3

An integrated approach to French liaison

Olivier Bonami, Gilles Boyé, Jesse Tseng

Consonant liaison at word boundaries in French is the result of a complex interplay of grammatical and extragrammatical factors. In this paper we offer a descriptive overview of syntactic factors influencing liaison. We provide a detailed analysis in the framework of HPSG, that integrates the morphophonological and syntactic conditions governing this feature of French grammar. Although we do not directly model other factors influencing liaison (such as frequency effects, prosodic considerations, or sociolinguistic variables) our analysis is modular enough to accommodate additional conditions resulting from more complete empirical studies.[1]

3.1 A descriptive overview of liaison

Many French words come in two shapes, which we will refer to as the "short form" and the "long form".[2] The short form is always used before consonants and in utterance-final position; the long form is used, in some cases, when the following word is vowel initial. The term *liaison* refers to the realization of the long form in appropriate contexts.[3]

---

[1]We thank Anne Abeillé, Elisabeth Delais-Roussarie, Danièle Godard and three anonymous *Formal Grammar* referees for their comments and suggestions.

[2]By "words" we mean fully inflected lexical items, as opposed to lexemes or paradigm headwords (which can correspond to many distinct surface forms).

[3]A few isolated words, such as *six* 'six', have two long forms, one for liaison contexts ([siz]) and one for utterance-final position and certain other contexts ([sis]).

We separate two issues in the grammar of liaison: the identification of the contexts where liaison is possible and the relationship between short and long forms. In the simplest cases of short/long form alternation, the long form is just the short form with an additional, so-called "latent" final consonant. For instance, the adverb *très* 'very' has a short form [tʁɛ], found e.g. in *très doué* 'very gifted' [tʁɛdwe], and a long form [tʁɛz] found e.g. in *très intelligent* 'very intelligent' [tʁɛzɛ̃tɛliʒɑ̃]. In section 3.1.1 we restrict our attention to this basic type of short/long form alternation, to examine the syntactic contexts where liaison is possible and/or mandatory. We defer the examination of more unusual short/long form pairs until section 3.1.2.

3.1.1 Syntactic contexts for liaison

Given a word w_1 with distinct short and long forms and a following vowel-initial[4] word w_2, liaison is observed to be obligatory (indicated by $w_1 = w_2$), optional (indicated by $w_1 \circ w_2$), or impossible (indicated by $w_1 \neq w_2$). This is reflected in prescriptive manuals of French, which include long, but rather arbitrary, lists of "correct" and "incorrect" liaisons (Delattre, 1966, Fouché, 1959). Early formal approaches to liaison concentrated on finding appropriate generalizations about the domain of liaison. Selkirk (1972, 1974) suggests that the domains of obligatory and optional (or "stylistically elevated") liaison can be determined by means of simple, cross-categorial rules referring to explicitly defined notions like "phonological word" and "inflected lexical head". Morin and Kaye (1982), however, present many counterexamples to Selkirk's generalizations, and conclude that the status of liaison (obligatory, optional, impossible) must more or less be stipulated on a construction-by-construction basis. In later work, liaison has been used as a prime argument for distinguishing syntactic constituency from prosodic constituency: according to Selkirk (1986), obligatory and optional liaison occur within two different kinds of prosodic constituents, determined by applying an algorithm to syntactic structures that results in new constituent boundaries. Later work has shown that obligatory liaison is much more limited than the prosodic approach suggests (de Jong, 1994), and that liaison domains do not coincide with independently identifiable prosodic domains (Post, 2000).

At the same time, many studies have emphasized the importance of sociolinguistic factors in the realization of liaison (see Encrevé 1988 for a detailed discussion), the lexical conditioning of many liaison contexts (Tranel, 1981, de Jong, 1994, Morin, 1998) and the influence of

---

[4]For the time being we ignore the problem of *h aspiré*; see section 3.2.1.

frequency (Bybee, 2001) or prosodic factors (Fougeron et al., 2001) on the actual realization of liaison. Thus the current consensus is that there is a wide variety of factors involved in liaison, and few (if any) studies attempt to provide actual grammar fragments predicting the contexts where liaison is realized.

Our goal in this paper is to present a grammar fragment that incoporates genuinely syntactic constraints on liaison realization. Using observations taken from the previous literature and new evidence, we outline an updated list of the contexts where syntax forces liaison or makes it impossible. We assume that optional liaison is the default situation, and that specific syntactic environments make liaison impossible or mandatory in particular cases.[5]

What follows is a survey of the main phrasal structures of French; in anticipation of our formal analysis, we are guided by the description of French phrase structure in HPSG of Abeillé and Godard (2000, 2002).

In subject-head combinations, liaison is impossible between the daughters, irrespective of the head's category (compare (1a) and (1b)). Note that by contrast, liaison is obligatory between a weak form subject pronoun and the verb (2). This is part of the motivation for analyzing these pronouns as affixes on the finite verb rather than daughters in a syntactic combination (Miller and Sag, 1997, Miller, 1992).

