The domain of liaison: theories and data*

GEERT BOOIJ and DAAN DE JONG

Abstract

In this paper it is argued that current theories of the domain of liaison are inadequate. It is shown that 'domain' is a variable intralinguistic constraint on the (variable) rule of liaison. The variable nature of liaison is also apparent from the fact that its rate of application covaries with extralinguistic factors such as style and social class. These results also show that rules of sentence phonology should be based on reliable corpora of spoken language.

1. Introduction

Liaison can be seen as the phenomenon of latent word-final consonants surfacing phonetically before a vowel-initial word. For instance, in *petit ami* [p(ə)titami] 'little friend' the final *t* of *petit* is realized phonetically, whereas it does not surface in *petit garçon* [p(ə)tigarsɔ̃] 'small boy', because the following word begins with a consonant. These latent consonants also surface in inflectional and derivational morphology, as *petite* [p(ə)tit] 'small', fem. sg., and *petitesse* [ptitɛs] 'smallness' illustrate. A survey of alternative views of liaison is presented in Klausenburger (1984). The liaison consonant is usually (but not always, see Encrevé 1983) tautosyllabic with the next word, the first fragment of the second word, as is illustrated by the syllabification of *petit ami*: $(\text{pti})_\sigma(\text{ta})_\sigma(\text{mi})_\sigma$ where σ = syllable. That is, liaison usually implies *enchaînement*.[1]

In this paper, we will focus on one particular problem in the analysis of liaison, its domain of application. It is well known that liaison only applies within certain environments. The proper characterization of these environments has been the subject of much discussion in recent phonological literature (for example, Selkirk 1974, 1978, 1984; Nespor and Vogel 1982; Kaisse 1985). Basing our arguments on data from a number of corpora, we will show that the domains in which liaison applies,

Linguistics 25 (1987), 1005–1025

0024–3949/87/0025–1005 $2.00

obligatorily or optionally, are not correctly defined by current theories. Related to this issue, we will also investigate whether Kaisse's (1985) division of the rules of sentence phonology into two discrete categories, external sandhi rules and fast-speech rules, can be maintained in the light of the facts of liaison. Our conclusion will be that the contextual restrictions on liaison are of a variable nature, that is, they are intralinguistic variable conditions in the sense of the Labovian theory of variable rules. Furthermore, it will be shown that Kaisse's categorization of the rules of sentence phonology is not tenable in its strict form.

2. The domain of liaison

Table 1 gives the contexts in which liaison is traditionally assumed to be obligatory, optional, or forbidden (adapted from Delattre 1966). In traditional descriptions of liaison, its domain is referred to as 'rhythmical unit' (Grammont 1938: 129; Malmberg 1972: 41) or 'breath group' (Pulgram 1965: 131). In Selkirk (1980 [1972], 1974) a first attempt was made to formally define the domain of liaison on the basis of syntactic structure. In Selkirk (1978) this theory was replaced by a general theory concerning the domains of prosodic rules. This latter theory assumes that, first, the syntactic surface structure of a sentence is mapped onto a prosodic structure (with prosodic categories such as syllable, foot, phonological word, phonological phrase, etc.). Prosodic rules then apply within prosodic domains like the phonological phrase. It is proposed in Selkirk (1978) that the phonological phrase (φ) is the domain of obligatory liaison. The phonological phrase may be derived as follows from syntactic surface structure (we present here the variant defended in Nespor and Vogel 1982: 228–229):

Table 1. *Obligatory, optional, and forbidden liaison*[2]

	Obligatory	Optional	Forbidden
Noun	noun determiner + pronoun adjective	plural noun	singular noun
Verb	personal pronoun + verb verb + personal pronoun	verb + following word	
Invariable words	monosyllables	polysyllables	*et* + following word
Special cases	frozen expressions		*h aspiré* word + *un*, *huit*, *onze*, and derivations

(1) φ construction:
Join into a φ a lexical head (X) with all the items on its nonrecursive side within the maximal projection, and with any other nonlexical item on the same side (such as prepositions, complementizers, conjunctions, copulas).

According to this definition of φ, all the liaison contexts mentioned above will be defined as phonological phrases. On the other hand, rule (1) will define the syntactic phrase *des athlètes américaines* 'American athletes' as consisting of two φs, *des athlètes* and *américainès*. Hence we predict liaison only between the first two words of this phrase, as required. However, in a more formal style of speech we may also find liaison between *athlètes* and *américaines*; to account for this, Selkirk (1974: 581) proposed the following rule:

(2) A head noun, verb, or adjective which is inflected may be in a liaison context with the word that follows, if that word is in its complement.

No attempt has been made, either in Selkirk (1978) or in Nespor and Vogel (1982), to reformulate this generalization concerning the formal style of speech within the framework of prosodic phonology. Nespor and Vogel proposed a rule of φ' construction for an analogical situation in Italian, in which heads and complements consisting of one word form, optionally, one domain with respect to the prosodic rule of consonant gemination. We might use the same strategy for liaison and formulate the following rule of φ' construction for French (compare Nespor and Vogel 1982: 230):

(3) φ' construction (formal style only):
A φ which is the first complement of an inflected X on its recursive side loses its label and is joined to the X under a new node labeled φ'.

