming GROOT WOORDENBOEK ENGELS - NEDERLANDS (compiled by Martin and Tops for van Dale), the user and his needs are reasonably well-defined. To those who are strangers to the Dutch dictionary-making scene, let us first point out that to the average linguistically naive Dutchman van Dale's GROOT WOORDENBOEK DER NEDERLANDSE TAAL which runs to some 3,000 pages in two volumes is the ultimate arbiter of what is or is not correct Dutch, and as such could be said to be comparable in status, if not in scope, to the SHORTER OXFORD DICTIONARY in Britain, or WEBSTER'S THIRD (W3) in the United States. The new English-Dutch dictionary is to form part of a series of six new bilingual dictionaries, French-Dutch, Dutch-French, German-Dutch, Dutch-German, English-Dutch and Dutch-English, which, together with a new dictionary of contemporary Dutch, are designed to become as authoritative in their fields as their prestigious namesake. The dictionaries will each run to about 1500 pages and will clearly be the most exhaustive general bilingual dictionaries available for the educated speaker of Dutch.

Of course, the Dutch market is a small one. There is, however, a comparatively substantial demand for bilingual dictionaries thanks to the prominent position of the teaching of English, German and, to a lesser extent, French in Dutch secondary education. When we try to envisage the Dutch user and his needs, we find that the position of English and the knowledge of English in the Netherlands is clearly superior to that of any other foreign language, including Dutch, some critics might say. English is a compulsory subject in all types of secondary schools and it will soon be made compulsory in the last two years of primary schools. In a sense, it could be argued that English has been a compulsory subject for much longer, at least for those under 30. From their earliest days, the TV-viewers among them have been exposed day in, day out, to - probably predominantly American - English sounds, even before they could read the subtitles. On Dutch television dubbing is, fortunately, rare. As a result, since the vast majority of programmes are imported from abroad and since the vast majority of the imports are English-speaking programmes, most Dutch people are now continually exposed to English. Of course, this situation has important implications for the dictionary maker. Today, the sound of English is more familiar to the average Dutchman than that of some of the more prominent regional varieties of his own language.

So much for the dictionary user's English-language background. It is clear that in view of this the level of presentation of the pronunciation component of an English-Dutch dictionary will have to be fairly high. In fact, we would maintain that there are three

specific requirements to be met:

- (1) both the Standard Southern British pronunciation (also called RP) and a Standard American accent should be represented;
- (2) the phenomenon of stress shift should be properly dealt with,
- (3) the marking of stress in compounds and idiomatic expressions should be adequate.

It is to a discussion of these three points that the rest of this paper is devoted.

Choice of accent and transcription system

One reason why RP should certainly be included is that an RP-type accent is still held up as a model for most Dutch student teachers to imitate, even though one cannot always help feeling that young people are finding this type of accent increasingly difficult to identify with. At the same time, in view of the increasing familiarity of the potential user with American-style pronunciations, particularly in songs and on television, and their great prestige among Dutch speakers, despite or perhaps because of the strong pressure towards RP in English classes, General American (GA), the accent that is usually used in teaching American English to foreigners, would seem to have an almost equally strong claim. Our decision has been to represent both accents.

Another decision concerns the notation. Broadly, there are three possibilities: a re-spelling system, a phonemic, or an allophonic transcription. Gimson (1973) clearly prefers the second solution:

My own prejudices lead me to favor a notational system based on the symbols of the IPA in their more conventional shape - as indeed they are used in so many foreign language dictionaries. The great advantage of this system lies in its simplicity and the possibility of consistency in the sound/symbol relationship.

Re-spelling would obviously only be at all helpful if it were done in terms of the spelling conventions of the native language, since we cannot assume our Dutch user to be sufficiently familiar with the English sound/spelling symbol relationships. Re-spelling would therefore have the obvious disadvantage of encouraging phonological interference, and especially phoneme substitution, i.e. the tendency to replace the phonemes of the target language with those of the native language. It would thus take the naive speaker as the norm, which would be extremely irritating to the more advanced speaker. In addition, as pointed out by Gimson, re-spelling cannot usually be done without the use of numerous diacritics, and it leads to different conventions being employed by different dictionaries in their efforts to compensate for orthographic inconsistency.

