



HAL
open science

Electronic dictionaries and the integration of prosody in the lexicographical treatment of polysemy

François Nemo, Mélanie Petit

► **To cite this version:**

François Nemo, Mélanie Petit. Electronic dictionaries and the integration of prosody in the lexicographical treatment of polysemy. *Lexicography and Dictionaries in the Information Age*, Airlangga University Press, 2013. hal-01452770

HAL Id: hal-01452770

<https://hal.science/hal-01452770>

Submitted on 2 Feb 2017

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.

Electronic dictionaries and the integration of prosody in the lexicographical treatment of polysemy

François Nemo, University of Orléans (LLL, UMR7270)
Mélanie Petit, University of Orléans (LLL, UMR7270)

Abstract

We discuss here the way the integration of sound files in lexical-lexicographical entries, which is made possible by the electronic nature of e-dictionaries, opens a new frontier in lexicographical work, by making possible the integration of prosodic features in the description of words. We start by showing the relevance of prosody at a lexical level, and hence by defining “lexical prosody”, using results obtained in linguistic semantics. We then present some of the most important conditions, consequences and challenges of the integration of prosody in lexicography, showing that it cannot be limited to the addition of sound forms in existing descriptions but implies corpus-based studies of the relationship between prosodic features and polysemy. Using monolingual and bilingual descriptions of French *enfin* (*at last, finally* etc.), we then describe the way taking into account prosody reshape the description of a lexical entry, and argue for the necessity of a plurisemic structuring of such descriptions.

Keywords: lexical prosody – polysemy – prosodic lexicography

1. Introduction

To a certain extent, electronic dictionaries could remain electronic versions of paper-made classical dictionaries, a purely technical change with no consequence for lexicographical work as such. Nevertheless, things could on the contrary prove to be quite different, and this change could prove to modify somehow considerably the very nature of dictionaries.

Our aim here is to explore the nature of the changes which could result of a full use of these new opportunities, concentrating on two correlated issues, namely the integration of prosody in lexicographical work and lexicographical representation, and the representation of polysemy.

1.1 Easing of ordinary constraints

A major consequence of the electronic turn is the fact that ordinary constraints limiting the size (and weight) of a dictionary (or a lexical entry) are radically dismantled, opening the possibility to provide a much more detailed and fine-grained representation of the actual use of words.

Another consequence is that lexicography is also freed of some of the classical constraints on lexicographical representation: clicks and links allow for instance for multi-layered encapsulated representation of lexical polysemy, and give the user the possibility to choose how far (s)he wants to go or whether (s)he wants to have access or not to such or such type of information.

Combining those two new opportunities in the description of polysemy then leads, for example in terms of exemplification, to leave the lexicographer with new challenges and new tools.

Being free of the constraint to limit the number of examples which can be given, one can avoid for instance the shortcomings associated with any exemplification, namely, as we shall see, that the lexicographer very often has to illustrate an A meaning with a Ab example, even though there are also Ac, Ad, Ae examples, none of which (including Ab) being indisputably a better illustration of A, with no “chemically pure” A being available.

1.2 Sound form in lexicography

Despite affecting lexicography in a very practical sense, getting free of the constraints of paper lexicography is a minor change compared to the possibility to give access to the oral

form of words, by the integration of sound files in lexicographical entries, sound being of course - but for the phonetic transcription ersatz - absent from paper dictionaries.

The more immediate change is the possibility for the user of an e-dictionary, by a simple click, either on the word or on the phonetic form, to *hear* a word and thus to associate a graphical orthographic form with a sound form.

The importance of this access cannot be underestimated, not only because in monolingual dictionaries, the graphemes used in phonetic forms are not common ground for the users, and thus often are a dead-end in terms of access to sounds, but that in bilingual dictionaries, the sound themselves may not be known and often cannot simply be pronounced by the users, this issue being extremely important for tonal languages for instance or in terms of access to accentual patterns.