The domain of liaison would then have to be defined as φ and φ'.

However, the problem with such an approach is that it would be completely ad hoc, at least for French, since we do not know of any independent motivation for such an additional prosodic category in that language. Moreover, it appears that liaison across the boundary between heads and complements also occurs in other styles (see below).

In Selkirk (1984: 334) we find a third approach to the definition of the domain of liaison, an approach, however, that embodies the same basic claim: syntactic structure determines, be it indirectly, the domain of application of liaison. This third theory claims that liaison applies obligatorily if the two relevant words are separated by at most one silent demibeat (a demibeat is a position on the metrical grid of a sentence). The

number of silent demibeats between the words is determined by the rule of silent demibeat addition (SDA), which reads as follows:

(4) Silent demibeat addition:
Add a silent demibeat at the end (right extreme) of the metrical grid aligned with
a. a word;
b. a word that is the head of a nonadjunct constituent;
c. a phrase;
d. a daughter phrase of S.

In this rule, function words (pronouns, prepositions, auxiliaries, etc.) are not counted as words (the 'principle of categorical invisibility of function words'). Hence, we will get configurations like the following (where each x indicates a silent demibeat) as far as silent demibeats between words are concerned:

(5) a. nous aimons
b. dans une minute
c. petit x ami
d. des athlètes x x américaines

The representations in (5) imply that liaison will only be blocked in the last example, where the rhythmic disjuncture between *athlètes* and *américaines* is too large. However, within this theory it is impossible to give a straightforward account of the extension of the domain of liaison as formulated in (2), as Selkirk herself points out (1984: 334). On the contrary, Selkirk's theory wrongly predicts that liaison in (5d) is possible in fast speech only, because in fast speech the actual time value of the silent demibeats will decrease and hence we will have less rhythmic disjuncture in such phrases. Thus, Selkirk is forced to completely separate the analysis of obligatory liaison from that of optional liaison:

> In such contexts [of optional liaison] liaison is no longer an essentially phonological phenomenon, but one being maintained by some rules that may be quite 'grammaticized' or 'syntacticized' and no longer reflect the processes of 'core phonology' (1984: 334).

The data presented below, however, suggest that such a division between purely phonological liaison and syntacticized liaison is unwarranted because optionality and domain variation are also found in nonformal speech.

A third theory concerning the domains of liaison is found in Kaisse (1985). Kaisse proposes a division of the rules of sentence phonology into two blocks. The first block contains 'external sandhi rules', which are

conditioned by syntactic and morphological factors, whereas their application is rate-independent and cannot be affected by pauses between the two words in question. The second block contains 'fast-speech rules', which are dependent on speech rate, syllabification, and the features of focus and determinant, but are insensitive to syntactic, morphological, or lexical factors. Another difference between the blocks is that only the first is sensitive to style and register, that is, it is socially conditioned.

The main syntactic condition on external sandhi rules is formulated in terms of the c-command relation. As far as liaison is concerned, Kaisse too makes a strict separation between liaison usage in formal and in informal speech and decides to study liaison usage in informal speech only, where obligatory liaison is predicted as follows:

(6) Word *b* must c-command word *a*.

Note that Kaisse predicts that liaison, being an external sandhi rule, is insensitive to pauses. However, the data presented in 3.3.8 show that liaison is highly sensitive to pauses. This implies that liaison is a counterexample to Kaisse's categorization of the rules of sentence phonology.

It should be realized that all these variant definitions of the domain of liaison are based on publications that implicitly or explicitly aim to set a standard for liaison: that is, they are normative. Moreover, they only claim to define standard French, the norm for well-educated speakers. Therefore a phonologist studying liaison is in need of a reliable corpus of spoken French that can be used to compare the theoretical claims made about liaison with the actual liaison behavior of French speakers. In particular, such a corpus will enable us to confront the claims made with respect to the domains in which liaison is obligatory, optional, or impossible with the data. Second, it may give us a deeper insight into the relation between the actual application of liaison and system-external, possibly interdependent factors like the style of speech and the social class. For these reasons we will present a survey of what is known about the various intra- and extralinguistic factors that affect the application of liaison in the following sections. Subsequently, we will discuss what consequences these variational data must have for our view of liaison, in particular as far as its domain of application is concerned.

As source materials we will mainly use five standard European French corpora (Agren 1973; De Jong et al. 1981; Encrevé 1983; Malécot 1979; Morin and Kaye 1982). These corpora, as well as a few others, are described in more detail in the Appendix at the end of this paper.

In section 3 we will show that the exceptions to the liaison domains as defined above are too numerous to be attributed to performance errors. We will also present a survey of the various intra- and extralinguistic

factors influencing the use of liaison. In section 4 we will discuss the role of social and stylistic factors.

3. The domains of liaison: empirical evidence

3.1. *The domain of obligatory liaison*

The corpus of De Jong et al. (1981) provides evidence about the obligatory domain of liaison. Thirty-eight informants — all native speakers of standard French — were asked to read aloud a text containing 11 contexts satisfying the conditions for obligatory liaison. Thus 418 contexts were gathered (see Table 2). In only 223 out of the 418 cases was liaison indeed realized (53.3%). The contexts listed in Table 2 were used (the number of individuals who used liaison is specified for each context in percentages and in absolute frequencies [N]).