(1)   a.   [Les enfants≠ont mangé].
            'The children have eaten.'

      b.   [Les enfants≠au lit], on servit le dessert.
            (With) the children in bed, dessert was served.'

(2)   [Ils=étaient contents].
      'They were happy.'

Liaison is obligatory between a specifier and the following head, as shown in (3).[6]

(3)   a.   mon=ami
            'my friend'

---

[5]It is important to interpret "optionality" in this paper as the absence of any *synactic* condition on the realization of liaison in a particular phrasal configuration. Lexical and other factors typically intervene to make liaison more or less likely in specific instances of this configuration, even to the point of making it obligatory or impossible. As a simple example, liaison is possible in *enfants [z] intelligents* 'intelligent children' but not in **enfant [t] intelligent* 'intelligent child' ; it is nearly always realized in *très [z] intelligent* 'very intelligent', while other adverbs in the exact same structure give rise to liaison less systematically.

[6]Note that in our analysis of French, [Det N$'$] combinations are the only instances of specifier-head combinations.

    b. [mon=[ancien collègue]]
       'my former colleague'

Head-complement structures in French are uniformly head-initial, and liaison is possible between the head and the first complement (4). There may be several complement daughters; liaison is only possible between a "lite" complement and a following complement (5a,b). Lite elements include the pronouns *tout* and *rien* and past participles in compound tenses (Abeillé and Godard, 2000, 2002). Other complements cannot give rise to liaison (6).[7]

(4) a. Paul [pensait∘à Marie].
        'Paul was thinking about Marie.'

      b. [dans∘une semaine]
        'in one week's time'

(5) a. Paul [donnera tout∘à Marie].
        'Paul will give everything to Marie.'

      b. Paul [a été mis∘à pied].
        'Paul has been put on suspension.'

(6) Jean [présentera ses enfants≠à Marie].
     'Jean will introduce his children to Marie.'

In head-adjunct combinations, liaison is generally possible between the head daughter and the adjunct daughter, which exhibit both possible word orders (7).

(7) a. [amis∘intimes]
        'close friends'

      b. [très∘intéressant]
        'very interesting'

      c. [bien∘équipé]
        'well equipped'

Prenominal attributive adjectives deserve special attention. It is traditionally assumed that liaison is obligatory between a prenominal adjective and the noun (e.g., *petit enfant* 'small child'), while it is optional between the noun and a postnominal adjective (7a). However Post (2000) provides decisive evidence that while liaison is more frequent prenominally, it is in fact optional in both cases (see Morin and

---

[7]Modifying adverbs realized among complements in the VP may give rise to liaison (even when they are not lite), as noted by Morin and Kaye (1982). We leave these aside since their status (adjunct vs. complement) is controversial.
Remember that we assume that pronominal clitics are affixes, not words; thus clitic liaison is a lexical phenomenon which falls outside of the scope of this paper.

Kaye (1982) and de Jong (1994) for earlier hints to this effect). In a reading task, Post observed that subjects realized liaison only 88% of the time with plural prenominal adjectives. This is especially significant since (i) liaison is more often realized in reading than in conversation (Fougeron et al., 2001), and (ii) the liaison rate is highly dependent on the choice of the noun and the adjective (falling to 61% for certain pairs), a situation typical of optional liaison contexts. We thus conclude that attributive adjectives conform to the general case of head-adjunct combinations, where liaison is not obligatory but only quite frequent.[8]

Liaison in filler-head combinations seems to be impossible (8a,b). Apparent counterexamples are predicative *quel(le)s*, which exhibits obligatory liaison (8c), and *dont*, with which liaison is optional (8d). But independent evidence shows that these items are not fillers: Comorovski (to appear) argues that predicative *quel* is a clitic combining directly with the verb; and obligatory liaison is just what we expect if *quel* combines with the verb in the lexicon. *Dont* is arguably a complementizer rather than a *wh*-word, since it does not give rise to pied-piping (9), and (ii) it cannot be followed by the complementizer *que* in varieties that allow this with *wh*-items (10); thus *dont* relatives are head-complement structures, and liaison is expected to be optional.

(8) a. Quelles tartes≠ont-ils mangées ?
        'Which pies did they eat?'

    b. les enfants [auxquels≠elle a parlé]
        'the children to whom she spoke'

    c. Quels [z] étaient les enjeux ?
        'What were the issues at stake?'

    d. l'homme dont∘il a parlé
        'the man he spoke about?'

(9) a. Voilà l'homme au frère de qui j'ai parlé.
        'Here is the man to whose brother I spoke.'

    b. *Voilà l'homme au frère dont j'ai parlé.

(10) a. % Voilà l'homme à qui que j'ai parlé.
        'Here is the man I spoke to.'

    b. *Voilà l'homme dont que j'ai parlé.
        'Here is the man I spoke about.'

---

[8]Note that Post provides data only for *plural* prenominal adjectives; in the absence of relevant evidence we suppose that the same situation holds in the masculine singular.

In coordinations, liaison is generally possible between the penultimate conjunct and the conjunction (11a) and between the conjunction and the final conjunct (11b). However, when there are more than two conjuncts, liaison is impossible between adjacent conjuncts (11c).