1.2.1. Prosodic form in lexicography

But when it comes to sound, one may still to go a step further and consider prosody. The issue in such a case is to know whether prosody is relevant or not to lexicographical work and to the lexicon.

In the discussion of this issue, it is important to distinguish strongly between: i) supra-segmental (prosodic) features which are irrelevant to meaning (e.g. accentual patterns) and do not vary from one use of a word to another; ii) supra-segmental (prosodic) features which play a provable role in interpretation, and which will vary from one meaning of a word to another.

As for the first ones (tones, accentual patterns, etc.), they will not be considered here any further and we shall concentrate on the other features which we shall call from now on “Lexical prosody”.

On that ground, the most basic issue is to know whether talking about the *lexical* association of such features with specific meaning is relevant or not, or whether such intonational constraints are external to the lexicon.

1.2.2. Is there anything prosodic in the lexicon, anything lexical with prosody?

Until the start of the 21st century, the idea that word semantics and polysemy could be somehow associated with prosodic features has only been spelled out locally (e.g. Fónagy about *enfin*) but was not questioned as a whole

It used to be the case for intonational meaning to be considered as being forcefully sentential or utterantial meaning and thus as irrelevant to lexical semantics or lexicology. Intonational meaning, to quote Ladd (1980, 6) was indeed conceived as referring: “*to the use of suprasegmental phonetic features to convey “post-lexical” i.e. sentence-level pragmatic meanings in a linguistically structured way*”.

Things started to change when semanticists with a specialization in the study of polysemy started to realize that important aspects of semantic interpretation could actually be dependant on the way a word was said, and consequently started studying this previously ignored or overlooked dimension of polysemy. With the first results, and first PhD theses on the issue (Petit, 2009), came the idea that semantics could not actually ignore prosodic variation nor to risk attributing interpretations driven by intonation to the signification of signs.

As for lexicography now, whose task is, contrary to semantics, descriptive rather than explicative, the question remains however of knowing whether the association of meanings and prosodic forms deserve to be considered as fully lexical or not.

A first answer is that this issue is distinct from the issue of knowing whether there would be such things as prosodic “morphemes” or to put it differently, if prosodic contours would deserve a morphemic status and thus to be described in a dictionary, as affixes for instance sometimes are.

What is questioned here is rather the lexical and lexicographical relevance of admitting that *linguistic form* cannot be reduced to a phonemic form φ when it comes to describe lexical polysemy, and hence that it should be described - whenever it proves to be useful - as a (φ, π) pairs, associating phonemic form and prosodic form, that could be exemplified by sound files.

There are very distinct ways to address this issue, all of which would deserve a paper on their own, since dealing with this issue is ultimately a matter of knowing where exactly the lexicon stops. For as long as a meaning is described, it is indeed contradictory to describe a meaning that could prove to be associated with such or such π form, and to pretend that there would be no such thing as lexical prosody in at least large portions of the lexicon.

A first illustration will be the way existing dictionaries describe among the meanings of French *enfin*, a class of uses which are said to be dominantly interjective and to express affective values, ranging from the expression of relief or satisfaction to the expression of impatience or resignation, and which are considered as important enough to represent half of the entry. If we name “implicit prosody”, this kind of daily lexicographical recognition of the existence of lexical prosody, and if we consider that the only way to associate the provided examples - for instance (TLFi, 2004) “*enfin seuls !*” (alone at last) - with the lexicographical description which they are supposed to illustrate, is actually to fill a prosodic gap by pronouncing them with the adapted prosodic contour, it is all too obvious that pretending that lexical prosody doesn't exist would be pure hypocrisy. And in any case, it is obvious that making available this implicit lexical prosody by the integration of sound files would considerably facilitate the understanding or identification of the targeted use and meaning.