According to the corpus of De Jong et al. and the corpus of Encrevé (1983: 48), liaison appears obligatorily only after determiners and in the combinations 'personal pronoun + verb' and 'verb + pronoun', as well as in certain frozen expressions.

3.2. *The domain of optional liaison (inflected head + complement)*

The most systematic empirical evidence about the optional domain of liaison can be found in Morin and Kaye (1982) and Agren (1973). Morin

Table 2. *Liaison usage in the 'domain of obligatory liaison'; N indicates the number of individuals who used liaison*

	%	N	
grandes_averses	86	32	'heavy rainshowers'
anciennes_amies	82	31	'old friends'
en_un bouquet	82	31	'in a bouquet'
trop_innocent	82	31	'too innocent'
certains_exécutants	76	29	'certains executors'
fort_interessant	71	27	'very interesting'
autres_activités	50	19	'other activities'
après_un moment	26	10	'after a moment'
pendant_un moment	20	8	'during a moment'
assez_humide	11	4	'rather moist'
depuis_un an	5	1	'since a year'
Total	53.3	223	

and Kaye raise the following objections against Selkirk's optional domain of liaison:

i. Liaison between an inflected head and a following complement is also found in casual conversation. Morin and Kaye (1982: 298) cite Agren, who found liaison between an inflected noun and a following adjective in casual speech in 14% of the cases.

ii. Liaison also appears between two complements, as in (7):

(7) des attaques nucléaires‿américaines
'American nuclear attacks'

iii. Liaison can also apply between an inflected word and a non-complement (Morin and Kaye 1982: 307):

(8) Ils réfléchissent‿avant de répondre.
'They think before answering.'

iv. As to verbs, liaison applies more frequently to 3rd person verbs than to 1st person verbs.

v. There are cross-category differences between a head noun, verb, and adjective. In sentences like (9) liaison appears frequently, but it is almost nonexistent in sentences like (10) (Morin and Kaye 1982: 314–315):

(9) Ils les rappellent‿à l'ordre.
'They call them back to order.'
(10) des rappels‿à l'ordre
'recalls to order'

Thus we see that the domain of optional liaison is not correctly defined by the theories in section 2.

3.3. *Intralinguistic factors determining the usage of liaison*

In all of the corpora we find indications that the rule for the application of liaison is sensitive to a number of intralinguistic factors. Below we present a brief discussion of the nine most relevant factors.

3.3.1. *Word category.* Morin and Kay (1982: 314) claim that, with regard to an inflected head and a following complement, liaison is considerably more frequent when the head is a verb than when it is a noun. Apparently, word category is a factor. This also applies to the obligatory domains as formulated in the previous section: liaison is obligatory after determiners and clitic pronouns, but not after the other word categories within this domain.

3.3.2. *Syntactic surface structure.* Agren's analysis provides indications that syntactic surface structure directly influences liaison. Depending on the syntactic category of the following constituent, liaison frequencies vary considerably. Examples are given in Tables 3, 4, and 5.

Table 3. *Liaison scores between* avoir *and following word in percentages (N = total number of liaison contexts)*

avoir +	%	N
Past participle	61	716
un + noun	37	277
Adverbial phrase	33	147
Total	51	1140

Table 4. *Liaison scores between* pas *and following word in percentages (N = total number of liaison contexts)*

pas +	%	N
Adjective	43	46
Infinitive	27	220
Past participle	25	219
Adverb	23	177
un + noun	19	185
Preposition	9	103
Adverbial phrase	0	15
Total	23	965

Table 5. *Liaison scores between polysyllabic adverbs and following word in percentages (N = total number of liaison contexts)*

polysyllabic adverbs +	%	N
Adjective	52	483
Past participle	39	210
Infinitive	35	56
Adverb	16	60
Preposition	16	179
Total	40	988

3.3.3. *Nature of the latent consonant.* A latent consonant may be a /t/, /z/, /n/, /r/, /p/, or /k/.[3] Several researchers have tested whether liaison could be dependent upon the nature of the latent consonant. They all conclude that this is the case. In the corpus of Encrevé (1983: 52) relative liaison frequencies for /t/, /z/, and /r/ are 72%, 39.8%, and 11% respectively. After some words /n/ was realized obligatorily, after others never.[4]

Morin and Kaye (1982: 295) point out that liaison is more frequent after prenominal adjectives with the plural suffix /z/ than after singular adjectives. As for verbs, they notice that liaison is more easily made after 3rd person verbs (*t* liaison) than after 1st person verbs (*z* liaison). This also holds good for the corpus of De Jong et al.: *étais* (2.3%), *était* (18.7%); *vais* (9.5%), *vont* (33%).

Malécot's (1979: 168) results are different from Encrevé's. Liaison scores for /z/, /t/, and /r/ are 60.5%, 52.4%, and 94.4% respectively (the data for /r/ were too scarce to be reliable). The difference in rank order with Encrevé's data is due to the fact that Malécot put together obligatory and optional liaisons, whereas Encrevé's analysis refers to optional liaisons only.