The potential power of an allophonic transcription in a bilingual dictionary is clear. It would be possible to explicitly indicate those features which the target accent does not have in common with the native accent. Examples could be a superscript h for fortis plosives, which are unaspirated in Dutch, or the use of length-marks to indicate allophonic duration of vowels and/or nasals and /l/. Here are some examples:

	phonemic	allophonic		phonemic	allophonic
<u>pence</u>	/pens/	[p ^h ens]	<u>cease</u>	/si:s/	[sis]
<u>pens</u>	/penz/	[p ^h en:z]	<u>seize</u>	/si:z/	[si:z]
<u>spend</u>	/spend/	[spen:d]	<u>route</u>	/ru:t/	[rut]
<u>spent</u>	/spent/	[spent]	<u>rude</u>	/ru:d/	[ru:d]

However, any degree of - necessarily incomplete - allophonic detail would only be obtained at a high cost: the reader's intuitive knowledge of the phoneme is ignored and his willingness to aim at a near-native pronunciation taken for granted. Yet, even when this is the foreigner's professed aim, he will often subconsciously aim at a covert norm, i.e. English with a Dutch accent.

On balance, then, the choice would clearly be to use an IPA-type phonemic notation, because it achieves consistency with a minimum number of symbols, is used in the best learner's dictionaries like the LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH (LDOCE) and the OXFORD ADVANCED LEARNER'S DICTIONARY OF CURRENT ENGLISH (ALD), does not encourage phonological interference and relies on the intuitive notion of the phoneme as sound-unit. In this connection, it is worth noting that the Netherlands are perhaps unique in that all Dutch teachers of English have a working knowledge of phonetic transcription and that IPA-based transcriptions are in fact quite regularly used in secondary-school textbooks. Our task, therefore, was to devise a set of notational conventions which would allow us to represent RP as well as GA pronunciations, while achieving the greatest possible economy of symbols. At an early stage we also decided that rather than giving the citation-form of a particular word, the user would perhaps be better served with a representation of the way words are pronounced in what Gimson (1980:297) calls a "careful colloquial style". This implies the listing of weak forms before the (rarer) strong forms and the use of round brackets to indicate phonemes which are likely to be elided in any but the more formal styles, e.g. -elision before /r/ and /l/, as in memory, bachelor, /t/ and /d/ elision in postman, landscape, and indicating t-voicing and flapping, and/or /t/ elision in GA as in meeting, twenty. A third question concerns the number of variants to be given for each entry. Since a bilingual dictionary should not seek or pretend to be exhaustive and since we were not writing a pronouncing dictionary, a purely descriptive approach, listing all acceptable pronunciations, would clearly not be appropriate. Instead, we have chosen to give variants only when it was clear that they were roughly equally common. In those cases where the more common American variant also occurs as a less common Southern British alternative or vice versa, this is not explicitly indicated. Thus, address is /ə'dres || 'ædres / ; harass is /'hærəs || hə'ræs/ ; either is /'aɪðə || 'i:ðə / , and advertisement /əd'vɜ:ti:smənt || 'ædvərtaizmənt/. The following symbols are used for RP and GA. Symbols in brackets are not used for GA.

<u>vowels</u>		<u>consonants</u>				<u>marginal phonemes</u>		
ɪ	i:	əɪ	p	f	m	!	x	
e	u:	eɪ	t	θ	n	r	ǣ	ð
æ	ɑ:(:)	ɔɪ	k	s	ŋ	j		
(ɒ)	ɔ:(:)	əʊ	b	ʃ		w		
ʊ	(ɜ:)	oo	d	v				
ʌ		ɪə	g	ʒ				
ə		eə		z				
i		ʊə		ʒ				
				h				

There are four special conventions in our notation:

- ˘ for T-flapping and voicing, as in /[˘]mi:tɪŋ, [˘]bɑ:tɪ/
- nt̘ to indicate nasal flap or t-deletion, as in /əd[˘]væŋt̘ɪdʒ, [˘]twentɪ/
- i for {i:} as in /[˘]hæpi, [˘]hæpiɪəs, ri[˘]ækʃn/
- ɜ for {RP predominantly ɪ, but also ə} as in /[˘]kɪsɜz, [˘]mendɜd /