A second illustration of the fact that specific meanings are not associated with phonemic forms but with intoned phonemic forms, i.e. (φ , π) pairs, as shown by the results obtained of extraction tests, an extraction test being a procedure which consists in:

- extracting a target word from a sound file from a data bank in which hundreds of uses of that word have been stored;
- using playback of the extracted segment to test the capacity ordinary people to associate a meaning to it;
- cross-checking the stability of the associations which are made and thus their inter-subjective status of the π form;
- showing that people are actually able from the sole (φ , π) form to spell out a correct interpretation of “what is said” and a rather robust hypothesis about the nature of the situation in which the word was used.

Last but not least is the fact that in bilingual dictionaries, providing access to a certain meaning cannot either be achieved by providing a simple translation and example. Saying that *enfin* may be translated by *well* when it marks resignation and providing an example such as “*Oh well, as you insist*”, the Oxford French Dictionary (Corréard & Grundy, Eds, 2007) similarly rely entirely on a correct (and at best uneasy) guessing of the targeted prosodic contour, for the same sentence could in fact be uttered with a completely distinct π form and interpretation such as false modesty or plain satisfaction. All considered, it is thus obvious that providing the targeted π form is not only relevant in bilingual dictionaries but absolutely essential to the effective understanding of the entry.

2. Issues in prosody conscious lexicography

Once acknowledged the relevance and importance of lexical prosody, almost all remains to be done, thus opening a new chapter and creating new challenges for lexicography.

The major issue, problem and challenge is indeed that it is not at all sufficient nor possible to make “implicit lexicographical prosody” explicit, and for instance for *enfin* to provide a sound file for each existing meaning and example present in the TLFi (2004) or OUP bilingual dictionary (2007).

Such an approach is indeed condemned by the fact that existing dictionaries are both based on written rather than oral corpora and, to state the obvious, on a strictly intuitive approach when it comes to describing the nature and value of prosodic contours.

This can be illustrated by the way an example such as “*Taisez-vous enfin !*” (Will you shut up”) is given by the TLFi (2004) to illustrate the use of *enfin* to mark impatience, when in reality: i) what most uses of such a sentence mark is actually “reprobation” or “irritation” about something which is going on (or has just started), and also a strong request for it to stop, but not

impatience per se; ii) there is actually another frequent use of *enfin* where residual impatience is strongly marked, the π contour being interpreted as “it’s more than time”, but it is not mentioned at all because of an overgeneralization in the description of other uses presented as marking “relief” or “satisfaction”.

Generally speaking, it must be clear that “lexical prosody” cannot in fact be guessed and must be studied to avoid these discrepancies between the attributed value and the real one, and more importantly to ensure a minimum of exhaustivity in terms of describing polysemy and its prosodic dimension.

A major concern about both problems is the systematic overgeneralization of interpretation-types (Nemo & Petit, 2010), for instance when the notion of “forced resignation” is promoted by the TLFi (2004) as the only descriptor of three prosodic variants of the same performative use of *enfin*, thus letting in oblivion both a prosodic contour marking “willingly resignation” and another use in which resignation is absent and where what is said thanks to the π form is better paraphrased by a problem-free “*forget it*”.

In any case, one admitted that the description of lexical prosody cannot be left to hazardous intuition, it is clear that a methodology must be defined in order to achieve descriptive adequacy and exhaustivity in the description of word’s uses and meanings.

2.1. Studying lexical prosody: methodology

As lexicography as a whole, the study of lexical prosody cannot but rely on corpus. Because of the oral nature of prosodic features, it implies the use of oral authentic data taken from the greatest possible diversity of pragmatic contexts. Even though working on oral data has never been considered a necessity in lexicography, the increasing availability of oral corpora makes such a change possible.

From this global corpus must then be extracted for each word a data bank of uses, allowing the concrete studies of the diversity of uses and its correlation with prosodic features. In some cases, for instance for highly polysemical or polyfunctional words, this data bank may and must imply collecting up to 1500 occurrences/uses of each item, extracting simultaneously the utterance or exchange in which all these occurrences are taking place and the word form (φ, π) itself.

