SEGMENTAL AND PROSODIC ISSUES IN ROMANCE PHONOLOGY

AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE

Series IV – CURRENT ISSUES IN LINGUISTIC THEORY

Volume 282

Pilar Prieto, Joan Mascaró and Maria-Josep Solé (eds)

Segmental and prosodic issues in Romance phonology

SEGMENTAL AND PROSODIC ISSUES IN ROMANCE PHONOLOGY

Edited by

PILAR PRIETO
ICREA – Universitat Autònoma de Barcelona

JOAN MASCARÓ
Universitat Autònoma de Barcelona

MARIA-JOSEP SOLÉ
Universitat Autònoma de Barcelona

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Segmental and prosodic issues in romance phonology / edited by Pilar Prieto, Joan Mascaró and Maria-Josep Solé.

p. cm. -- (Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory, ISSN 0304-0763 ; v. 282)

Includes bibliographical references and index.

1. Romance languages--Phonology--Congresses.

PC76 .S44 2007

440--dc22 2007060751

ISBN 978 90 272 4797 1 (Hb; alk. paper)

John Benjamins Publishing Co. • P.O.Box 36224 • 1020 ME Amsterdam • The Netherlands

John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

CONTENTS

INTRODUCTION

The second *Phonetics and Phonology in Iberia* (PaPI) conference, hosted by the Universitat Autònoma de Barcelona in June 2005, proved a great success in bringing together scholars from around the world, all of them involved in researching contemporary issues in phonetics, phonology, and related areas, including language acquisition, language variation and change, and speech technology.

This volume provides a selection of the papers presented at that conference. Most of them are concerned with the relationship between phonetics and phonology, and most of them also use a methodological approach that has come to be known as 'laboratory phonology'. This approach, whose foundations were first laid about twenty years ago, sets out to answer a wide array of research questions through the use of experimental methods. In other words, experimental methodology previously associated with phonetic studies is applied to the realm of phonology with the goal of exploring the crucial correspondence between empirical data and theoretical claims. Over these last two decades, this experiment-based approach has proved extremely fruitful in its ability to test controversial claims in phonological theory, resolve phonological issues, and discover the principles guiding linguistic mechanisms (for a good overview, see articles by Pierrehumbert et al. 2001 and D'Imperio 2005).

The specific focus of the papers in the present collection is descriptive and theoretical issues in the phonology of Romance languages. These papers provide new empirical data on a number of phonetic and phonological phenomena in a variety of Romance languages and their dialectal varieties, including Catalan (Eastern and Western Catalan, Valencian and Majorcan), French (European and Quebec), Italian (Neapolitan), Portuguese (Standard European and Northern European) and Spanish (Andalusian, Argentinian, Central Peninsular and Chilean). Importantly, most of the contributions take a crosslinguistic or crossdialectal perspective, paving the way to a better understanding of linguistic differences in typologically close language varieties. This focus on Romance languages is motivated by our belief that there is a need for multilanguage data to test current theoretical claims and models, which often lack precisely this sort of broad crosslinguistic basis. The virtue of crosslinguistic research is that it constitutes a valuable tool to explore similarities and differences between languages and thus allows us to construct general linguistic theories, while at the same time ensuring that the peculiaritites of individual languages can be characterized within the theory.

An important goal of this volume is to bridge the gap between traditional Romance linguistics—already with a long and rich tradition in data collection, cross-language comparison, and phonetic variation—and laboratory phonology. In our view, subjecting the theoretical claims and data from traditional Romance linguistics to the scrutiny of experimental techniques in the laboratory can only strengthen the scientific basis of this discipline and better integrate its findings in current phonetic and phonological theory. Though in recent years laboratory phonology has proved to be a broad and fertile interdisciplinary approach in Romance linguistics and has grown in popularity among researchers, it is still not a well-known and established approach among Romance scholars. The body of experimental work devoted to Romance languages is still far smaller than the work that has examined, for example, Germanic languages. This volume is thus an attempt to help redress that imbalance: as the reader will see, the studies collected herein present cutting-edge laboratory phonology research as applied to Romance languages.

The volume has been organized into three main topic areas, which reflect the main themes of the conference. The first is concerned with segmental processes (coarticulation and assimilation processes, sandhi processes, feature cooccurrence and sequential restrictions), the second with prosodic structure (prosodic characterization of parentheticals, syllable structure, prosodic phrasing, stress and pitch accent prominence), and the third with the acquisition of segmental and prosodic features (the acquisition of vowel reduction, L1 and L2 vowel perception, initial word segmentation, and L2 rhythmic patterns). Thus we begin with the smaller segmental units, move on to larger phonological constituents, and conclude with a look at acquisition on both levels.

Section 1, *Segments and processes*, comprises four papers which address the phonetic properties of segments, their mutual influence, and phonological processes across words. Three of these papers focus on the interaction of production and perception in phonological structure and sound change, mostly within the framework of gestural phonology (Browman & Goldstein 1986), which allows modelling of the biomechanical and aerodynamic constraints of speech gestures in a way that accounts for the actual observed patterns. These papers also serve to illustrate the notion that sound change can be partly explained by universal phonetic factors and has its origins in synchronic variation (Ohala 1989, 1991).

Daniel Recasens examines the coarticulation patterns found in different types of VCV sequences in Catalan and Spanish and how they provide support for the 'Degree of Articulatory Constraint' (DAC) model of coarticulation. In particular, he argues that the workings of the speech production mechanism, and in particular the DAC model, can explain the directionality and extent of coarticulation found in VCV sequences. Acoustic data on the size of anticipatory and carryover effects between vowels [a] and [i] and a set of consonants clearly show that: (a) some VCV sequences with salient consonant anticipatory effects—like sequences with dark [l]—exhibit more vowel-to-

consonant anticipation effects than carryover effects, while sequences involving the alveolopalatals [ɲ] and [ʎ] show the reverse pattern; (b) vowel anticipation does not exhibit a comparable size for all consonants but is greater for unconstrained consonants than for more constrained ones. As predicted by the model, the direction and extent of vowel coarticulation varies inversely with the degree of constraint of the intervocalic consonant. Importantly, Recasens provides evidence in support of a close match between the predictions of the model and a number of observed sound changes and assimilatory processes in various Catalan dialects.