Following Wells (1982: 72-85), on the segmental level, the RP and GA accents differ in the four following respects:

a) phoneme system:

- RP /ɪ,ə,eə,ʊə/ do not occur; GA has /ɪr,er , ʊr /
- GA has /ɹ/ and /ɹ̄/, represented as /ɜr/ and /əɹ/, i.e. further is transcribed /[˘]fɜ:ðə || [˘]fɜrðəɹ /
- the RP opposition /ɒ/ -/ɑ: /, as in bomb - balm, does not occur.

b) phonotactic distribution:

- GA is rhotic or r-pronouncing, e.g. far /fɑ: || far /
- GA typically has /hw/ in wh-words, e.g. white /waɪt || hwɑɪt /
- GA frequently has /j/ where RP has /hj/, huge /hju:dʒ || (h)ju:dʒ /
- GA has /u/ where RP has /ju: / after /θ, t, d, n, l/ in accented syllables

c) lexical incidence or lexical distribution:

- the distribution of RP / æ, ɑ:, ɒ, ɔ: / versus GA /æ, ɑ, ɔ /.

RP	GA	Examples
æ	æ	gather, map lack, lamb
ɑ:		rather, half ask, dance
ɒ	ɑ	father, calm balm, lava
		mock, fodder bomb, collar
ɔ:	ɔ	forest, gong coffee, gone call, cause caught, broad

Table based on Gus-
senhoven and Broeders
(1976)

- RP/-ɪd, -ɪz/versus GA /-əd, -əz/ in grammatical endings
- words ending in -ile, -ary, -ery, -ory, -mony, -ative, -ization, -ersion and -ursion
- RP/ɪ/ versus GA/i:/ in words like happy, coffee
- wordstress in recent French loanwords like cliché, attaché

d) phonetic realization:

Since our notation is phonemic and because words whose phonemic make-up is the same in RP as in GA are given only one transcription, differences in the realization of corresponding phonemes cannot be indicated. Inevitably, the use of identical symbols may lead the reader to think that the symbols stand for the same (phonetic) sound. A warning to this effect will be included in the front matter of the dictionary, as will a short discussion of the main differences in the realization of RP and GA phonemes, and a description of the main differences between these and their Dutch counterparts. A major realization difference is that in words like goal, home. In view of the considerable qualitative difference a compromise symbol is used: **ɔv**. It has the additional advantage of striking the reader as less exotic and of being easier to interpret than the for RP more accurate **əv**.

A number of our conventions could be said to be pedagogically motivated in that they are expected to prevent certain errors. Examples are the retention of /t,d/ in words like bench, inch, change to avoid [bɛ̃tʃ, ɪ̃tʃ, tʃeɪ̃dʒ], forms involving an undesirable Dutch palatalization; the bracketing of /t,d/ in words like postman, landscape, which may prevent the production of hypercorrect forms,

and the retention of length-marks in GA /i:/, u:/, to bring out their greater length than that of the corresponding (short) Dutch /i/ and /u/. Finally, there is the use of high accent marks to indicate secondary stresses preceding the primary stress, which may promote the correct positioning and realization of pre-nuclear pitch-prominent accents as in as'simi'lation, de'mocrati'zation, where Dutch speakers of English tend to have a non-pitch prominent rhythmic stress on the first syllable.

Stress shift

The second innovation in the dictionary will be the indication of 'stress shift', also known as 'stress retraction' or 'iambic reversal' (cf. Liberman and Prince 1977; Thompson 1980), and to some as the 'fifteen-men rule' (Chomsky and Halle 1968). What all these terms refer to is the tendency that can be observed in polysyllabic English words for the main stress to shift to an earlier syllable when the word itself is followed by another stress, as in 'fifteen 'men as opposed to fif'teen. Unfortunately, in the dictionaries we have consulted which deal with this phenomenon at all, the impression is created that stress shift is associated with individual words or compounds. Thus, according to LDOCE, stress shift is a feature of "a number of compounds" (p. xviii), which "can also happen with some single words..." (p. xix). This may explain why for example absolutely and altogether are marked for stress shift, while already is not.