Once collected this reference corpus, the next step is to obtain a double semantic and prosodic characterization of each use, each characterization process being initially made separately. As for our example of *enfin*, this means on the one hand testing each use for semantic-pragmatic features (e.g. “Is *enfin* dealing with on-going discourse or something in the situation?”, “Is the enunciation of *enfin* supposed to put an end to a a situation?”, etc.) and in other hand considering in itself the diversity and distribution of the prosodic contours (i.e. π forms).

The third and more important step (Petit, 2009) is then to cross-check the lexical stability of such or such π form and the prosodic stability of such or such semantic interpretation, and to account for all observed discrepancies. It is this process, which can be somehow automated (cf. the on-going DIASEMIE project), which makes possible

- a much more precise description of lexical polysemy, mainly by destroying over-generalizations and pointing to overlooked uses;
- a statistical approach to the frequency issue.
- the spelling out or confirmation of semantic invariants;

leading to the possibility of providing a complete list of all the interpretation-types and use-types, the formers having no π form (or better said being associated with various π forms), contrary to the latters, which are associated with a specific π form.

The application of such a methodology has far-reaching consequences in the description of lexical entries, for it forces to modify somehow considerably not only the content of the entry but its very architecture;

Such a transformation, nevertheless, is not making the lexicographer's task more difficult, dealing with polysemy, mainly taking π into account paradoxically makes things much easier when it comes to relate uses one with another.

2.2. Reconsidering the lexicographical representation of lexical polysemy

Producing a coherent, complete and readable presentation/formulation of the polysemy of a word - capable of avoiding the systematic and hazardous blend of "interpretation-types" and "use-types" which is observed in existing standards – is our next and final step.

Because contrasting thoroughly existing monolingual or bilingual entries of our example *enfin* with the modified one (Petit, 2009, 455-460), would require far more space than is available here, and because discussing existing lexicographical treatment of polysemy would be even longer, we shall restrict ourselves to the discussion of adopting a plurisemic architecture and content in the writing of a lexical entry. We shall do so by defining plurisemy and then describing how adopting a plurisemic architecture allows simultaneously to integrate lexical prosody and clarify the relationship between meanings.

Plurisemy, contrary to polysemy, is not a matter of dealing with multiple meanings of a word but with various *layers* within a single meaning. Because describing lexical prosody leaves no choice but adopting a plurisemic approach, we may contrast a classical polysemic degrouping such as:

- Enfin*, (1), Finally
 (2) Marking resignation
 (3)
 (4) Marking impatience
 (5) In other words
 (6) Introducing a correction
 (7) All things considered
 (8) Marking perplexity

by a plurisemic architecture and a lexical tree (Petit, 2009, 455-460) such as:

<i>Enfin</i> (φ)		
a	ab	<i>ab1</i> ($\varphi, \pi 1$) <i>ab2</i> ($\varphi, \pi 2$) <i>ab3</i> ($\varphi, \pi 3$)
	ac	<i>ac1</i> ($\varphi, \pi n$) <i>ac2</i> ($\varphi, \pi n'$) <i>ac3</i> ($\varphi, \pi n''$)
	ad	etc.

in which all uses of a word are introduced by a stable invariant stratum (which is inherited by all the meanings), then separated by a second stratum with variable values (b,c) and then by an ultimate stratum whose variable values (1,2,3, n) depends on the π form.

Instead of opposing the marking of relief, impatience (or the corrective use), one remarks that when *enfin* is associated with relief (=1) or irritation (=2), it always concerns a pre-existent problem (=a), and that this problem is either in previous discourse (=b) or in the real world (C), a past or an on-going problem, etc.

What is important and provable is that whereas the a and ab/ac strata (sign-meaning and interpretation-type) are not associated with specific π forms, use-types as ab2 are directly dependant on the π forms.

What is even more important for lexicography as a whole is to realize that the "ab1" use-type meaning - with its three layers of meaning, a, b, and 1 as a comment on ab – is actually the ultimate lexical and lexicographical level, and its most solid ground.