3.3.4. *Word length.* After short words liaison is more frequent than after long ones. Encrevé (1983: 52) divided his data into monosyllables and polysyllables. After monosyllables the liaison frequency was 77%, after polysyllables 29.3%. In the corpora of De Jong et al., Agren, and Malécot liaison is effected more often after monosyllabic prepositions and adverbs than after polysyllabic ones.

3.3.5. *Word frequency.* Agren (1973: 28) points out that after frequent words liaison is more easily produced than after less-frequent words. For instance, the frequent 3rd person sg. form of *être*, *est* (2669 cases), triggers liaison almost automatically (97%), whereas for the more rare form *êtes* (34 cases), the score falls to 71%. Although this hypothesis seems to be supported by many more examples, Agren does not take it into consideration as an independent factor, among other things because of the close correlation with word length; that is, short words are often frequent words.

In this context, one might wish to suggest that not only word frequency but also word-combination frequency is a factor. Thus in common expressions such as *premier étage* 'first floor', liaison is produced more often than in rare combinations like *singulier ami* 'singular friend' (De Jong et al. 1981: 93–94).

3.3.6. *Preceding segments.* Morin and Kaye (1982: 295) assume that a latent consonant preceded by a vowel is realized more often than one preceded by a consonant. Delattre (1966: 60) adds that liaison is more frequent when one consonant precedes than when two do so. However, this only seems to hold good for cases of optional liaison. For cases of obligatory liaison, we do not find this, as is shown in Kovac's pilot study (1979: 150) and in Agren (1973: 127–130), who examined this factor for all sequences 'noun + adjective'.

3.3.7. *Length of the following constituent.* Morin and Kaye (1982: 296) suppose that liaison is more easily produced before short complements than before long ones. For example, in (11) liaison would be more frequent than in (12):

(11) Ils travaillent d'abord et mangent‿après.
'They work first and eat later.'
(12) D'habitude, ils mangent‿apres avoir joué leur partie d'échecs.
'Usually, they eat after having finished their chess game.'

Delattre shares this point of view. This factor has not been investigated in any of the analyses.

3.3.8. *Pauses.* Liaison is highly sensitive to pauses. In Agren's corpus we find liaison in only 5% of the cases where there was a pause between the two words in a liaison context. However, latent consonants can be realized before a pause; that is, we find liaison without resyllabification.

A second important point is that this pause sensitivity contradicts Kaisse's hypothesis about the insensitivity of external sandhi rules to pauses. Agren's data show that the use of liaison is heavily constrained by the presence of a pause between the two words in a liaison context. In other words, sensitivity to phonetic pauses is not a distinguishing property between external sandhi rules and fast-speech rules.

3.3.9. *Frozen expressions.* A last factor to be taken into account is the obligatory liaison in frozen expressions like *accent aigu, fait accompli, le cas échéant* (for a more complete list see Malécot 1979: 166–167).

3.4. *Discussion*

Liaison is obligatory in a very limited number of contexts. In most contexts where liaison can apply, it applies optionally. Under the influence of a number of intralinguistic factors, liaison frequencies vary

considerably. This influence is systematic, and therefore we claim that liaison is not simply an optional rule, but a variable rule in the Labovian sense.

The syntax-based definitions of the domains of liaison given in section 2 indicate at most where liaison can apply, but even then we find several exceptions, as has been shown in this section. The domain definitions of section 2 specify only a set of contexts, which, depending on their specific nature, constrain or favor rule application.

Further research will have to determine (a) which of the nine factors we mentioned exert a significant influence upon liaison (perhaps not all of them are important) and (b) to what degree they exert an influence upon liaison (the effect of, for example, word category might be stronger than the effect of word length).

The variable rule method developed by Sankoff and Labov (Sankoff 1978; Labov 1980) is an excellent heuristic tool for the analysis of this kind of problem. Although the status of variable rules in relation to generative grammar is a much-debated issue (Kay and McDaniel 1979; Romaine 1984, 1985), we think that no adequate theory about liaison can be developed without taking into account its fundamentally variable nature. Because of the complex nature of the variability in liaison usage, neither traditional grammars nor intuitions can be used as empirical evidence: the direct study of liaison in natural speech by means of corpora is an absolute requirement for gaining an insight into the constraints upon its usage.[5]

4. External factors in the use of liaison

In this section we will present data confirming the variable-rule status of liaison in that its rate of application covaries with social class, age, sex, and degree of monitoring. The influence of external factors upon liaison usage was the main concern of De Jong et al. (1981). Encrevé's research is limited to the speeches of politicians; Malécot and Morin and Kaye investigate the language of upper-class speakers; while Agren examines only the influence of style. For these reasons we will restrict our discussion of the role of external factors to De Jong et al. (1981).