(11) a. [petits∘et grands]
          'small and large'
      b. [gentil mais∘idiot]
          'nice but dumb'
      c. livres [petits≠abîmés∘et chers]
          'small, damaged, and expensive books'

Coordination reveals some previously overlooked liaison data. As we have just seen, liaison is in general optional before a conjunction, and so we can use coordination to identify further constraints associated with the right edge of particular phrasal combinations. For example, liaison remains possible when the conjunct preceding the conjunction is a specifier-head (12) or adjunct-head (13) combination (the last example also illustrates optional liaison after a coordinated structure).

(12) [les amis]∘et les collègues de Marie
      'Marie's friends and colleagues'

(13) [[très bien]∘et très chaleureusement]∘accueilli
      'very well and very warmly received'

What is more surprising is that liaison is blocked when the conjunct is a subject-head (14), filler-head (15), or head-complement (16) combination (note that liaison is blocked even when the final complement is *lite*).This suggests that phrase types constrain the possibility of liaison not only between their daughters, but also between the phrase as a whole and following material.[9]

(14) [Paul dormait]≠et Marie travaillait.
      'Paul was sleeping and Marie was working.'

(15) [Qui dormait]≠et qui ne dormait pas ?
      'Who was sleeping and who wasn't?'

(16) a. Paul doit [acheter ces livres]≠ou les emprunter.
          'Paul must buy those books or borrow them.'
      b. impressionné [par les arguments]≠et par les exemples
          'impressed by the arguments and by the examples'

---

[9]The constraints on filler-head and subject-head combinations probably follow from a more general constraint against liaison between a clause and the following material. Such a constraint is necessary to block liaison after a single-word clause: *Sortez≠et restez dehors !* 'Get out and stay out!'

    c. les [femmes de marins]≠et leurs amants
      'sailors' wives and their lovers'

    d. Paul était [fier de tout]≠et enthousiasmé par n'importe quoi.
      'Paul was proud of everything and enthusiastic about just
      anything.'

3.1.2   The shape of the long form

As we stated above, for most words exhibiting a liaison alternation, the long form is identical to the short form except that it contains an extra final consonant. In many cases, the same consonant is also relevant for morphological processes. For instance, the masculine singular adjective *petit* 'small' has a long form [pətit] that ends in [t] just like the feminine singular form of the same adjective *petite* [pətit]. The same consonant shows up in derived words such as *petitesse* 'smallness' [pətitɛs], where the derivational suffix is [ɛs]. This type of data motivates the traditional idea that French phonological representations may contain a final "latent" consonant which shows up only when followed by material in the same word or at word boundaries where liaison is realized. Determining the shape alternation in these cases is a question of realization vs. non-realization of the latent consonant.

Although it is clear that the notion of latent consonant has some role to play in the grammar of French, it is important to make a clear distinction between the presence (or absence) of a latent consonant and the possibility (or impossibility) of liaison. First, singular nouns never give rise to liaison, despite the fact that some of them do have a latent consonant that is relevant morphologically (e.g., *dent* 'tooth' is realized as [dã] in all contexts, but is the base for the derived words *dentaire* [dãtɛʁ] 'dental' and *dentiste* [dãtist] 'dentist'). Thus the lexical representation of a word may include a latent consonant that is not involved in liaison.

Second, and more importantly, prenominal masculine singular adjectives may have a long form that is not related in this simple way to the corresponding short form. Three adjectives (*vieux* 'old', *beau* 'beautiful', and *nouveau* 'new') have a masculine singular long form phonologically identical to the feminine form of the adjective (resp. [vjɛj], [bɛl], and [nuvɛl]) but quite distinct from the masculine singular short form (resp. [vjø], [bo], and [nuvo]).[10] A dozen adjectives have a masculine singular long form whose final consonant is distinct from the one found in the feminine or in derived words (e.g., *gros* 'big', masculine singular

---

[10]The adjectives *mou* 'soft' and *fou* 'crazy' are usually also cited in this context, but their long forms have fallen out of use in contemporary French, except in a few fixed expressions.

short form [ɡʁo], masculine singular long form [ɡʁoz], feminine singu-
lar [ɡʁos], typical derived noun: *grosseur* 'bigness' [ɡʁosœʁ]).[11] Finally,
many adjectives are simply not possible in the masculine singular be-
fore a noun triggering liaison (Miller, 1992, Morin, 1998). A typical
example is *chaud*:

(17) a. une ambiance chaude / une chaude ambiance
          'a lively atmosphere'

      b. un débat chaud / un chaud débat
          'a lively debate'

      c. un entretien chaud / *un chaud entretien
          'a lively discussion'

   Prenominal position introduces a stilted stylistic effect for many ad-
jectives, making liaison judgments difficult to evaluate. It is clear, how-
ever, that *chaud* is not an isolated case. Dozens of other adjectives have
the same curious property of simply not having an acceptable masculine
singular prenominal liaison form.