Maria-Josep Solé argues that the articulatory-acoustic stability of phonological features may be affected not only by concurrent features, but also by features in adjacent segments when they coincide in time due to coarticulatory overlap. Specifically, she addresses the question of whether aerodynamic factors are at the origin of the incompatibility between nasality and frication. She presents the results of a series of experiments designed to explore the effects of velopharyngeal opening (or degrees of nasality) on the stability of segments requiring a high pressure build-up in the oral cavity, such as fricatives. Acoustic and aerodynamic evidence shows that in fricative + nasal sequences anticipatory velum lowering during the acoustic duration of the fricative reduces or even extinguishes the pressure difference required for frication. This is clear evidence that frication is unstable when it comes in contact with nasalization in adjacent segments. She presents important additional evidence that this instability is at the origin of a number of phonological patterns found historically and synchronically in Romance languages by which fricatives tend to lose their friction when they precede nasal consonants. In addition, she argues that the same aerodynamic and acoustic factors responsible for the combination of features within a segment can be used to explain how features interact in contiguous segments.

Francisco Torreira deals with the well-known phonological process of /s/ aspiration in coda position present in a large number of Spanish dialects. The author furnishes instrumental data showing that in Andalusian Spanish, a southern variety of Peninsular Spanish, /s/ aspiration before voiceless stops is accompanied by consistent postaspiration of the stop consonant. His analysis of spontaneous speech clearly shows that /s/ preceding a stop, though usually realized as a period of aspiration or breathy voice, may be absent in a considerable number of cases. Crucially, voiceless stops following /s/ show a consistent pattern of postaspiration. This asymmetry suggests that the conditioning factor for postaspiration in Andalusian voiceless stops may be the presence of a preceding underlying laryngeal gesture that is not strictly timed with the supralaryngeal gestures. Torreira argues that a gestural analysis (Browman & Goldstein 1986) offers a plausible account for a phenomenon that is especially difficult to explain in terms of segments. Within this framework, the timing of the starting point of the glottal opening gesture with respect to the supraglottal closure may not be very accurately specified, while the timing of the ending point with respect to the end of the stop closure is more precise.

Finally, Torreira reviews diachronic and synchronic examples from other languages which illustrate various paths of change for the same type of sound sequence. He suggests that this pattern in Andalusian Spanish, showing unstable preaspiration and more consistent postaspiration, might eventually lead towards a new category of aspirated voiceless stops, as has occurred in other languages.

Noël Nguyen and colleagues address the well-known sandhi phenomenon of French liaison and the question of how liaison consonants are processed in speech perception. Specifically, the study analyses whether during speech comprehension liaison consonants are processed and represented differently from word-initial and word-final consonants. With this object, the authors undertook a series of perception experiments that examined potential differences in the detection rate of liaison consonants vs fixed consonants and attempted to determine whether these differences might be attributable to the phonetic properties of the consonants involved. The results provide evidence that liaison consonants are more difficult to detect than word-initial consonants (detection scores were lower and response times tended to be slower for the former than for the latter) and that these differences are not attributable to potential phonetic differences between the two. Nguyen and colleagues argue that the difficulty in processing liaison consonants seems to provide partial support for the autosegmental representation of liaison consonants as floating segments. Yet as detection accuracy also seemed to vary depending on the degree of lexicalisation of the carrier word sequence, the authors point out that more research is needed to evaluate the potential effects of probability of occurrence and thereby properly evaluate the predictions of exemplar-based models in the perception of liaison consonants.

Prosodic structure and intonational phonology have been recurring themes in laboratory phonology work, partly because of the unreliability of introspective work on these issues. The four papers in Section 2, *Prosodic structure*, address important issues in this area of study, taking advantage of the benefits offered by using a common crosslinguistic experimental approach.

Lluïsa Astruc-Aguilera and Francis Nolan examine the prosodic characteristics of extra-sentential elements (dislocated phrases, vocatives, adverbials, etc.) in Catalan and English. The authors present the results of three experiments set up to study phrasing and accentuation patterns in these constructions. Though phonological studies have traditionally considered that extra-sentential elements form prosodically independent units and are unaccented, results from these experiments show that they do not always form independent tonal units, nor are they always deaccented. Rather, they show variation in their phrasing and intonation, revealing a trade-off between prosodic independence and tonal subordination: deaccentuation only seems to be compulsory in those cases in which the extrasentential elements and the main phrase belong to the same prosodic domain. Also, the experiments reveal that accentual cues seem to be stronger and more consistent cues than phrasing cues and the degree of inter-speaker variation is lower in the former than in the

latter. In their third experiment, using both a database in which three levels of stress were specified and a masking noise technique for recording, the authors found that right-dislocated phrases were totally deaccented. Astruc and Nolan conclude that prosody signals the pheripheral status of extrasentential elements by general deaccentuation or compression of the pitch range, independent phrasing being a more optional cue.

Laura Colantoni and Jeffrey Steele provide a detailed account of the behavior of stop-liquid clusters in two varieties of Spanish (Chilean and Argentinian) and two varieties of French (Quebec and European). They show that choice of cluster simplification, whether by assimilation or dissimilation, is correlated with the voicing properties of the stop and the manner properties of the liquid (tap, fricative, approximant). The results of the production experiment in the four varieties under investigation show that: a) similarity in manner and voicing between the two members of the cluster determines the degree of cluster simplification; b) in the case of stop-rhotic clusters, the phonetic characteristics of the rhotic determines the strategy used; in the case of Spanish, with the tap being highly similar to the stop, dissimilation via vowel epenthesis is the preferred outcome; and c) stop voicing plays a role in determining the degree of assimilation and dissimilation: given that voiceless stops are longer than their voiced counterparts, a compensatory lengthening effect is observed, and thus shorter epenthetic vowels are found. The effects that trigger synchronic variation can also account for the evolution of stop-liquid clusters from Latin to Romance. Finally, the authors provide an optimality theory-based analysis of their experimental results.