4.1. *Social class*

In virtually all linguistic contexts and on all stylistic levels liaison correlates with social class (see Tables 6 and 7). Without any difficulty,

Table 6. *Percentage of liaison users in five social classes: A = highest social class; E = lowest social class (for further details and for number of individuals per social class, see Appendix); RS: reading style; WLS: word-list style*

	A	B	C	D	E	Style
debacles_inattendues	88	67	17	50	25	RS
certains_executants	100	100	67	50	50	RS
moins_important	100	100	75	67	40	WLS
font_interessant	100	100	58	67	20	WLS

Table 7. *Liaison scores in percentages after* est, sont, *and* suis *in five social classes; N = total number of liaison contexts (style: careful speech in interviews)*

est		*sont*		*suis*		Social class
%	N	%	N	%	N	
43	58	22	18	14	29	A
65	43	35	17	16	18	B
40	66	14	21	27	22	C
24	29	13	8	13	8	D
29	17	0	4	10	10	E
Total 43	213	20	68	87		

Table 6 could be extended with dozens of examples, all showing the same pattern: members of the higher social classes use more liaison than those of the lower ones (though there are some irregularities, which, due to the relatively limited number of speakers per social class, is not surprising). We also see that optional liaisons are used by members of the lower social classes as well.

4.2. *Age*

Age is another relevant factor. Examples are given in Tables 8 and 9. The examples show that in general older people tend to use more liaison than younger people. Two explanations can be given. On the one hand the results might mean that people start using more liaison when they grow older, a phenomenon currently referred to as 'age grading'. On the other hand it might be an indication (in apparent time) of an ongoing change in liaison usage. The latter possibility shows that it is not a priori clear whether the statements of Klausenburger (1984: 55) and Encrevé (1983: 48) that there has been no shift in liaison usage from the 1940s to

Table 8. *Percentage of liaison users in three age groups: 1 = youngest generation; 3 = oldest generation (for further details and for number of individuals per age group, see Appendix); RS: reading style; WLS: word-list style*

	1	2	3	Style
depuis_un mois	33	35	67	RS
depuis_un an	30	89	50	WLS
pendant_un moment	0	22	50	RS
pendant_un moment	22	67	67	WLS

Table 9. *Liaison scores in three age groups: 1 = youngest generation; 3 = oldest generation; N = total number of liaison contexts (style: careful speech in interviews)*

	est		*sont*		*suis*		
	%	N	%	N	%	N	Age
	33	72	4	24	18	33	1
	35	84	16	31	18	34	2
	67	57	62	13	21	20	3
Total	42	213	20	68	19	87	

the 1970s can be generalized as applying to the whole range of speakers of standard French, for instance to those of Tours.

4.3. *Sex*

Finally, in several contexts, liaison frequencies were different for men and women (see Table 10). Women use somewhat more liaison than men do, a phenomenon also shown by other sociolinguistic variables (see Trudgill 1983: 84–102).

Table 10. *Percentage of liaison users among men and women (M: men; W: women) RS: reading style; WLS: word-list style*

	M	W	Style
quand_on est	80	95	RS
quand_ils sont	75	100	WLS
souvent_ensemble	0	20	RS
souvent_absent	55	25	WLS
depuis_un mois	15	40	RS
depuis_un an	25	50	WLS

4.4. *The role of style in the use of liaison*

Figures 1 to 6 confirm that liaison usage increases when speakers pay more attention to their speech. This kind of style shifting appears within social classes (Figures 1 to 5) as well as within age groups (Figure 6). The data are from De Jong et al. (1981).

The examples differentiate between reading style (formal) and word-list style (very formal). Labov (1972) elicited five different styles (casual and careful speech, reading style, word-list style, and minimal pairs) for several linguistic variables. Most likely the same can be done for liaison, and, as for the variables examined by Labov (1972), we would probably find a gradual increase in the degree of liaison usage correlated with a