   The data just discussed show that for the case of prenominal ad-
jectives in the masculine singular, the relationship between short and
long forms can involve more than just the realization or non-realization
of a latent consonant. To account for this data, we assume that the
paradigm of French adjectives contains an extra slot for the masculine
singular prenominal liaison form (Bonami and Boyé, 2003). This form
is identical to the feminine singular form by default; thus when the fem-
inine singular is suppletive (as in the cases of *vieux/vieille*, *beau/belle*,
nouveau/nouvelle discussed above), the masculine singular long form is
identical to the feminine form, not the masculine singular short form.
The default identity with the feminine singular is overridden in the case
of *gros* and similar adjectives. Finally adjectives like *chaud* are simply
defective—this lexeme is missing one of its inflectional forms.

   To sum up, the relation between the short and long forms of a word
can be determined by one of two factors: either the word has a latent
consonant and the long form is the short form with the latent consonant
realized at the end, or the short and long forms occupy distinct slots in

---

[11]The masculine singular long form of these adjectives is sometimes taken to be
derived from the feminine singular by a phonological process turning [s] into [z] and
[d] into [t] (Steriade, 1999). However this process would affect a non-natural class
of segments, perform totally opposite operations on them (voicing vs. devoicing),
and affect just a few lexical items sharing the same category and morphosyntactic
features in a syntactically defined environment—and even then, not fully produc-
tively (for example, *chaud* 'hot' and *froid* 'cold' are never realized as *[ʃot] and
*[fʁwat]). In light of these properties of the phenomenon, a lexical treatment is
clearly preferable.

the inflectional paradigm of the lexeme, and are related by inflectional morphology.[12]

3.2 An HPSG analysis for optional liaison

In this section we outline the general analysis and show how it applies to optional liaison contexts; we defer discussion of obligatory and impossible liaison to section 3.3.

3.2.1 Feature inventory

We introduce a number of new features to lexical and phrasal representations in order to encode the morphophonological conditions and effects of liaison.

$$(18) \quad sign \rightarrow \begin{bmatrix} \text{LEFT} & \begin{bmatrix} \text{LTRIG} & boolean \end{bmatrix} \\ \text{RIGHT} & \begin{bmatrix} \text{LFORM} & boolean \\ \text{APP} & list(segment) \end{bmatrix} \end{bmatrix}$$

First, a boolean-valued attribute LIAISON-TRIGGER indicates whether a word (potentially) licenses liaison to its left. Consonant-initial words (like *doué* in *très doué*) carry the feature [LTRIG −], except when the consonant is a glide ([j], [w], [ɥ]). In that case both possibilities exist: some words are [LTRIG +] (*mes [z] yeux* 'my eyes', *des [z] oiseaux* 'birds', *belles [z] huîtres* 'beautiful oysters'), others are [LTRIG −] (**les [z] hiéroglyphes* 'hieroglyphics', **bon [n] week-end* 'good weekend', **des [z] huées* 'jeers'). Vowel-initial words (e.g., *intelligent, ami, à, et*) are typically [LTRIG +], but there are exceptions too. So-called "*h aspiré*" words are (phonetically) vowel-initial, but they must be lexically specified as [LTRIG −] because they fail to trigger liaison: **curieux [z] hasard* 'funny coincidence', **tes [z] onze enfants* 'your eleven children'.

Next, two features are needed for representing liaison target status. The feature APPENDIX encodes the latent consonant (for both the liaison alternation and morphological derivation); thus a word such as *très* has a [z] in its APP, relevant for liaison, and the noun *dent* has a [t], relevant only morphologically. Of course many words simply have an empty appendix, if there is no reason to postulate a latent consonant.

The feature LIAISON-FORM indicates whether or not a word realizes liaison—in other words, for a word with distinct long and short forms,

---

[12]The plural forms of adjectives also give rise to liaison, but in these cases the shape alternation is systematically of the simple type, involving the latent consonant [z]. Note also that there are four determiners (three possessives *ma/mon, ta/ton, sa/son*, and the singular demonstrative *ce/cet(te)*) which give rise to lexically-controlled alternations similar to those found with adjectives.

the LFORM value determines which one will be chosen as the phonological realization of the word. For words with only one shape (e.g., feminine singular adjectives), the value of LFORM has no consequence on the phonology. All singular nouns are lexically specified as [LFORM −], so that even though they may have a latent consonant in APP (e.g., *dent*), they do not have long forms in liaison contexts. Masculine singular adjective forms are lexically specified as either [LFORM +] (e.g., *vieil*, *bel*) or [LFORM −] (e.g., *vieux*, *beau*).[13] Finally, words with distinct short and long forms differing only in the realization or non-realization of a latent consonant (e.g., *très*, plural adjectives and nouns), are lexically underspecified for the feature LFORM. The contextually instantiated value of the feature will determine whether the appendix is realized or not (as explained in section 3.2.3).