In the ENGLISH PRONOUNCING DICTIONARY (EPD), the vast majority of the stress-shifted variants that are given occur with compound adjectives of the hard-headed type. ALD does not even mention, let alone indicate stress shift, while the CONCISE PRONOUNCING DICTIONARY OF BRITISH AND AMERICAN ENGLISH (CPDBAE), though rightly stating in the front matter that in connected speech any syllable marked for stress may have its stress reduced from principal to subordinate or even suppressed altogether, gives no indication of the phenomenon its editor so admirably describes in the dictionary-entries themselves. Nor is it clear why in CPD words like 'after'noon, 'prin'cess, 'third'rate and 'thir'teen should have been given two accent-marks, while al'ready, cam'paign, i'dea and week'end are given only one. A much more satisfactory treatment is that in W3 where beneficial and campaign are transcribed ;benə;fɪʃəl and ⁽¹⁾kam;pæn. As is implicit in this notation, stress shift is in fact best looked upon as a general phonological rule which affects all polysyllabic nouns, verbs, adjectives and adverbs. However, not all these open-class items are always affected to the same extent. Factors like the prosodic and morphological make-up of the word (cf. Gussenhoven forthcoming) and the style and speed of utterance will all affect the likelihood of application of this essentially variable rule. In view of this, in the new GROOT WOORDENBOEK words like idea, campaign, introduce, already are assigned two high accent marks to indicate that either of these marked syllables can receive pitch prominence, depending on the phonological context:

<u>idea</u>	ˈaɪˈdɪə	<u>introduce</u>	ˈɪntrəˈdjuːs -ˈduːs
<u>campaign</u>	ˈkæmˈpeɪn	<u>already</u>	ˈɔːlˈredɪ

Admittedly, the notation we are proposing is a trifle unorthodox; at the same time, however, it is extremely simple and succeeds in capturing the important generalization that we feel can be made about the prosodic/accentual behaviour of polysyllabic English words. Whenever a word is assigned two high accent-marks it will tend to be subject to stress shift. Here are some examples of stress shifted words in context:

Nouns

'cam'paign <> a 'campaign 'issue
 as'soci'ation <> As'sociation 'Football
 'i'dea <> the 'idea of 'going there
 'week'end <> a 'weekend re'turn

Verbs

'intro'duce <> 'introduce 'Peter
 'repre'sent <> 'represent 'you
 'look 'up <> 'look up the 'word
 'come 'in <> 'come in 'now

Adjectives

'Nor'wegian <> a 'Norwegian 'tanker
 'i'deal <> an 'ideal oppor'tunity
 'inter'national <> an 'international 'athlete
 'Anglo-'Irish <> 'Anglo-Irish re'lations

Adverbs

'al'ready <> he's 'already 'done it
 'abso'lutely <> 'absolutely im'possible
 'alto'gether <> 'altogether 'wrong

An adequate stress shift notation is probably of far greater importance for the advanced foreign speaker than for the native speaker. In the latter's pronunciation, stress shift is a fully automatic process of which he is normally completely unaware. However, the Dutch speaker of English, in whose language stress shift is a much more limited phenomenon (cf. Gussenhoven and Broeders 1981), will soon reach the stage where he finds himself wondering whether compounds like weekend or goodwill ought to be pronounced like AN 'weekend, 'goodwill, i.e. with the accentual pattern of the English loanwords in Dutch, or 'week'end, 'good'will. The problem is aggravated by the fact that (a) some compounds have no fixed accentual pattern, e.g. short cut, ceasefire, high chair, public school or ice-cream; (b) the accentual pattern of compounds is extremely unpredictable for the foreign speaker (e.g. 'Christmas 'pudding, 'front'door; (c) there are differences between British and American accents, as in lawn tennis.

