

Jacques Durand (CLLE UMR5263 CNRS & Université de Toulouse II, IUF)

Julien Eychenne (Hankuk University of Foreign Studies, Yongin, South Korea)

*The word in French: its status, its function, its phonotactic properties*

Abstract

French is often described as a language where the word is difficult and often impossible to delimit from a phonetic point of view. From a typological perspective, it has been argued that French is a prime example of a so-called ‘cursus’ language by opposition to ‘nexus’ languages like English (Pulgram 1970). As a cursus language, French is assumed to have a syllabic rhythm which gives priority to breath groups within which individual lexemes lose all prosodic and segmental identity. Thus it is pointed out that in French stress is not lexical but should be viewed as a property of the last syllable of a rhythmic group (see inter alia Troubetzkoy, 1939[1970]: 296, note 3, Fox 2000: 94). By contrast, in a language like English word-stress is claimed to act as a cue which allows the hearer to retrieve lexical units. Segmentally, fixed final consonants and liaison consonants are said to behave like onsets of following vowel-initial words, an “enchaînement” which blurs away crucial word-boundaries. One therefore wonders if the word is not ultimately a projection of orthographic norms, a position defended most trenchantly by Laks (2003, 2005) who, within a usage-based approach, appears to see the word as, at best, the by-product of larger chunks memorized by speakers/hearers (and writers/readers).

Our position is that the word in French, whatever its ultimate mental status, cannot be dispensed with at various levels, including the phonological/phonetic one (an argument already developed by Lyche and Girard, 1995). Before all else, however, a distinction must be established between lexemes, morphosyntactic words and word-forms. Our argument, once made more precise, is that word-forms can often be identified in a way close to orthographic conventions. We remind the audience that words in that sense correspond to minimal free forms, have an internal stability different from that of phrases, are accessible epilinguistically in informal paraphrases and can sometimes correspond to independent gestural signs. From an acquisition point of view, proper names and imperative forms both often based in French on independent word-forms constitute one of the fundamental building blocks of language. Having established the plausibility of the word as a basic linguistic unit, we turn to phonological and phonetic considerations showing its relevance: stress phenomena (Lyche and Girard 1995, Lacheret and Lyche 2008), prosodic contours extracted from large-scale oral corpora (Ada-Decker, Nemoto, and Durand 2009), the stabilization of schwa in word initial position (Hansen 1994, Walker 1996), the difference between monosyllabic grammatical words and monosyllabic prefixes (Côté 2007), various “enchaînement” phenomena (Durand and Lyche 2008). We finally turn to phonotactic arguments, in a classical sense (see in particular Basbøll, 2001, to appear, Basbøll and Lambertsen (to appear)). We focus on some examples of clusters which give strong indication as to word-boundaries (e.g. /tʃ/ is a strong indicator of non word-initial position, /tʃ/ with devoicing of the /r/ is word-final). This leads us to a discussion of gradience within linguistic systems and the appropriateness of models such as Optimality Theory (OT). OT has changed our conception of phonotactic regularities by invoking constraints which are not simply present or absent but can have a gradient effect through ranking. We argue that recent developments such as Maximum Entropy (MaxEnt)

grammars (Goldwater & Johnson 2003; Hayes & Wilson 2008) provide an interesting mid-way between classical generative models and purely data-driven, usage based approaches. These models are able to capture stochastic effects while still explicitly modeling symbolic generalizations. We show how a large corpus such as PFC (Durand, Laks, Lyche 2002, 2009) gives an empirical database on which such grammars can be built, focussing our attention on word-edge effects and on the role of the constraint NOCODA in external sandhi contexts.