3.2.2 Propagation

Up to now we have only seen how the attributes LTRIG, LFORM, and APP operate in lexical entries. But since liaison can occur between words that are not sisters in a local tree, we need to specify a mechanism for the propagation of these features in syntactic combinations so that the relevant liaison information is visible at the phrasal level. It is clear that this propagation is not uniformly head-driven, or indeed driven by any syntactic considerations; it depends only on the linear order of the daughters. If the first word in a phrase is vowel-initial, then of course the phrase itself is vowel-initial, and similarly if the last word in a phrase is a long form that must appear in a liaison context, then the phrase as a whole must appear in a liaison context. More formally, a dominating phrase will always have the same liaison trigger status (LTRIG value) as its left-most daughter, and the same liaison target status and latent consonant (LFORM and APP) as its right-most daughter. Our liaison features can therefore be treated as "edge features", which have also been used for the formal analysis of phrasal affixes in French (Miller, 1992, Tseng, 2003b), and in an earlier HPSG treatment of liaison (Tseng, 2003a). The Edge Feature Principle (19) allows feature propagation along the right and left edges of phrases.[14]

---

[13]These values are part of the morphosyntactic properties regularly associated with the two relevant slots of adjectival paradigms; they are not stipulated word by word.

[14]We adopt an encoding of phrase structure in the spirit of Sag et al. (2003), whereby all daughters of a phrase (including the head daughter) are listed in the DTRS value, which is the locus of linear precedence constraints. Note that further work is needed to determine how the current analysis can be integrated with linearization-based analyses of French syntax (see e.g. Bonami, Godard, and Marandin, 1999).

(19) Edge Feature Principle

a.
$$phrase \rightarrow \begin{bmatrix} \text{LEFT} & \boxed{1} \\ \text{DTRS} & \left\langle [\text{LEFT}\,\boxed{1}] \right\rangle \oplus list(sign) \end{bmatrix}$$

b.
$$phrase \rightarrow \begin{bmatrix} \text{RIGHT} & \boxed{1} \\ \text{DTRS} & list(sign) \oplus \left\langle [\text{RIGHT}\,\boxed{1}] \right\rangle \end{bmatrix}$$

In combination with the feature geometry introduced in (18), liaison information appears correctly on all phrasal signs. (20) illustrates the percolation of features in a simple phrase.[15]

(20) *amis intimes* 'close friends':

$$\begin{bmatrix} \text{L} & \begin{bmatrix} \text{LTRIG} & \boxed{3}+ \end{bmatrix} \\ \text{R} & \begin{bmatrix} \text{LFORM} & \boxed{4}\ bool \\ \text{APP} & \boxed{5}\langle z \rangle \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{PHON} & \langle \text{ami} \rangle \\ \text{CAT} & \textsf{N} \\ \text{L} & [\text{LTRIG} \quad \boxed{3}+] \\ \text{R} & \begin{bmatrix} \text{LFORM} & bool \\ \text{APP} & \langle z \rangle \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{PHON} & \langle \tilde{\varepsilon}\text{tim} \rangle \\ \text{CAT} & \textsf{Adj} \\ \text{L} & [\text{LTRIG} \quad +] \\ \text{R} & \begin{bmatrix} \text{LFORM} & \boxed{4}\ bool \\ \text{APP} & \boxed{5}\langle z \rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

3.2.3 Phonological realization

We still need to explain formally how the various combinations of values of LFORM, LTRIG, and APP give rise to the characteristic phonological aspects of the liaison alternation. We define a function dtrs-to-phon, taking a list of signs and producing a list of phonological strings, for this purpose.[16]

(21)
$$sign \rightarrow \begin{bmatrix} \text{PHON} & \textsf{dtrs-to-phon}(\boxed{\Sigma}) \\ \text{DTRS} & \boxed{\Sigma} \end{bmatrix}$$

---

[15]In this figure L abbreviates LEFT and R abbreviates RIGHT. In later figures the features LEFT and RIGHT are omitted, since no ambiguity can arise. In addition, we abbreviate CAT values to traditional category labels ($\textsf{V}$, $\textsf{VP}$, $\textsf{S}$, etc.).

[16]For the sake of concreteness, we treat phonological representations as lists of segment sequences, and we treat phonological combination as list concatenation. This extremely simplified view of phonology is sufficient for our purposes.

The first clause of the definition (22) takes care of the realization of liaison between two daughters in a phrase. If the first daughter is a liaison form ([LFORM +]) and the next daughter is a liaison trigger ([LTRIG +]), this clause adds both the PHON value of the first daughter and its APP (latent consonant, if any) to the phonology of the phrase. The recursive call to dtrs-to-phon specifies what must be done with the phonology of the remaining daughter(s).

(22)
$$\text{dtrs-to-phon} \left(\left\langle \begin{bmatrix} \text{PHON} & \boxed{1} \\ \text{APP} & \boxed{2} \\ \text{LFORM} & + \end{bmatrix}, \boxed{3}[\text{LTRIG} \ +] \right\rangle \oplus \boxed{\Sigma} \right)$$
$$= \boxed{1} \oplus \boxed{2} \oplus \text{dtrs-to-phon} \left(\langle\boxed{3}\rangle \oplus \boxed{\Sigma} \right)$$

The second clause of dtrs-to-phon (23) covers cases where liaison is not realized after the first daughter. This daughter's PHON value is incorporated into the phrasal phonology, its APP value is ignored, and dtrs-to-phon is called recursively to handle the rest of the list (which must be non-empty). Note that this clause does not check the LTRIG status of the following daughter. This is in accordance with our decision to treat optional liaison as the default situation: since liaison is optional in the general case, the absence of liaison does not need to be licensed by properties of the next daughter.