The chapter by Sónia Frota and colleagues explores the phonetic realization of intonational phrasing in five Romance language varieties: Catalan, two varieties of European Portuguese, Neapolitan Italian, and Spanish. Data from a common experimental database (the 'Romance Languages Database') is used to analyze the phonetic realization of phrasing in the five varieties under examination. First, the authors provide a typology of combinations of nuclear pitch accents plus boundary tones used across languages, as well as their relative frequency. The dominant boundary tone used in the five varieties is the high (H) boundary tone, in the form of either a continuation rise or sustained pitch. Second, they offer a detailed analysis of the phonetics of the H boundary tone across languages. Specifically, the data reveal that nuclear pitch accent choice affects the scaling of the H boundary tone in a similar and consistent way, namely, the tone is higher after High nuclear accents than after Low nuclear accents. They interpret this as resulting from the upstep of the H boundary tone after an accentual H. Also, the data reveal mixed effects of constituent length on the scaling of H boundary tones, with the languages observed clustering in two main groups: the Catalan-Spanish group (with almost no length effects) and the Italian-European Portuguese group (with clear length effects).

Marta Ortega-Llebaria and Pilar Prieto's paper examines the acoustic correlates of stress prominence in Spanish in both accented and unaccented

environments (e.g. parentheticals). Traditional studies typically describe the correlates of stress in accented environments, thus suffering from covariation between stress and accent. This paper goes beyond traditional accounts in that the pitch accent factor is controlled for. The results of the production experiments described reveal that the stress contrast is maintained in deaccented contexts and that syllable duration and spectral tilt (intensity at high frequencies of the spectrum) are reliable acoustic correlates of this contrast in Spanish. These results contribute to the discussion about the nature of stress across languages, advocating for the view that stress prominence has its own phonetic cues, and against other views which claim that stress cues are parasitic on vowel reduction cues. Thus while American English, Dutch and Spanish differ in the degree of vowel reduction involved in marking stressed positions, they do not differ greatly in the way they use other acoustic correlates (i.e. duration and intensity) to signal the presence of stress.

Section 3, *Acquisition of segmental and prosodic structure*, comprises four papers which address the acquisition of segmental and prosodic contrasts by infants and second language learners. The papers in this section illustrate how laboratory phonology is in fact starting to bridge the gap between psycholinguistics and phonology.

Maria João Freitas focuses on the early production of vowels in unstressed position by European Portuguese-speaking children, and specifically, on how these children start acquiring two specific phonological processes of vowel reduction, namely, that of /ɛ, e/ turning into [ɨ] and that of /a/ turning into [ɐ] in unstressed positions. Since the acquisition of vowel reduction processes has not received much attention in the acquisition literature, this paper provides new empirical data from European Portuguese, a Romance variety which presents a number of reduced vowels in unstressed position deriving from the productivity of the vowel reduction process. On the basis of longitudinal data collected for four children aged 0;10 to 2;8, it is observed that Portuguese children acquire vowel reduction relatively early in the path of development, and that syllable deletion is one of the common strategies found in children. Freitas claims that the complexity of the target vowel system increases children's early sensitivity to vowel differences and promotes the speed of phonological development. Interestingly, the results show that vowel reduction emerges either simultaneously in word-medial and word-final position or possibly earlier word-finally. Freitas suggests that the presence of morphological content in the word-final vowel might be promoting phonological development in this position.

Geoffrey Stewart Morrison introduces logistic regression analysis as applied to L1 and L2 speech perception data involving Spanish vowels. The chapter is intended as a tutorial for L2-speech-perception students and researchers who are not familiar with the technique. Using data taken from previous identification experiments on L1 Spanish vowel perception and on L1 and L2 English vowel perception, the author applies logistic regression model fitting techniques to determine which acoustic cues are attended to by listeners

when identifying stimuli. He shows that logistic regression coefficients can be successfully used to produce intuitive representations and quantify how listeners use those acoustic cues, as well as to model sequential stages in L2 learners' perception. At the same time, these statistics can also be used to determine whether there are significant differences in the perception of stimuli by L1 vs L2 groups of listeners. In sum, Morrison shows how the logistic-regression technique can be successfully used in L2 speech perception research.

Ferran Pons and Laura Bosch focus on how infants under one year of age deal with word segmentation and which prosodic features they pay attention to in order to perform this task. Sensitivity to prosodic information has been observed very early in development in studies with English and Dutch children. For example, at nine months, American English infants show a trochaic bias, meaning that they prefer to listen to lists of strong-weak disyllabic words (trochaic), as opposed to lists of weak-strong disyllabic words (iambic), a stress pattern which is atypical of English. Using a slightly modified version of the Head-Turn Preference Procedure, a paradigm that has been used successfully in infant speech perception research over the past twenty years, Pons and Bosch set out to explore the metrical preferences of six-month-old Spanish- and Catalan-learning infants. The data revealed no pattern of preference for trochees. An additional experiment with nine-month-old infants revealed that, unexpectedly, even at this age they do not show a pattern of preference for trochaic or iambic stress. The authors partly attribute this crosslinguistic difference to a weaker predominance of the bisyllabic trochaic pattern in Catalan and Spanish relative to English. Yet the results cast some doubt on the usefulness of this prosodic cue (i.e. stress pattern) alone to help early word segmentation of fluent speech. The authors suggest that phonotactic information—the fact that heavy CVC syllables appear generally in stressed positions—might be combined with stress cues at a very early age in order to predict the patterns of preference.

Finally, Laurence White and Sven Mattys set out to test the discriminative performance of different metrics of rhythmic distinctions across languages. One of the novelties of this article is that it collects data from second language rhythm, in the hope that the metrics will prove useful as a tool to identify the rhythmic differences between native English speakers, for example, and Spanish speakers of L2 English. The authors' first production experiment is designed to test how well different metrics support the distinction between the rhythm of 'syllable-timed' French and Spanish and that of 'stress-timed' Dutch and English, with the effect of L1 on L2 rhythm also considered. The results show that rate-normalised metrics of variation in vocalic interval duration clearly and effectively (a) discriminate between the classic distinction between stress-timed and syllable-timed languages; and (b) are informative about the adaptation of speakers to rhythmically-similar (Dutch and English) or rhythmically-distinct (Spanish and English) second languages. Their second production experiment examines the rhythmic contrasts between

different accents of British English, with results showing evidence of rhythmic gradience between them. Finally, results from a perceptual test find a normalised metric of vocalic interval variation to be the strongest predictor of the rating of the second language speaker's accent as native or non-native.