3. Conclusion

The fact that “lexical prosody” can and could be made accessible to the user of dictionaries as a sound file and also recognized and made explicit as a crucial stratum of lexical meaning is both a consequence of the electronic turn and the discovery of a new frontier for lexicographical work.

As we have just seen, if it does imply shifting to oral corpora and adopting a new methodology, it also provides an objective base to improve the precision, exhaustivity and coherence of lexical entries, and to make possible a plurisemic shift in the treatment of polysemy.

References

- Barnes, B. 1995. Discourse particles in french conversation: (*eh*) *ben*, *bon* and *enfin*. *The French Review*, 68, 5, 813-821.
- Beeching, K. 2000. La fonction de la particule pragmatique *enfin* dans le discours des hommes et des femmes. *Femmes et français*, Paris: L'Harmattan.
- Bertrand, R & Chanet, C. 2005. Fonctions pragmatiques et prosodie de *enfin* en français spontané, *Revue de Sémantique et de Pragmatique* n° 17, 41-68.
- Buchi, E. & Städtler, T. 2008. La pragmatization de l'adverbe *enfin* du point de vue des romanistes (« Enfin, de celui des francisants qui conçoivent leur recherche dans le cadre de la linguistique romane »). *CMLF (Paris, 9-12 juillet 2008). Recueil des résumés, CD-ROM des actes*, 159-171.
- Cadiot, A. & al. 1985. *Enfin*, marqueur métalinguistique, *Journal of Pragmatics*, 9, 199-239.
- Corréard, M-H & Grundy, F. Eds. 2007. *Oxford Hachette French Dictionary: French-English, English-French*. Oxford: Oxford University Press. 4th Edition.
- Dostie, G. 2004. *Pragmatization et marqueurs discursifs, analyse sémantique et traitement lexicographique*. Bruxelles: Duculot.
- Ducrot, O. 1987. L'interprétation comme point de départ imaginaire de la sémantique. *Dire et ne pas dire* ([1972], 1991). Paris: Hermann.
- Fischer, K. Ed. 2006. *Approaches to discourse particles*. Oxford : Elsevier.
- Fónagy, I. 1983. *La vive voix: essai de psychophonétique*. Paris : Payot.
- Luscher, J.M. & Moeschler, J. 1990. Approches dérivationnelles et procédurales des opérateurs et connecteurs temporels : les exemples de *et* et de *enfin*. *Cahiers de Linguistique Française*, 11. 77-104.
- Mosegaard-Hansen, M-B 2005. From prepositional phrase to hesitation marker: The semantic and pragmatic evolution of French *enfin*. *Journal of historical pragmatics*, 6. 37-68.
- Nemo, F. 1999. The pragmatics of signs, the semantics of relevance, and the semantic/pragmatic interface. *The Semantics-Pragmatics Interface from Different Points of View*. Amsterdam : Elsevier Science. 343-417.
- Nemo, F & Petit, M. 2010 Sémantiques des contextes-types, *Etudes de sémantique et pragmatique françaises*. Saussure, Louis de / Rihs, Alain (éds). Berne: Lang. 379-397.
- Paillard, D. 2003. À propos de *enfin*. *Ordre et distinction dans la langue et le discours. Actes du Colloque international de Metz (18, 19, 20 mars 1999)*, Paris: Champion. 387-408.
- Petit, M. 2009. *Discrimination prosodique et représentation du lexique : application aux emplois des connecteurs*. Thèse, Université d'Orléans.
- Petit, M. 2010. Prosodie et sémantique des mots : pour une morphologie élargie, *Liens Linguistiques. Etudes sur la combinatoire et la hiérarchie des composants*. Berne: Peter Lang. 99-114.
- Trésor de la Langue Française informatisé (TLFi)*. 2004. Cd-Rom. Atilf (Autor and editor). Nancy.