(23)
$$\text{dtrs-to-phon} \left(\left\langle \begin{bmatrix} \text{PHON} & \boxed{1} \\ \text{LFORM} & - \end{bmatrix} \right\rangle \oplus \boxed{\Sigma} \ nelist(sign) \right)$$
$$= \boxed{1} \oplus \text{dtrs-to-phon} \left(\boxed{\Sigma} \right)$$

Finally, the last sign of every phrase's DTRS list is handled by clause (24). No matter what the final sign's LFORM value is, we simply add its PHON value to the phrasal phonology, without appending the APP list, and the calculation of the phrasal phonology terminates.[17]

___

[17]Of course, the right-most daughter of a phrase can be [LFORM +], and it can have a non-empty APP, but clause (24) ignores these features. The information is not lost, however, because the phrase itself shares its LFORM and APP values with this right-most daughter, in accordance with the EFP (19). Consequently, when this phrase combines with other material to form a larger phrase, it will appear on a higher DTRS list, and its right-edge liaison features will either be taken into account there (see e.g. example (30)), or they will continue to propagate to the next higher phrase. We assume that complete utterances are always [LFORM −], so they can never end with a liaison form (like *vieil*), and the formulation of dtrs-to-phon ensures that a latent consonant is never realized at the end of a complete utterance (i.e. maximal DTRS list).

(24)
$$\text{dtrs-to-phon}\left(\left\langle\begin{bmatrix}\text{PHON} & \boxed{1}\end{bmatrix}\right\rangle\right)=\boxed{1}$$

Going back to example (20), note that the value of LFORM on *amis* is underspecified. Thus there are two possible outputs of dtrs-to-phon: either *amis* is [LFORM +] and (22) applies, or it is [LFORM −] and (23) applies. This situation is typical of optional liaison contexts.[18]

(25) a.
$$\text{dtrs-to-phon}\left(\begin{bmatrix}\text{PHON} & \langle\text{ami}\rangle \\ \text{LFORM} & + \\ \text{APP} & \langle\text{z}\rangle\end{bmatrix},\begin{bmatrix}\text{PHON} & \langle\tilde{\varepsilon}\text{tim}\rangle \\ \text{LTRIG} & -\end{bmatrix}\right)=\langle\text{ami,z,}\tilde{\varepsilon}\text{tim}\rangle$$

b.
$$\text{dtrs-to-phon}\left(\begin{bmatrix}\text{PHON} & \langle\text{ami}\rangle \\ \text{LFORM} & - \\ \text{APP} & \langle\text{z}\rangle\end{bmatrix},\begin{bmatrix}\text{PHON} & \langle\tilde{\varepsilon}\text{tim}\rangle \\ \text{LTRIG} & -\end{bmatrix}\right)=\langle\text{ami,}\tilde{\varepsilon}\text{tim}\rangle$$

In the structurally identical case of (26), since the second element on DTRS is [LTRIG −], clause (22) cannot apply, and thus only one realization (with no liaison) is possible.

(26) *amis chers* [amiʃɛʁ] 'dear friends':
$$\begin{bmatrix}\text{PHON} & \text{dtrs-to-phon}(\boxed{1},\boxed{2})=\langle\text{ami,}\text{ʃɛʁ}\rangle \\ \text{LTRIG} & \boxed{3}+ \\ \text{LFORM} & \boxed{4} \\ \text{APP} & \boxed{5}\langle\text{z}\rangle \\ \\ \text{DTRS} & \left\langle\boxed{1}\begin{bmatrix}\text{PHON} & \langle\text{ami}\rangle \\ \text{CAT} & \text{N} \\ \text{LTRIG} & \boxed{3}+ \\ \text{LFORM} & bool \\ \text{APP} & \langle\text{z}\rangle\end{bmatrix},\boxed{2}\begin{bmatrix}\text{PHON} & \langle\text{ʃɛʁ}\rangle \\ \text{CAT} & \text{Adj} \\ \text{LTRIG} & - \\ \text{LFORM} & \boxed{4} \\ \text{APP} & \boxed{5}\langle\text{z}\rangle\end{bmatrix}\right\rangle\end{bmatrix}$$

An important feature of the dtrs-to-phon function is that it decides whether liaison occurs not on the basis of the content of the APPENDIX, but on the basis of the LFORM feature. Thus words with an empty appendix can still be liaison forms licensed by clause (22) (prenominal

---

[18]Note that a property of our analysis is that the liaison consonant occurs as an autonomous, unsyllabified element on the PHON list. This leaves it to the phonology proper to determine where the liaison consonant is syllabified, in accordance with the well-known observation that both syllabifications (rightward and leftward) are possible (Encrevé, 1988), and that pauses may occur on either side of the liaison consonant (Morin and Kaye, 1982).

adjectives like *vieil*), and words with a non-empty appendix can fail to give rise to liaison (e.g., *dent*) because they are [LFORM −]. More generally, many (if not most) words both have an empty appendix and are underspecified for the LFORM feature (e.g., *frère*, *vrai*, *avec*). For these words, applying clause (22) (if allowed by the following DTRS element) or (23) leads to exactly the same result.