The marking of stress in compounds and idiomatic expressions

This brings us to our third point, the marking of stress in compounds and idiomatic expressions. In Dutch, noun compounds normally have the main accent on the first element (unless the word is no longer felt to be a compound). In English, there is a similar tendency, but there are very many (groups) of exceptions. Here are some examples (based on Gussenhoven and Broeders 1981:121-122).

names of streets and buildings: 'Leicester 'Square, 'Waterloo 'Station;
 prefixes non, arch, ex, self, under, over, vice: 'non-'starter, 'arch'bishop, 'ex-'wife, 'over'statement, 'self-'service, 'vice-'chancellor;
 where the first element refers to place or time: 'back'yard, 'town'hall, 'north'pole, 'armchair'gardening, 'spring'cleaning;

names of dishes: 'apple 'pie, to'mato 'soup, 'Christmas 'pudding.

The Dutch speaker will naturally be inclined to accent all of these on the first element, while in combinations like try-out, pin-up, lay-out the opposite is likely to happen, as in Dutch try-'out, pin-'up, lay-'out. In the new GROOT WOORDENBOEK, compounds whose constituents occur as separate entries are not transcribed, but are assigned stress-marks in the headword. A brief look at the standard learner's dictionaries shows that their treatment is not particularly felicitous. On page xxii of the ALD for example we are told that "A compound that is made up of two separately written words is normally spoken with the strong stress in the second word: ,national 'park." This is subsequently called the 'normal pattern' and is not normally marked. If, however, the main stress comes in the first element, this is always shown, as in 'post office. As a result the unpredictable pattern is NOT marked, while the predictable pattern is.

Thus under country the compounds country club, country dance, country cousin, country house, and country seat are all unmarked, while country party is marked. For the same reason, under air and credit, the vast majority of the compounds that are written as two separate words have to be marked because they have initial stress. Recording angel and capital goods are unmarked, and therefore presumably have even stress. However, when we look at the entries hot and home we find that all noun compounds are explicitly marked.

By contrast, the treatment in LDOCE is exemplary: all compounds are marked explicitly; there is no default notation. Here are some examples:

bus stop /'../, town hall /..'./, Association Football /.....'../

The marking of the accentual pattern of idiomatic expressions is a second area which is still frequently treated in a rather cavalier fashion, again possibly because the native lexicographer is not sufficiently sensitive to the needs of the foreign user. Here the exception is ALD. The problem is probably partly due to the native speaker's unawareness of the unpredictability, or if you like idiomaticity, of the accentual pattern of idiomatic expressions. We ourselves had never been aware of the crucial difference in Dutch between 'n hapje Eten (have something to eat) and 'n HAPje eten (have a bite), until the English co-author of this paper used the second when he clearly meant the first. Compare English WAIT a minute (i.e. I haven't finished) and wait a MINute (i.e. sixty seconds). Here are some English expressions to illustrate the nature of the problem.

Compare:

never put a foot 'wrong
put your best foot 'forward
put your 'foot down
put your 'foot in it

not have a leg to 'stand on
have a 'bone to pick with
have money to 'burn

look the other 'way
learn the 'hard way
no 'way

In English, the rules for stress assignment are not yet fully understood (cf. Gussenhoven 1983). However, what is clear is that it is certainly not the case, as ALD and the LONGMAN DICTIONARY OF ENGLISH IDIOMS (LDEI) suggest in their introductions, that the tonic falls on the "last 'important' word in the idiom" (ALD p. xii) or "on the last word in a sentence if it is a content word" (LDEI p. xiii). The only virtue the ALD convention has is its consistency. Thus under house we find put one's 'house in order, get on like a 'house on fire and bring the 'house down all duly marked. LDOCE is sadly lacking in this respect: there is no marking of the accentual pattern of idiomatic expressions whatsoever. Apart from the unrevealing nature of the convention it shares with ALD, LDEI has the additional disadvantage of not practising what it preaches, i.e. it does not consistently mark the main stress when this is not on the last content word. The reason for this is obvious; it lies in the ill-conceived nature of their convention. The stress mark is frequently omitted where its position would in fact be regular: in SV sentences on the S, as in see how the land lies, all hell broke loose, till the cows come home. On the other hand, search me, and I must say, where a closed-system item or form word has the tonic, are not marked either.

Further examples of expressions that are unmarked in LDEI, although they should be in view of the marking convention, are make one's 'hair stand on end, let the 'side down, you can say 'that again, not set the 'Thames on fire. Again note that the position of the tonic is quite regular in these cases, namely on the object.