As a final word, we would like to thank the scholars who agreed to review the contributions included in this volume. We are greatly indebted to the anonymous reviewers at the John Benjamins office as well as the external reviewers who have participated in the assessment of articles: Laura Colantoni, Néstor Cuartero, Eva Estebas, Paula Fikkert, Chip Gerfen, Barbara Gili-Fivela, José Ignacio Hualde, Conxita Lleó, Francis Nolan, Hugo Quené, Daniel Recasens, Marija Tabain, and Laurence White.

We owe the smooth progress in the production of this book to Anke de Looper of Benjamins and E.F.K. Koerner, the series editor, who we would like to thank for their active and continuing support from the start of the project. Many thanks are also due to Michael Kennedy-Scanlon, Marianna Nadeu and Maria del Mar Vanrell for their help in textual editing and proofreading. This work was partially supported by grants BFF2003-08364-C02, HUM2004-20318-E, and HUM2006-01758/FILO awarded by the *Ministerio de Ciencia y Tecnologia* and *FEDER*, and by grant 2005 ARCS1 00174 awarded by the *Agència de Gestió d'Ajuts Universitaris i de Recerca* (*AGAUR, Generalitat de Catalunya*).

We believe this volume will constitute a very useful companion for phoneticians, phonologists, and researchers investigating sound structure in Romance languages. It is our desire that it will spark further interest in laboratory phonology and will contribute to enlarging the body of research focusing on these languages.

Barcelona, November 2006

Pilar Prieto
Joan Mascaró
Maria-Josep Solé

References

Browman, Catherine P. & Louis Goldstein. 1986. "Towards an articulatory phonology". *Phonology Yearbook* 3. 219-252.

D'Imperio, Mariapaola. 2005. "La Phonologie de Laboratoire: Finalités et quelques applications". *Phonologie et phonétique: Forme et substance* ed. by Noël Nguyen, Sophie Wauquier-Gravelines & Jacques Durand. 241-264. Paris: Hermès.

Ohala, John J. 1989. "Sound change is drawn from a pool of synchronic variation". *Language Change: Contributions to the study of its causes* ed. by Leiv E. Breivik & Ernst H. Jahr. 173-198. Berlin: Mouton de Gruyter.

----------. 1991. “What's cognitive, what's not, in sound change”. *Diachrony within Synchrony: Language History and Cognition* ed. by Günter Kellermann & Michael D. Morrissey. 309-355. Duisburg: Peter Lang Verlag.

Pierrehumbert, Janet, Mary E. Beckman & D. Robert Ladd. 2001. “Conceptual foundations of phonology as a laboratory science”. *Phonological Knowledge: Conceptual and Empirical Issues* ed. by Noel Burton-Roberts, Philip Carr & Gerard Docherty. 273-303. Oxford: Oxford University Press.

PART 1

SEGMENTS AND PROCESSES

DETECTION OF LIAISON CONSONANTS IN SPEECH PROCESSING IN FRENCH

EXPERIMENTAL DATA AND THEORETICAL IMPLICATIONS*

NOËL NGUYEN[1], SOPHIE WAUQUIER-GRAVELINES[2], LEONARDO LANCIA[1] & BETTY TULLER[3]

[1]*Laboratoire Parole et Langage, CNRS & Université de Provence,* [2]*Structures formelles du langage, CNRS et Université de Paris VIII,* [3]*Center for Complex Systems and Brain Sciences, Florida Atlantic University*

Abstract

The goal of the present study is to better understand the mechanisms involved in the processing of liaison consonants by listeners in French. Previous work (Wauquier-Gravelines 1996) showed that liaison consonants are more difficult to detect than word-initial consonants in a phoneme-detection task. We examined to what extent such differences are attributable to the consonants' phonetic properties, and we also compared the perception of liaison consonants with that of fixed word-final and word-medial consonants, as well as word-initial ones. The results suggest that liaison consonants have a specific perceptual status. Implications for both autosegmental and exemplar-based theories of liaison are discussed.

1. *Introduction*

French liaison is a well-known phenomenon of external sandhi that refers to the appearance of a consonant at the juncture of two words, when the second word begins with a vowel, e.g. *un* [œ̃] + *enfant* [ɑ̃fɑ̃] → [œ̃nɑ̃fɑ̃] "a child", *petit* [pəti] + *ami* [ami] → [pətitami] "little friend". Liaison consonants are usually *enchaînées*, i.e. realized as syllable-onset consonants, although they can also appear in coda position, compare [pə.ti.ta.mi] (with *enchaînement*) and [pə.tit.a.mi] (without *enchaînement*, Encrevé 1988). In the following, the two words at the juncture of which liaison consonants appear will be referred to as Word 1 and Word 2, respectively.

* This work was partly supported by the *ACI Systèmes complexes en SHS* Research Program (CNRS & French Ministry of Research) and by NSF Grant #0414657. We thank Sharon Peperkamp and Stéphanie Ducrot for drawing our attention to the missing-letter effect. We are also grateful to Robert Espesser for sharing his statistical expertise, and to Pierre Encrevé, Zsuzsanna Fagyal, Cécile Fougeron, Mariapaola D'Imperio, Maria-Josep Solé, Marina Vigário, and three anonymous reviewers for useful comments.

Among the many different approaches to French liaison that have been proposed over the last thirty years or so (see Tranel 1995; Côté 2005, for reviews), a major bone of contention relates to whether liaison is a phonological or a lexical phenomenon. The phonological approach dates back to early generative studies on French phonology, in which liaison was seen as an exception to a general process of final consonant deletion, referred to as the 'French Truncation Rule' by Schane (1968). By contrast, according to another proposal made later in the same general framework (e.g. Klausenburger 1974, 1977), liaison consonants arose in the course of the derivation owing to an insertion mechanism (views differed as to whether this epenthesis occurred at the end of Word 1 or at the onset of Word 2). More recent treatments of liaison in nonlinear phonology have reconceptualized the deletion/insertion dichotomy, as pointed out by Tranel (1995). Thus, in the autosegmental account proposed by Encrevé (1988) and Encrevé and Scheer (2005), liaison consonants are viewed as floating segments, with respect to both the segmental and syllabic tiers. Such consonants must be associated with both tiers to be phonetically realized, and this association takes place only under certain conditions. In both the linear and nonlinear phonological approaches, liaison is generally portrayed as being subjected to prosodic, morphological, syntactic and stylistic factors.