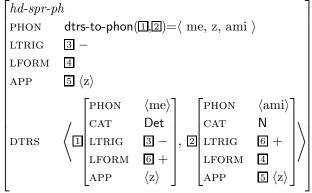## 3.3 Constraining liaison contexts

In accordance with the observations in section 3.1, we now provide an account of contexts where liaison is obligatory or impossible. Remember that since the formulation of dtrs-to-phon treats optional liaison as the default case, nothing needs to be added to the grammar to account for the optional cases.

3.3.1 Obligatory liaison

To account for obligatory liaison between specifier and head, we assume that the LFORM feature of the specifier must be identical to the LTRIG of the head (27). This makes the realization of liaison entirely dependent on the trigger status of the second daughter. As (28) illustrates, the effect of this specification is that dtrs-to-phon can only produce one result when the second daughter is [LTRIG +], unlike in optional liaison contexts (e.g., (20), (25))

(27) $hd\text{-}spr\text{-}ph \rightarrow \left[\text{DTRS} \quad \left\langle [\text{LFORM} \; \boxed{1}], [\text{LTRIG} \; \boxed{1}] \right\rangle \right]$

(28) Obligatory liaison with latent consonant: *mes amis* [mezami]

$$
\begin{bmatrix}
hd\text{-}spr\text{-}ph \\
\text{PHON} \quad \text{dtrs-to-phon}(\boxed{1},\boxed{2})=\langle \text{ me, z, ami } \rangle \\
\text{LTRIG} \quad \boxed{3} - \\
\text{LFORM} \quad \boxed{4} \\
\text{APP} \quad \boxed{5} \langle \text{z} \rangle \\
\text{DTRS} \quad \left\langle \boxed{1} \begin{bmatrix} \text{PHON} & \langle \text{me} \rangle \\ \text{CAT} & \text{Det} \\ \text{LTRIG} & \boxed{3} - \\ \text{LFORM} & \boxed{6} + \\ \text{APP} & \langle \text{z} \rangle \end{bmatrix}, \boxed{2} \begin{bmatrix} \text{PHON} & \langle \text{ami} \rangle \\ \text{CAT} & \text{N} \\ \text{LTRIG} & \boxed{6} + \\ \text{LFORM} & \boxed{4} \\ \text{APP} & \boxed{5} \langle \text{z} \rangle \end{bmatrix} \right\rangle
\end{bmatrix}
$$

3.3.2 Impossible liaison

We now turn to cases of impossible liaison. To block liaison in a particular context, it is sufficient to constrain the relevant sign to be

[LFORM −]. We illustrate first with subject-head combinations. The constraint in (29) forces both daughters in a head-subject phrase to be [LFORM −]. Thus when we combine the NP in (28) with a VP, the *head-subj-ph* type forces a [LFORM −] specification on the NP, irrespective of the VP's LTRIG value. This is illustrated with an [LTRIG +] VP in (30). The [LFORM −] specification on the second daughter in (29) blocks liaison between the whole clause and a further constituent, even though the clause has an appendix (which originates on the verb). This accounts for the observation in (14) that there can be no liaison between a subject-head combination and a following constituent.

(29) *head-subj-ph* → $\left[\text{DTRS} \ \Big\langle [\text{LFORM} \ -],[\text{LFORM} \ -] \Big\rangle \right]$

(30) *Mes amis arrivent* [mezamiariv] 'My friends are coming'

$$\begin{bmatrix} \textit{hd-subj-ph} \\ \text{PHON} \quad \text{dtrs-to-phon}(\boxed{1},\boxed{2}) = \langle \text{ me, z, ami, aʁiv } \rangle \\ \text{LTRIG} \quad \boxed{3} - \\ \text{LFORM} \quad \boxed{4} - \\ \text{APP} \quad \boxed{5} \langle \text{t} \rangle \\ \\ \text{DTRS} \quad \Bigg\langle \boxed{1} \begin{bmatrix} \text{PHON} & \langle \text{me, z, ami} \rangle \\ \text{CAT} & \text{NP} \\ \text{LTRIG} & \boxed{3} - \\ \text{LFORM} & - \\ \text{APP} & \langle \text{z} \rangle \end{bmatrix}, \boxed{2} \begin{bmatrix} \text{PHON} & \langle \text{aʁiv} \rangle \\ \text{CAT} & \text{VP} \\ \text{LTRIG} & + \\ \text{LFORM} & \boxed{4} - \\ \text{APP} & \boxed{5} \langle \text{t} \rangle \end{bmatrix} \Bigg\rangle \end{bmatrix}$$

Head-filler phrases are subject to a constraint exactly parallel to that on head-subject phrases (31). For head-complement phrases, we need to account for the fact that liaison is possible after the head and lite complement daughters, but that it is impossible after nonlite complements. The constraint in (32) licenses this behavior (assuming that the head always left-most in head-complement structures in French). In addition, the whole head-complement phrase is [LFORM −], blocking liaison to the right of the whole phrase, in accordance with the observation in (16).