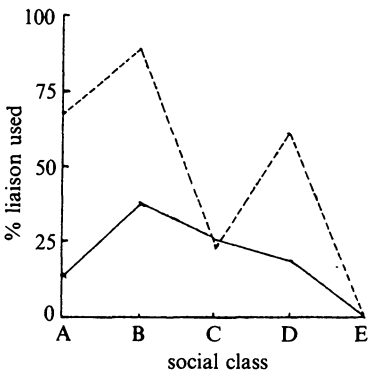

Figure 1. *pendant un moment*

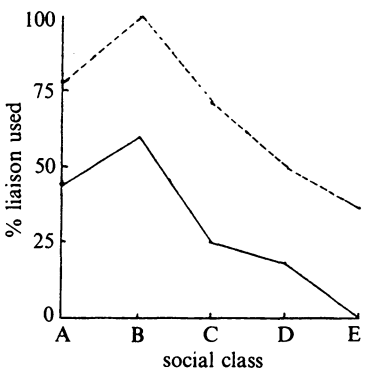

Figure 2. *après être*

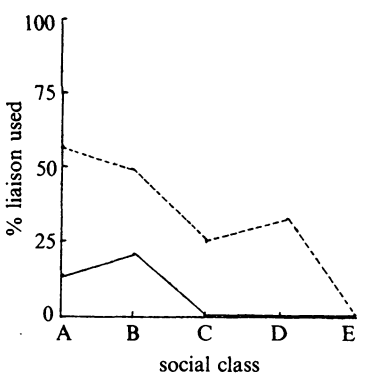

Figure 3. *toujours en*

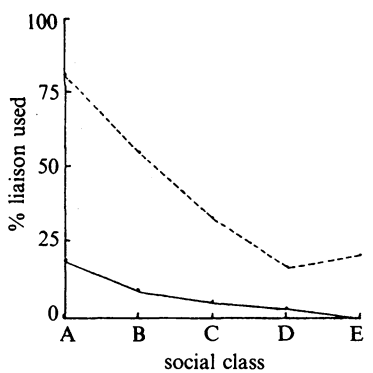

Figure 4. *assez humide*

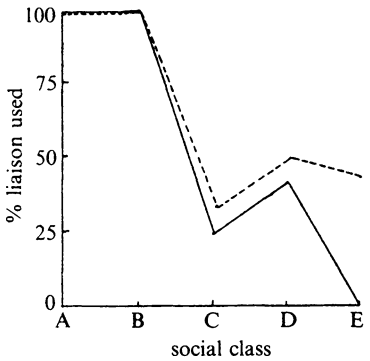

Figure 5. *autres activités*

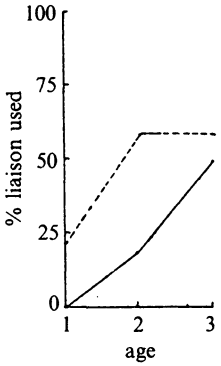

Figure 6. *pendant un moment*

gradual increase in the degree of monitoring. This once again sheds doubt on Selkirk's distinction of two discrete liaison domains, one for obligatory and one for optional application. The existence of a continuum is much more likely.

5. Conclusions

Current theories of the domain of liaison do not provide an adequate definition of the domain restrictions on liaison. We have shown that in a number of contexts for which it has been claimed that obligatory application of liaison is required, it is in fact optional. This also implies that theories which are forced to distinguish between normal (core or phonological) liaison and abnormal (syntacticized) liaison, such as the theories of Selkirk (1984) and Kaisse (1985), impose an arbitrary distinction on the liaison data. We conclude, therefore, that 'domain' is an intralinguistic variable constraint on the variable rule of liaison rather than an absolute condition on its application. The variable-rule status of liaison is confirmed by the fact that its rate of application covaries with extralinguistic factors such as style, social class, sex, and age.

Our liaison data have also thrown some light on Kaisse's hypothesis that the rules of sentence phonology consist of two discrete categories, external sandhi rules and fast-speech rules: liaison, although subject to structural restrictions, is also affected by phonetic pauses, according to Kaisse a distinguishing property of fast-speech rules.

A final, methodological conclusion is that theories of sentence phonology should be based not only on intuitions or (possibly prescriptive)

grammars, but also on corpora of spoken language with sufficient diversification with respect to extralinguistic dimensions of variation such as style and social class.

Received 14 January 1986
Revised version received
10 December 1986

Free University, Amsterdam

Appendix

In this paper we refer to the following corpora:

1. *The radio corpus of Agren (1973)*

Agren's corpus consists of 134 radio programs, recorded in 1961 and 1962, representing in total about 40 hours of speech. The informants were mainly upper-middle-class people, such as journalists, writers, and politicians. The corpus has been subdivided into three parts according to the speech style in the programs. In 43 programs the speech could be characterized as elevated and in 41 as informal, while for the remaining 50 the style was in between formal and informal. In total 8441 liaison contexts have been analyzed.

2. *The political leaders' corpus of Encrevé (1983)*

Encrevé's corpus consists of radio and TV performances (speeches, debates, and press conferences) of 21 political and union leaders in France and was recorded in the period 1978–1981. The total number of liaison contexts amounts to 10,816, of which 5029 belong to contexts where liaison is realized variably (see section 4.1).

Encrevé also analyzed the liaison usage of six political leaders for the period 1928–1978. This part of the corpus contains 4023 optional liaison contexts.

3. *The Parisian corpus of Malécot (1979)*

Malécot's corpus consists of 50 conversations averaging 30 minutes in duration with the members of the educated middle class of Paris. The social variables taken into account are occupation, age, and sex. The corpus is also subdivided according to the subject matter of the conversations and the attitude or posture of the informants, as well as rate of speech. 4409 liaison contexts have been analyzed.

4. *The pilot study of Kovac (1979)*

In a pilot study on liaison Kovac used four male native speakers of French (ages 19, 33, 47, and 52): three of them were university professors, the fourth was a university student. All of them were living in the United States.

The informants were first interviewed for 10–15 minutes. They were then asked to read a short text, and finally they had to read a phrase list. Using this procedure, Kovac tried to elicit three speech styles: careful speech, reading style, and word-list style (see Labov 1972).

5. *The corpus of Morin and Kaye (1982)*

The corpus of Morin and Kaye consists of 66 sentences (all containing a liaison context) administered to nine speakers, six from France (university professors or high school teachers) and three from Quebec (graduate and undergraduate students at the Université de Montréal). The informants were asked to read the sentences as naturally as possible in an informal style, and then in elevated speech. There were no significant differences in liaison usage between the French and the Canadian speakers.

Additionally, Morin and Kaye recorded the cases of elevated liaison (that is, in the context of inflected head plus complement) and their absence in the speech of TV reporters and the people interviewed on Radio Canada news programs in Montreal. They also recorded hypercorrections like *un gros-t enfant* 'a big child'.