Lexical approaches to liaison can be divided into two main strands. Suppletive analyses as advocated by Klausenburger (1984) among others, assume that words such as *petit* are associated in the lexicon with two distinct allomorphs, a longer one ending in a liaison consonant (/pətit/) and a shorter one without liaison consonant (/pəti/). In contrast, in exemplar-based models, such as the one recently proposed by Bybee (2001, 2005), liaison consonants are said to take place within specific grammatical constructions, e.g. [NOUN -z- [vowel]-ADJ]$_{\text{Plural}}$, in *enfants intelligents* [ɑ̃fɑ̃zɛ̃teliʒɑ̃] "clever children". Constructions display different degrees of generality/abstractness, and range on a continuum from very abstract (as in the example given above), to fixed, lexicalized phrases like *c'est-à-dire* [setadiʁ] "that is to say". This provides a unified account of both false liaisons, which are attributed to the overgeneralization of a high-frequency construction, as in *quatre enfants* [katʁəzɑ̃fɑ̃] "four children", and word-specific differences in the realization of liaison. Frequency of use is of central importance, as liaison is assumed to occur more often within a sequence of words characterized by a higher frequency of co-occurrence. This approach is neutral with respect to the issue of whether liaison consonants result from a deletion or insertion process, nor does it make any specific claim as to whether the consonant belongs to Word 1 or 2.

As noted above, liaison consonants when realized are usually *enchaînées*, i.e. syllabified into onset position. This results in a mismatch between word and syllable boundaries. Specifically, the syllable whose onset position the liaison consonant comes to occupy straddles the boundary between Word 1 and Word 2 (e.g. [pə.ti.ta.mi], where the word boundary takes place between [t]

and the following [a]). Recent psycholinguistic studies (Wauquier-Gravelines 1996; Gaskell, Spinelli & Meunier 2002; Spinelli, McQueen & Cutler 2003) have shown that this mismatch does not necessarily make it more difficult for listeners to identify the second word, but may in fact facilitate the recognition of that word with respect to a baseline condition. This raises questions for models of speech perception in which the syllable is viewed as a primary unit of segmentation in lexical access in French (see Content, Kearns & Frauenfelder 2001, for a recent discussion in that domain).

A related issue concerns the way in which liaison consonants are processed in speech perception. One may ask which perceptual mechanisms allow a *liaison enchaînée* to be distinguished from word-initial as well as word-final consonants, and to which of the two words the liaison consonant is associated by the listener. More generally, the question arises whether in speech comprehension liaison consonants are processed and represented in a way that is different from fixed consonants. It is this issue that is addressed in the present paper. A series of experiments are reported which together suggest that liaison consonants do have a distinct perceptual status. Implications for current models of liaison in French will be discussed.

2. *Empirical evidence for a specific status of liaison consonants in speech perception*

Wauquier-Gravelines (1996) examined the speed and accuracy with which listeners can detect the presence of a liaison consonant in the speech chain. Because it has not been published, we present this work here in some detail and in light of more recent findings. Wauquier-Gravelines compared listeners' responses to liaison consonants and word-initial consonants in a phoneme detection task. Listeners were presented with a series of sentences and were asked to detect a pre-specified phoneme in each sentence. The material contained pairs of sentences that were designed so that the target phoneme appeared as a word-initial consonant in one sentence (e.g. *son navire* [sɔ̃naviʁ] "his ship") and as a liaison consonant (e.g. /n/ in *son avion* [sɔ̃navjɔ̃] "his plane") in the other. The two sentences in each pair were matched with respect to their syntactic, lexical and phonemic make-up. A number of filler sentences were also used. Both the test and filler sentences were recorded by a native speaker of standard French.

Two experiments were conducted. Each experiment was comprised of a training phase and a test phase. The target consonant was /t/ in the first experiment and /n/ in the second one. There were fourteen subjects, all native speakers of standard French, with no known hearing impairment, and naive as to the purpose of the experiment.

The data showed that listeners experienced greater difficulties in detecting the liaison than the word-initial consonant. There were significantly fewer correct responses for the liaison than for the word-initial consonant for both /t/ (liaison: 67.8%, word-initial: 92.8%, $\chi^2 = 9.56$, $p<0.01$) and /n/

(liaison: 44.6%, word-initial: 87.5%, χ^2 = 21.07, $p<0.01$), although this difference was smaller for /t/ than /n/.

These results suggest that liaison consonants are not processed in the same way as fixed consonants by listeners. There is a potential parallel between this phenomenon and the status liaison consonants have in autosegmental phonology. As indicated above, liaison consonants display both syllabic and skeletal flotation in Encrevé's (1988) autosegmental model. When followed by a word with a null onset (i.e. an onset with no corresponding segmental constituent and no skeletal slot), the liaison consonant is attributed a skeletal slot and, in the unmarked case, is syllabified into onset position. Thus, liaison consonants are not lexically anchored to a timing unit and are in this regard characterized by structural instability. It may be hypothesized that listeners' behaviour in the phoneme-detection experiments is a reflection of this instability. In other words, it would be more difficult for listeners to map a liaison consonant onto a phonemic label because unlike 'ordinary' phonemes, i.e. fixed consonants, liaison consonants are underlyingly floating with respect to the skeleton associated with the word to which they belong. The absence of a pre-established link between the liaison consonant and one of the available timing units in the underlying lexical representation would make that consonant harder to detect in an explicit manner.

This phenomenon is reminiscent of Sapir's (1933) observation that speakers of British English are convinced they do not pronounce *sawed* and *soared* in the same way, because *soared* is viewed as underlyingly containing an *r*, even though both words may be phonetically transcribed [sɔ:d] (in non-rhotic varieties of BE). In Encrevé's model, the difference between *sawed* and *soared* is attributed to the presence of a floating *r* in the latter but not in the former. Likewise, Wauquier-Gravelines' (1996) findings may suggest that a liaison consonant is perceived by listeners in a way that mirrors its specific phonological status as a floating segment. In other words, syllabic/skeletal flotation may be perceptually and cognitively relevant.

Although differences in the phonological status of liaison and word-initial consonants thus provide an appealing explanation for the observed perceptual patterns, other factors such as the frequency of occurrence of the Word 1-2 sequences, the target's acoustic properties, and the target's position within the carrier word, may have also played a role. We begin with the issue of lexical frequency.

