

# Phonological Licensing of Cranberry Words

Jan-Philipp Soehn

SFB 441, University of Tübingen

jp.soehn@uni-tuebingen.de

On my poster, I would like to attract attention to cranberry words (CWs). This term goes back to Aronoff (1976) and denotes lexical items that reveal idiosyncratic behaviour concerning their distribution just like the morpheme *cran*. This means that they impose restrictions on the context in which they occur that go beyond normal syntactico-semantic requirements of their part-of-speech, e. g. *in a trice* / \**at the trice* vs. *in a moment* / *at the moment*.

CWs are quite numerous (about 450 in German) and exhibit a fascinating variety of idiosyncrasies. Zooming in even closer, we find CWs that impose phonological constraints on their context, or, put differently, that are lexically licensed by certain phonological features of the context. One very prominent example is the *a/an* alternation in English (*a paper* vs. *an article*), discussed in Bloomfield (1935). This phenomenon is also known as external sandhi or ‘shape conditions’ (cf. e. g. Pullum and Zwicky, 1988; Spencer, 1991). I take *an* as a CW that co-occurs exclusively with particular segments (mainly vowels) immediately to the right. Even if describing the phenomenon informally is more or less straightforward, it is difficult to handle for a grammar theory. On the one hand, it doesn’t seem to be part of the word’s lexical entry because there is a reference to the following syntactic context. On the other hand, it can’t be a phonological rule because the *-n* affixation is not a productive process in English and it is reserved to the English indefinite article. Thus, this kind of lexical idiosyncrasy provides (further) evidence for the need of a “collocation module” in grammar which provides the tools to formulate the constraints that these words impose.

An allomorphy such as *a/an* can be found for the definite article *the* as well. The form [ði:] is used before a sandhi-triggering sound and before a speech pause. Also, the alternation [ə]/[i:] is idiosyncratic for this lexical element. Similar phenomena can be found in other languages and I will discuss such cases in French and Welsh: Certain prenominal adjectives in French have an irregular masculine singular form in front of vowels, such as *beau* (*un bel arbre* ‘a beautiful tree’) or *vieux* (*un vieil homme* ‘an old man’).<sup>1</sup> As far as Welsh is concerned, some function words exhibit word-

---

<sup>1</sup> For reasons of space, I have to defer a more detailed discussion of this phenomenon linked to “liaison” in French to a later occasion.

final alternation as well (cf. Lapointe, 2001), which is of exactly the same sort as the English *a/an* pair. However, the Welsh conjunction *a/ac* ‘and’ occurs with its C-final form also before a very limited but heterogeneous set of consonant-initial function words comprising prepositions, conjunctions, adverbs, polarity items, a quotative form, and two forms of *be*. I conclude with Lapointe that the conjunction *a/ac* behaves idiosyncratically, not only with respect to the following segment but also with respect to the following lexeme, thus revealing an interplay between different kinds of lexical licensing.

Our analysis within HPSG is based on Soehn (2004) and provides a way to capture the idiosyncratic distribution of a sign. In this approach a new list-valued feature, `COLL` (Context of Lexical Licensing), is introduced, by which a sign can express constraints on the environment in which it is allowed to occur. `COLL` contains licensing objects (cf. Sailer, 2007), which in turn bear the features `LICENSER` for the constraint itself and `LOCUS`. The latter feature determines the scope of the constraint and thus adequately restricts the collocation module. The value of `LOCUS` is a phrase of a certain kind (*utterance, complete-clause, np,...*) which is identified as a node in the syntactic configuration above the sign in question. The `LICENSING-PRINCIPLE` guarantees that a phrase specified in `LOCUS` minimally dominates the sign and meets all the criteria mentioned within the value of `LICENSER`. The conception of `LOCUS` provides a “window” in which collocation restrictions must be satisfied.

Restrictions on the phonology of adjacent signs affect the phenogrammatical structure, not the tectogrammar, cf. *\*a/an apple* vs. *a/\*an big apple*, and thus a linearization component is needed as well. Although the determiner is the first constituent within an NP headed by *apple* in both cases, it is the word to the right that decides on the form *a* or *an*. I assume that each sign has a word-order domain, expressed in the `DOM` list, which is subject to linear precedence rules and other constraints which influence constituent order. Elements of that list are *dom\_objects* that contain information below `PHONOLOGY` and `SYNSEM` (cf. e. g. Crysmann, 2002). In the HPSG literature about a linearization component, there is a discussion about what kind of information should be included in the *dom\_objects*. On the one hand, there has to be enough information to be able to describe all relevant phenomena, on the other hand, one doesn’t want to repeat all of a sign’s information on its `DOM` list – this would make the linearization component too powerful. The Welsh cases provide evidence for the need to have not only phonologic but syntactic (lexeme) information as well.

Concerning phonological restrictions, the lexical item which imposes the restriction does not have information about adjacent words. Only some phrase which dominates it has sufficient “knowledge” about its parts and their relative position (on `DOM`). Therefore, it makes sense to “look upwards” in the tectogrammatical structure in order to access phenogrammatical information. The poster will contain some examples of the analysis, e. g. the sketch of the lexical entry of *vieil* ‘old’: its lexical entry comprises phonological, syntactic, semantic and distributional information. The French adjective *vieil* [vjɛj] is the masculine singular form of *vieux*. The main semantic contribution of that sign is the meaning ‘old’. The `COLL` value specifies that this sign is li-

censed within an NP (it is an attributive adjective) and before a sandhi-triggering segment, i. e. a vowel or a particular kind of consonant.

In sum, I make use of a module of grammar which has been designed for lexical licensing and which has already been implemented into a grammar fragment of German. I propose an analysis of CWs licensed in environments with particular phonological properties. With the approach being included in a grammar, idiosyncratic cases which escape regular rules of sandhi and other phonological assimilation processes can thus be handled correctly. Our data provide linguistic evidence in favor of a lexical licensing module in formal grammar frameworks and in favor of a certain shape of a linearization component.

## References

- Aronoff, Mark (1976). *Word Formation in Generative Grammar* (Dritter Druck 1985 ed.). MIT Press, Cambridge MA. Linguistic Inquiry Monographs.
- Bloomfield, Leonard (1935). *Language*. George Allen & Unwin, London. Linguistic Inquiry Monographs.
- Crysmann, Berthold (2002). *Constraint-based Coanalysis. Portuguese Cliticisation and Morphology-Syntax Interaction in HPSG*. Ph. D. thesis, Saarland University.
- Lapointe, Steven G. (2001). Stem selection and OT. In G. Booij and J. van Marle (Eds.), *Yearbook of Morphology 1999*, pp. 263–297. Kluwer Academic Publishers, London.
- Pullum, Geoffrey K. and Zwicky, Arnold M. (1988). The syntax-phonology interface. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey, Volume I of Linguistic Theory: Foundations*, pp. 255–280. Cambridge University Press.
- Sailer, Manfred (2007). NPI Licensing, Intervention and Discourse Representation Structures in HPSG. In St. Müller (Ed.), *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford*, pp. 214–234. CSLI Publications.
- Soehn, Jan-Philipp (2004). License to COLL. How to bind bound words and readings to their contexts. In St. Müller (Ed.): *The Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Leuven*, pp. 261-273. CSLI Publications.
- Spencer, Andrew (1991). *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Blackwell, Oxford.