6. *The Tours corpus of De Jong et al. (1981)*

The Tours corpus consists of interviews with 38 informants. For selection of the informants the stratification of the INSEE (Institut National de la Statistique et des Etudes Economiques), shown in Table A was used. The informants were divided into three age groups (18–30 years [1], 30–50 years [2], and 50 and over [3]). The number of men and women was about equal (see Table B).

The informants, who had all lived in Tours for at least 20 years, were first interviewed for about 30 minutes. The main goal of the interviews was to make the informants speak as freely as possible. The speech can be characterized as 'careful' (see Labov 1972). Second, the informants were asked to read a text, containing

Table A. *INSEE stratification used for the Tours corpus*

A.	*Professions libérales et cadres supérieurs* (higher professional)
B.	*Cadres moyens* (lower professional)
C.	*Employés* (skilled nonmanual)
D.	*Contremaîtres, ouvriers qualifiés* (skilled manual)
E.	*Ouvriers spécialisés, manoeuvres* (semi- and unskilled)

Table B. *Tours informants according to social class, age, and sex*

Social class	Age:						Total	
	18–30		30–50		50+		M	W
	M	W	M	W	M	W		
A.	3	1	2	0	1	2	6	3
B.	0	1	1	2	1	1	2	4
C.	1	4	2	2	2	1	5	7
D.	3	0	1	2	0	0	4	2
E.	1	1	1	1	0	1	2	3
Total	8	7	7	7	4	5	19	19

about 50 liaison contexts. Because of the relatively high degree of self-monitoring, this kind of speech is rather formal. It will be referred to as 'reading style'. Finally, the informants had to read a list of 45 short sentences, all containing a liaison context. In this way, the most formal style of speech was elicited.

7. *The Montréal corpus of Cedergren and Sankoff (1972)*

For the Montréal corpus, 120 informants were randomly selected on the basis of a six-level income stratification. The breakdown of the corpus into subgroups of income level, age, and sex is shown in Table C.

107 of the informants were born in Montréal; the others had lived there from at

Table C. *Quotas set for the number of respondents to be interviewed according to age, sex, and subsection of the sample (adapted from Sankoff and Sankoff 1976: 19): 1 = $5100 and over; 2 = $4100–5099; 3 = $3600–4099; 4 = $3100–3599; 5 = $2600–3099; 6 = $2200–2599*

Level of sample	Age:								Total	
	15–19		20–34		35–54		55+		M	W
	M	W	M	W	M	W	M	W		
1.	3	2	2	3	2	3	3	2	10	10
2.	2	3	3	2	3	2	2	3	10	10
3.	2	3	4	1	1	4	3	2	10	10
4.	3	2	2	3	3	2	2	3	10	10
5.	2	3	2	3	3	2	3	2	10	10
6.	3	2	2	3	3	2	2	3	10	10
Total	15	15	15	15	15	15	15	15	60	60

least the age of six. The informants were interviewed about topics likely to provoke spontaneous speech. The average length of the interviews was one hour. A detailed description is given in Sankoff and Sankoff (1976).

Over 13,000 liaison contexts in the speech of 100 informants were taken from the corpus and analyzed by Tousignant (see Tousignant and Sankoff 1979).

8. *The Orléans corpus*[6]

For the Orléans corpus, 147 informants, randomly selected, were interviewed. A division into subgroups is shown in Table D. As in the case of the Montréal and the Tours corpora, the main goal of the interviews was to obtain as spontaneous speech as possible. The average length of the interviews is 70 minutes.

Additionally, the following recordings were made:

1. 43 informal recordings with the interviewed informants in the presence of friends and/or family;
2. 36 recordings of the interviewed informants during their work;
3. 29 recordings of debates and discussions;
4. 87 interviews with locally important people;
5. 51 telephone conversations with the interviewed informants;
6. 84 conversations in shops and in the street, recorded with a hidden microphone.

The size of the Orléans corpus is estimated at 4,500,000 tokens. A detailed description of the Orléans corpus can be found in Blanc and Biggs (1971) and Mullineaux and Blanc (1982).

Table D. *Orléans informants according to age, sex, and social class; numbers in the left column indicate social class (1: higher professional; 2: lower professional; 3: skilled nonmanual; 4: skilled manual; 5: semi- and unskilled) (adapted from Mullineaux and Blanc 1982: 21)*

Social class	Age:						Total	
	50+		30–49		18–30		M	W
	M	W	M	W	M	W		
1.	4	4	8	5	3	3	15	12
2.	1	6	10	7	6	5	17	18
3.	6	5	5	8	4	8	15	21
4.	5	7	6	5	5	2	16	14
5.	5	5	2	4	2	1	9	10
Total	21	27	31	29	20	19	72	75

Notes

* This paper is part of research project no. LETT 83/7, Faculty of Letters, Free University, Amsterdam. The research of the second author has been made possible by the Stichting Taalwetenschap, which is subsidized by the Dutch Organization for the Advancement of Pure Research, Z.W.O., grant no. 300-165-019.

We would like to thank Yves-Charles Morin and two anonymous referees of *Linguistics* for their constructive criticism of a previous draft of this paper.

Correspondence address: Dr Geert Booij, Department of General Linguistics, Free University, P.O. Box 7161, 1007 MC Amsterdam, The Netherlands.