It might be the case that liaison consonants appeared in a context that rendered them less predictable by listeners than word-initial consonants. Recent studies (e.g. Adda-Decker, Boula de Mareüil & Lamel 1999; Fougeron, Goldman & Frauenfelder 2001; Fougeron, Goldman, Dart, Guélat & Jeager 2001) suggest that the realization of liaison is partially conditioned by a complex interplay between the lexical frequencies of Words 1 and 2. Specifically, Fougeron, Goldman and Frauenfelder (2001) found that the rate of realization of liaison shows both a positive correlation with the frequency of Word 1, and a small, but significant, negative correlation with the frequency of

Word 2. Fougeron et al.'s results also revealed that the rate of liaison increases with the frequency of co-occurrence of the two words. In Wauquier-Gravelines' experiments, however, potential lexical frequency effects were fully neutralized for Word 1 since that word was identical for both sentences in each sentence pair. In addition, Fougeron, Goldman and Frauenfelder (2001) point out that because high-frequency words are often short function words, the relationship found between frequency of Word 2 and rate of liaison may actually reflect the fact that liaison is realized less often before short function words than before longer words. Since Wauquier-Gravelines only used nouns and adjectives (most of them di- or trisyllabic) in Word 2 position, it seems unlikely that, in her material, liaison consonants had a lower probability of occurrence than word-initial consonants. Note also that words starting with a vowel are much more numerous in French than consonant-initial words with either of the two target consonants used in the experiments, /t/ or /n/. In such contexts, listeners should have been biased towards identifying the target as a liaison, rather than a word-initial consonant. This again suggests that the lower detection rate obtained for the liaison consonant was not related to the frequencies of occurrence associated with both targets.

Let us now turn to the target consonant's acoustic properties. Differences may arise in that domain between liaison and word-initial consonants, which would make the former less perceptually salient than the latter. Such differences have indeed been found in the vicinity of the consonant in previous work (e.g. Delattre 1940; Dejean de la Bâtie 1993; Gaskell et al. 2002; Spinelli et al. 2003). Thus, Dejean de la Bâtie (1993) found that the duration of the closure and that of the following burst are both shorter for liaison /t/ compared with word-initial /t/. In Gaskell et al. (2002), the duration of /t/, /r/ and /z/ also proved to be on average slightly but significantly shorter in liaison (73 ms) than in word-initial position (88 ms; consonant duration was taken as the time interval between the offset of the preceding vowel and the onset of the following vowel). A similar durational difference was found between liaison (64 ms) and word-initial consonants (71 ms) by Spinelli et al. (2003), for /p, r, t, n, g/. Note that the shorter duration for liaison consonants reported in the above studies could be due to actual liaison shortening and/or word-initial lengthening (Fougeron 2001).

Wauquier-Gravelines carried out a series of acoustic analyses on sentences analogous to those she used as stimuli in the two experiments reported above. For /t/, she found that the closure and burst had a significantly shorter duration in *liaison enchaînée* (mean overall value: 50 ms) than in word-initial position (70 ms), in keeping with previous findings. For /n/, however, the acoustic duration of the consonant was not found to be statistically different in *liaison enchaînée* (58 ms) and word-initial position (61 ms). Thus, it seems that variations in duration in liaison vs word-initial position are both subtle and specific to certain consonants (possibly obstruents) only. Although such data suggest that the observed differences in the listener's responses to the liaison and word-initial consonants are not related to how these two types of

consonant are phonetically realized, this issue will be taken up again in the next section.

Yet another factor that may have contributed to making the liaison consonant less easily detectable than the word-initial consonant relates to the position that these consonants occupied in the carrier word. In the phonological approach espoused by Encrevé (1988), among others, the liaison consonant lexically belongs to Word 1 and occurs in final position in that word. Because a greater perceptual weight is attributed to word onsets compared with word offsets in sequential models of word recognition such as Cohort (Marslen-Wilson & Zwitzerlood 1989), it may be speculated that the word-initial consonant was perceptually more prominent than the liaison consonant. Thus, to test the hypothesis that the lower detection rate for the liaison consonant is attributable to syllabic/skeletal flotation, rather than position in the word, it would be necessary to include word-final fixed consonants in the potential targets, and to show that listeners' responses are more accurate for these consonants than for liaison consonants. In Encrevé's model, so-called final fixed consonants are characterized by the fact that the corresponding coda constituent on the syllabic tier is floating with respect to the skeleton. This allows the model to account for the *enchaînement* of final fixed consonants prior to a vowel-initial word. A crucial difference between final fixed and liaison consonants, however, is that only the former are anchored to the skeleton.

Wauquier-Gravelines' material was not designed to undertake systematic comparisons between listeners' responses to liaison and final fixed consonants. These methodological issues were addressed in the experiment described in the following section.

3. *Further evidence on the specific perceptual status of liaison consonants*

The goal of this experiment was to confirm and extend Wauquier-Gravelines' findings in two directions. First, we examined to what extent differences in the detection rate of liaison consonants vs word-initial consonants are attributable to the phonetic properties of these consonants, by systematically manipulating these properties. Second, the potentially distinctive status of liaison consonants compared with fixed consonants in perception was further explored by inserting fixed word-final and word-medial consonants, as well as word-initial ones, in the material.

3.1 *Method*

3.1.1 *Material.* The material was made up of twenty sets of four test sentences. These sentences contained a target consonant which appeared in the vicinity of the boundary between two words. The target consonant was /z/ for twelve of the twenty sets and /n/ for the remaining sets. Within each set, the target consonant was located at the onset of Word 2, at the end of Word 1, in word-medial position, and as a liaison consonant at the juncture between Words 1

and 2. As an example, the position of the target /z/ in each of the four sentences for one of the sentence sets is shown in Table 1. The critical words are underlined in the orthographic transcription. A phonemic transcription of these two words is also shown, with the target consonant displayed in bold.

Sentence type	Target position	Example
1	W2-initial	*Il y a <u>des zéros</u>* /de**z**ero/ *partout dans le tableau.* "There are zeros everywhere in the table".
2	W1-final	*On a eu <u>seize élèves</u>* /sɛ**z**elɛv/ *qui ont réussi au bac* "Sixteen pupils of ours have passed the baccalaureate exam".
3	Word-medial	*J'ai rapporté <u>du raisin</u>* /dyrɛ**z**ɛ̃/ *du marché ce matin.* "I brought some grapes back from the market this morning".
4	Liaison	*J'ai remis <u>des écrous</u>* /de**z**ekru/ *en haut du radiateur.* "I put some nuts back on top of the radiator".