(31) *head-fill-ph* → $\left[\text{DTRS} \ \Big\langle [\text{LFORM} \ -],[\text{LFORM} \ -] \Big\rangle \right]$

(32) *head-comps-ph* →
$$\begin{bmatrix} \text{LFORM} & - \\ \text{DTRS} & \langle \ [\ ] \ \rangle \oplus \textit{nelist}\Big([\text{WEIGHT} \ \textit{lite}] \vee [\text{LFORM} \ -]\Big) \end{bmatrix}$$

3.4   Conclusion

In this paper we provided both new data illustrating the syntactic constraints on French liaison, and an HPSG grammar fragment that accounts for the observed distribution of obligatory, optional, and impossible liaison.

The account is incomplete in that it does not consider non-syntactic factors that make optional liaison more or less probable (in extreme cases, almost mandatory or almost impossible). Two approaches can be taken to integrate such data into the current analysis. One possibility is to leave the grammar as it is, and introduce a post-grammatical component to determine the likelihood of a particular liaison, taking into account lexical, prosodic, collocational, and other information. A second, more promising line of analysis is to make a more sophisticated use of the dtrs-to-phon function. Since this function operates on complete representations of the signs it combines, it has access to all the necessary information. Thus dtrs-to-phon could output, in addition to the phonological form of a phrase with or without liaison, an indication of the probability of actually realizing that particular option.

References

Abeillé, Anne and Danièle Godard. 2000. French word order and lexical weight. In R. D. Borsley, ed., *The Nature and Function of Syntactic Categories*, vol. 32 of *Syntax and Semantics*, pages 325–360. San Diego: Academic Press.

Abeillé, Anne and Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78:404–452.

Bonami, Olivier and Gilles Boyé. 2003. La nature morphologique des allomorphies conditionnées. Les formes de liaison des adjectifs en français. In Fradin et al. (2003).

Bonami, Olivier, Danièle Godard, and Jean-Marie Marandin. 1999. Constituency and word order in French subject inversion. In G. Bouma, E. Hinrichs, G.-J. Kruijff, and R. T. Oehrle, eds., *Constraints and Resources in Natural Language Syntax and Semantics*, pages 21–40. Stanford, CA: CSLI Publications.

Bybee, Joan. 2001. Frequency effects on French liaison. In J. Bybee and P. Hopper, eds., *Frequency and the emergence of linguistic structure*, pages 337–359. Amsterdam: John Benjamins.

Comorovski, Ileana. To appear. Quel. In F. Corblin and H. de Swart, eds., *Handbook of French Semantics*. Stanford, CA: CSLI Publications.

de Jong, Daan. 1994. La sociophonologie de la liaison orléanaise. In C. Lyche, ed., *French Generative Phonology: Retrospective and Perspectives*, pages 95-130. Salford: AFLS/ESRI.

Delattre, Pierre. 1966. *Studies in French and comparative phonetics*. The Hague: Mouton.

Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement*. Paris: Seuil.

Fouché, Pierre. 1959. *Traité de prononciation française*. Paris: Klincksieck.

Fougeron, Cécile, Jean-Philippe Goldman, Alicia Dart, Laurence Guélat and Clémentine Jeager. 2001. Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison français. *Actes de TALN 2001*, pages 173–182.

Fradin, Bernard, Georgette Dal, Françoise Kerleroux, Nabil Hathout, Marc Plénat, and Michel Roché, eds. 2003. *Les Unités morphologiques / Morphological Units*. Villeneuve d'Ascq: Silex.

Miller, Philip H. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.

Miller, Philip H. and Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15:573–639.

Morin, Yves-Charles. 1998. Remarks on prenominal liaison consonants in French. Ms., Université de Montréal.

Morin, Yves-Charles and Jonathan D. Kaye. 1982. The syntactic bases for French liaison. *Journal of Linguistics* pages 291–330.

Post, Brechtje. 2000. Pitch accents, liaison and the phonological phrase in French. *Probus* 12:127–164.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: a Formal Introduction*. Stanford, CA: CSLI Publications.

Selkirk, Elisabeth O. 1972. *The phrase phonology of English and French*. Ph.D. thesis, MIT.

Selkirk, Elisabeth O. 1974. French liaison and the $\bar{\text{X}}$ convention. *Linguistic Inquiry* 5:573–590.

Selkirk, Elisabeth O. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371–405.

Steriade, Donka. 1999. Lexical conservatism in French adjectival liaison. In J.-M. Authier, B. E. Bullock, and L. Reed, eds., *Formal perspectives on Romance linguistics*, pages 243–270. Amsterdam: John Benjamins.