In the following table we list a number of examples of idioms which require marking, arranged according to grammatical structure, with indications of the ways in which they are treated in LDEI, ALD and the recently-published second volume of the OXFORD DICTIONARY OF CURRENT ENGLISH (ODCIE II); a plus sign indicates the presence of a stress mark, a minus sign indicates its absence, while an empty space means that the expression is not listed in the dictionary.

	LDEI	ALD	ODCIE II
in 'any case, at 'any rate	-	+	-
not by a 'long chalk	+	+	-
a 'big shot	-	+	+
the 'big time	-		+
the 'small fry	+	+	+
that's 'X's funeral	-	+	+
like 'nobody's business	-		-
be 'nobody's business	-		+
play 'silly buggers	-		
a 'fine kettle of fish	-	+	-
a 'nasty piece of work	-		-
Jack of 'all trades	-	+	-
a 'whale of time	-	+	-
in the 'first place	-	+	-
all 'over the place	-		
at the 'best of times	-	+	-
look on the 'bright side	-		
at 'one time	-	+	
the 'silly season			-

	LDEI	ALD	ODCIE II
in the 'long run	-	+	-
be on 'speaking terms		+	
a 'stick-in-the-mud	-	+	
a 'good-for-nothing	-	+	
spot 'on	-		
get stuck 'in	-		
get what 'for	+		
be 'coining money	-	+	-
'look the part (?)	-		-
'stretch a point	-	-	+
there's nothing 'in/'for it	-		+
what 'of it	+	+	+
think nothing 'of it	-	+	
no two ways a'bout it	-		-
'you should be so lucky	-		
'I'll be bound	-		-
'that's more like it	-		-
'that's the stuff	-		-
'I should cocoa	-		

Summing up, we can say that when it comes to stress marking LDOCE is all right for compounds, but gives no information on idioms, ALD is excellent for idioms, but misleading for compounds, while the marking of idioms in LDEI is clearly extremely unsatisfactory. As for ODCIE II, although we have not had enough time to study it in detail, our first impression is that again the stress marking leaves much to be desired. What is clearly unsatisfactory is that the stress mark does not appear in the headphrase itself, but is given right at the end of the entry, very much, it would seem, as an afterthought.

In this paper we have attempted to indicate a number of shortcomings we have encountered in the pronunciation component of dictionaries claiming to cater for the needs of foreign users. We have tried to demonstrate that simply providing a transcription for all the headwords will no longer do. It is our view that the pronunciation component of such dictionaries should give explicit descriptions of phenomena like stress shift and the stress-pattern of compounds and idiomatic expressions. We can only hope that, in the future, the editors of dictionaries like LDEI, LDOCE, ALD and ODCIE will show greater appreciation of the needs of the foreign user, and incidentally his teacher, in this area.

References

- Chomsky, N. and Halle, M. (1968) The Sound Pattern of English. New York: Harper & Row
- Gimson, A.C. (1973) "Phonology and the lexicographer" in Lexicography in English (Annals of the New York Academy of Sciences Vol. 211) ed. by R.I. McDavid and A.R. Duckert. New York: ANYAS
- Gimson, A.C. (1980) An Introduction to the Pronunciation of English. London: Arnold
- Gussenhoven, C. (forthcoming) "Stress shift and the nucleus" Linguistics 21

- Gussenhoven, C. (1983) "Focus, mode and the nucleus" Journal of Linguistics 19: 377-417
- Gussenhoven, C. and Broeders, A. (1976) The Pronunciation of English: A Course for Dutch Learners. Groningen: Wolters-Noordhoff-Longman
- Gussenhoven, C. and Broeders, A. (1981) English Pronunciation for Student Teachers. Groningen: Wolters-Noordhoff-Longman
- IPA (1975) The Principles of the International Phonetic Association. London: University College
- Liberman, M. and Prince, A. (1977) "On stress and linguistic rhythm" Linguistic Inquiry 8: 249-336
- Thompson, H.S. (1980) Stress and Salience in English: Theory and Practice. Palo Alto, Calif.: Xerox
- Wells, J.C. (1982) Accents of English. Cambridge: U.P.