Table 1: *Position of the target consonant* /z/ *in each of the four sentences, for one of the twenty sentence sets.*

In all cases, liaison consonants appeared in an unmarked context which made their pronunciation obligatory: determinant + noun (e.g. *des* [z] *écrous* "nuts"), adjective + noun (e.g. *lointain* [n] *ami* "distant friend"), monosyllabic adverb (e.g. *très* [z] *ému* "very touched") or preposition (e.g. *en* [n] *Asie* "in Asia") before another word.

In addition, both Type-1 and Type-4 sentences were locally ambiguous as to the morpho-phonological status of the target consonant, i.e. the first part of the sentence, up to the post-consonantal vowel, was in both cases consistent with the consonant being a W2-initial or a liaison consonant. This is true, for example, of the W2-initial [z] in *Il y a des* [z] *zéros* "There are zeros" (where the morpho-syntactic and phonological make-up of the first part of the sentence up to the post-consonantal vowel may allow the listener to interpret [z] as a liaison consonant, until the following word is identified) and, reciprocally, of the liaison [z] in *J'ai remis des* [z] *écrous* "I put some nuts back" (where the first part of the sentence up to the post-consonantal vowel could lead to [z] being temporarily interpreted as the initial consonant of the upcoming word by the listener). Importantly, for most Type-1 sentences, Word 1 contained a liaison consonant whose realization would be obligatory prior to a word-initial vowel. For example, the liaison consonant /z/ associated with the determinant *des* in *des zéros* is obligatorily pronounced when the following word begins with a vowel. [There were only two exceptions to this. In *les délégués zaïrois* "the Zairian delegates" (plural noun + adj., target cons.: word-initial /z/), the realization of the liaison consonant /z/ at the end of *délégués* prior to a word-initial vowel is optional. In *un bien naturel* "a natural resource" (sing. noun + adj., target cons.: word-initial /n/), the realization of a liaison /n/ at the end of *bien* used as a noun before a word-initial vowel, is excluded. The

corresponding Type-4 sequences are *les avis* “the notices” (det. + noun, liaison /z/) and *bien appris* “well learned” (adv. + part participle, liaison /n/), respectively.] Such constructions allowed us to ensure that the listeners could not predict whether the target consonant was a W2-initial or liaison consonant from the preceding words in the sentence.

All sentences had about the same number of syllables (mean = 13, s.d. = 1.4) and the rank of the word in which the target consonant appeared was approximately the same across sentences (average rank, from the beginning of the sentence = 4.5 words, s.d. = 1.1). The target-bearing word was as short as possible and contained two syllables on average (s.d. = 0.6) in Type-1 sentences, one syllable (s.d. = 0) in Type-2 sentences, two syllables (s.d. = 0.2) in Type-3 sentences, and one syllable (s.d. = 0.2) in Type-4 sentences. The purpose of using such short words was to minimize the possibility for the target consonant to be anticipated by the listener in Type-2, -3 and -4 sentences.

The pre- and post-target vowels were as phonetically similar as possible across the four sentences in each set, differing from each other by at most one distinctive feature (in a standard distinctive-feature system) for most sets. The pre-target vowel itself was preceded by a consonant (e.g. /d/ in *des zéros*) on which two constraints were imposed for Type-1 and Type-4 sentences. First, consonants appearing in that position in the two sentences had to share as many phonetic properties with each other as possible. Second, whenever possible we used consonants characterized by a well-defined acoustic transition with the following vowel, such as voiceless obstruents. These constraints were motivated by the splicing procedure to which Type-1 and Type-4 sentences were later subjected (see below). A further phonetic constraint was that the sounds preceding the target consonant were as different from the target as possible, to avoid any perceptual interference (Stemberger, Elman & Haden 1985).

In addition, the sentences had similar syntactic structures, and Word 2 was chosen to be as semantically unpredictable as possible from the first part of the sentence (on the basis of the first and second authors’ intuitions as native speakers of French).

Finally, we constructed 240 filler sentences (120 without /z/ and 120 without /n/), which were similar to the test sentences with respect to overall length and syntactic structure. Furthermore, part of the words occurring in Word 1 position in Type-1 and Type-4 test sentences also appeared in the filler sentences prior to a word-initial consonant that differed from the target in the test sentences, e.g. *des crêpes* /dekrɛp/ “pancakes”. This means that these words were not systematically associated with the presence of the target consonant in the material, and that the listeners were thus prevented from developing a response strategy based on learning such an association over the course of the experiment (thus, *des* was not always followed by /z/, whether as a word-initial or liaison consonant).

3.1.2 *Speaker, recording and acoustic labelling*. The material was recorded by the first author, whose speech can be characterized as intermediate between Southern and standard French. In particular, this speaker does not pronounce word-final schwas, as is the case in Southern French (see Nguyen & Fagyal 2006, for further details). The recording took place in a sound-proof room using high-quality recording equipment (sampling frequency = 22050 Hz). The speaker first read the list of test sentences five times, then the filler sentences. Both the test and filler sentences were randomized. The speaker's task was to read the sentences naturally, while maintaining the same rate, rhythm and pitch contour throughout the corpus.

The acoustic data were transferred onto a personal computer for further processing. For each test sentence, markers were placed at the acoustic onset and offset of each segment in each V-target C-V sequence. The location of these acoustic boundaries was determined from both the digital speech waveform and a corresponding wideband spectrogram.

3.1.3 *Stimuli and experimental design*. The initial set of stimuli consisted of the 80 test sentences and 240 filler sentences. For each of the Type-2 and Type-3 sentences, one repetition out of the five available was selected, which we judged as being articulated fluently, clearly, and at a normal rate. In addition, two different versions of Type-1 and Type-4 sentences were created. In the *identity-spliced* version, the target consonant and preceding vowel originated from another repetition of the same sentence. In the *cross-spliced* version, the target consonant and preceding vowel came from either the Type-1 or Type-4 corresponding sentence, for Type-4 and Type-1 sentences, respectively. To construct the identity- and cross-spliced stimuli, we selected those among the five available repetitions per sentence which allowed the vowel + consonant sequence to be spliced into the carrier sentence with no audible discontinuities across the splicing points. As for Type-2 and Type-3 sentences, fluency, clarity of articulation and rate were also taken into consideration. Although the consonant's duration and that of the preceding vowel did not significantly differ when the consonant was in W2-initial compared with liaison position (as reported in Section 3.2.1 below), variations related to the consonant's position may be shown in the vicinity of that consonant by other acoustic parameters. Cross-splicing allowed us to assess the perceptual relevance of such potential acoustic variations. These were expected to result in a lower target detection rate and/or a longer reaction time in the cross-spliced sentences than the identity-spliced sentences, which we used as a baseline condition.

The experimental task was a speeded phoneme detection task, with two targets, /n/ and /z/. Thirty-four native speakers of French with no known hearing deficit participated and were partitioned into two main groups. The stimuli were blocked by target, and the order of presentation of the targets was counterbalanced across groups. Test and filler sentences were fully randomized within each block. The two subject groups were further divided into two

subgroups. For each of the Type-1 and Type-4 sentences, one subgroup was presented with the identity-spliced version and the other with the cross-spliced version. Which subgroup heard the identity-spliced vs cross-spliced version systematically changed from one sentence to the next. In this way, each subject heard each sentence only once, either the identity-spliced (for half of the sentences) or the cross-spliced version (for the other half). One of the four subgroups contained ten subjects and the others had eight subjects. The stimuli were played over headphones at a comfortable sound level. Subjects had to press a button on a response box, using their dominant hand, if and as soon as they detected the target in the sentence. Reaction time was measured from the acoustic onset of the target phoneme. The test phase was preceded by a short training phase with ten sentences. The experiment lasted about thirty minutes, and each subject received a small fee for her/his participation.

3.2 *Results*

3.2.1 *Durational measurements*. In a first attempt to characterize the acoustic properties potentially associated with the target consonant depending on its position and phonological status, we measured the duration of that consonant, along with that of the preceding vowel. Figure 1 shows the average duration for each segment in each of the four types of sentence. Repeated-measure ANOVAS revealed that duration significantly varied as a function of sentence type for /z/ ($F(3,33) = 6.282$, $p<0.01$) and the preceding vowel ($F(3,33) = 17.669$, $p<0.001$), as well as for /n/ ($F(3,21) = 3.185$, $p<0.05$) and the preceding vowel ($F(3,21) = 7.101$, $p<0.01$). Scheffé post-hoc tests showed that the duration of /z/ was significantly longer in W2-initial position than in W1-final position ($p<0.01$). In addition, and for both /z/ and /n/ sentences, the preceding vowel's duration was significantly longer in W1-final than in W2-initial (/z/ sentences: $p<0.001$; /n/ sentences: $p<0.01$), word-medial (/z/ sentences: $p<0.01$; /n/ sentences: $p<0.05$), and liaison position (/z/ sentences: $p<0.001$; /n/ sentences: $p<0.01$). Pairwise comparisons between the mean values associated with the four types of sentence yielded no significant difference for /n/ duration.

To summarize, vowels in word-final closed syllables were longer than vowels in other positions and /z/ was longer when it appeared in onset position in word-initial syllables as opposed to coda position in word-final syllables. Importantly, however, the comparison between W2-initial and liaison positions, which formed the main focus of interest in this work, revealed no significant difference in the duration of either the target consonants or the preceding vowel. Note that this is not consistent with the tendency for consonants to be shorter in liaison than in W2-position reported previously (Dejean de la Bâtie 1993; Gaskell et al. 2002; Spinelli et al. 2003). This may be due, at least in part, to the phonetic make-up of the material used in each study. Dejean de la Bâtie's (1993) analyses focused on /t/; the present work examines /z/ and /n/. The two other studies used a variety of target consonants that included /z/ (Gaskell et al. 2002) and /n/ (Spinelli et al. 2003), but it is unclear

to what extent /z/ and /n/ actually contributed to the observed position-dependent differences in duration because the authors only provide mean duration values across all target consonants. A more relevant comparison is with Wauquier-Gravelines (1996), who measured the duration of /n/ in liaison vs W2-initial position, and, as in the present study, found no significant difference between the two.

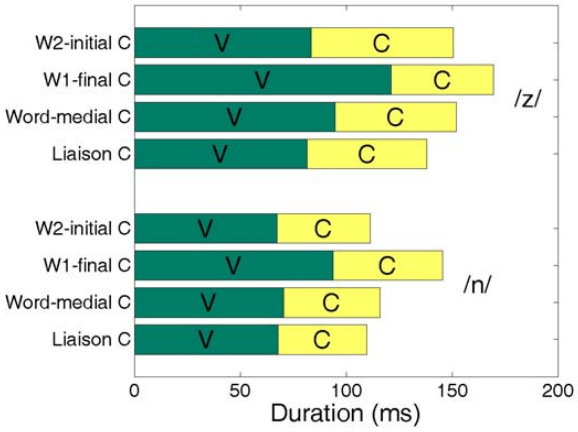

Figure 1: *Average duration of the target consonant and pre-consonantal vowel as a function of consonant position, for* /z/ *and* /n/.

3.2.2 *Perceptual data pre-processing*. Data from one subject out of the thirty-four were omitted due to the unusually high error rate (61%); data from two other subjects were omitted because their mean reaction times were more than two standard deviations above the overall mean RT. After these exclusions, the four subgroups of subjects contained seven, eight, nine and seven members. For these thirty-one subjects, the proportion of correct detections ranged from 65% to 93% over both targets, and the mean reaction time ranged from 538 ms to 1396 ms. There was a significant negative correlation between percent correct detection and mean RT per subject ($R^2 = 0.36$, $t(29) = -4.02$, $p<0.001$), i.e. subjects who tended to miss the target more often were also slower to respond when they did detect the target.

3.2.3 *Target detection rates*. To assess the effect of cross-splicing on phoneme detection, a by-subject repeated-measures ANOVA was carried out, with target identity, splicing type (identity-spliced vs cross-spliced) and position (W2-initial, liaison) as independent variables and percent correct detection as the dependent variable. All of the independent variables were within-group factors. The experimental design allowed us to put these three independent variables together in a by-subject ANOVA but not in a by-item ANOVA. The analysis was restricted to the W2-initial and liaison positions since cross-splicing was performed for these two positions only. Percent correct detection was