1. See Clements and Keyser (1983) and Booij (1984) for a defense of this view of liaison.
2. Encrevé proposes one modification: liaison after monosyllables has become variable.
3. /p/ is a latent consonant in *beaucoup* and *trop* only. /k/ liaison is limited to the word *long*, like in *un long espoir* [œ̃ lɔ̃kɛspwar] 'a long hope'.
4. /n/ liaison is accompanied by a complex denasalization process, as in *mon ami* [mɔ̃nami] 'my friend' with nasalization, and *bon ami* [bɔnami] 'good friend' without nasalization.
5. This is also illustrated by several remarks made by Kaisse; we mention the following two:

   'It is an unfortunate fact that much of the literature on liaison is prefaced with a paragraph disagreeing with the basic data on which some previous analysis was based' (Kaisse 1985: 163).

   'I have not treated the case of a copula plus predicate in the text because such contexts are the subject of disagreement among grammarians (and of wavering judgment on the part of my informants as well as Selkirk's); some say they optionally exhibit liaison even in casual style, while others place such liaisons in the more formal *conversation soignée*' (Kaisse 1985: 166).

6. A Z.W.O.-sponsored research project is being carried out to determine which intra- and extralinguistic factors influence liaison. As source materials the Orléans corpus will be used.

References

Agren, J. (1973). *Etude sur quelques liaisons facultatives dans le français de conversation radiophonique*. Uppsala: Ka-We Tryck.

Blanc, M., and Biggs, P. (1971). L'enquête sociolinguistique sur le français parlé à Orléans. *Le français dans le monde* 85, 16–25.

Booij, G. E. (1984). French C/Ø alternations, extrasyllabicity and lexical phonology. *Linguistic Review* 3, 181–207.

Clements, G. N., and Keyser, S. J. (1983). *CV Phonology. A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press.

Delattre, P. (1966). *Studies in French and Comparative Phonetics*, 39–62. The Hague: Mouton.

Encrevé, P. (1983). La liaison sans enchaînement. *Actes de la recherche en sciences sociales* 46, 39–66.

Grammont, M. (1938). *La prononciation française*. Paris: Librairie Delgrave.

De Jong, D., Poll, E., and Woudman, W. (1981). La liaison: l'influence sociale et stylistique

sur l'emploi de la liaison dans le français parlé à Tours. Unpublished master's thesis, University of Groningen.
Kaisse, E. M. (1985). *Connected Speech. The Interaction of Syntax and Phonology*. Orlando: Academic Press.
Kay, P., and McDaniel, C. K. (1979). On the logic of variable rules. *Language in Society* 8, 151–187.
Klausenburger, J. (1984). *French Liaison and Linguistic Theory*. Stuttgart: Steiner.
Kovac, C. (1979). The role of speech style in the realization of optional liaison in French: a pilot study. *La Linguistique* 15, 149–154.
Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
—(ed.) (1980). *Locating Language in Time and Space*. New York: Academic Press.
Malécot, A. (1979). French liaison as a function of grammatical, phonetic and paralinguistic variables. *Phonetica* 32, 161–179.
Malmberg, B. (1972). *Phonétique française*. Malmö: Hermods.
Morin, Y.-C., and Kaye, J. (1982). The syntactic bases for liaison. *Journal of Linguistics* 18, 291–330.
Mullineaux, A., and Blanc, M. (1982). The problem of classifying the population sample in the sociolinguistic survey of Orléans (1969) in terms of socio-economic, social and educational categories. *Review of Applied Linguistics ITL* 55, 3–38.
Nespor, M., and Vogel, I. (1982). Prosodic domains of external sandhi rules. In H. van der Hulst and N. Smith (eds.), *The Structure of Phonological Representation, Part I*, 225–256. Dordrecht: Foris.
Pulgram, E. (1965). Prosodic systems: French. *Lingua* 13, 125–144.
Robach, I.-B. (1974). *Etude sociolinguistique de la segmentation syntaxique en français parlé*. Lund: Gleerup.
Romaine, S. (1984). The status of sociological models and categories in explaining language variation. *Linguistische Berichte* 90, 25–35.
—(1985). Variable rules O.K.? Or can there be sociolinguistic grammars? *Language and Communication* 5, 53–67.
Sankoff, D. (1978). *Linguistic Variation: Models and Methods*. New York: Academic Press.
—, and Sankoff, G. (1976). Sample survey methods and the computer assisted analysis in the study of grammatical variation. In R. Darnell (ed.), *Canadian Languages in their Social Context*, 7–64. Edmonton: Linguistic Research.
Selkirk, E. O. (1974). French liaison and the X′ notation. *Linguistic Inquiry* 5, 573–590.
—(1978). On prosodic structure and its relation to syntactic structure. Unpublished manuscript, University of Amherst, Mass.
—(1980 [1972]). *The Phrase Phonology of English and French*. Bloomington: Indiana University Linguistics Club.
—(1984). *Phonology and Syntax. The Relation Between Sound and Structure*. Cambridge, Mass.: MIT Press.
Tousignant, C., and Sankoff, D. (1979). Aspects de la compétence productive et réceptive: la liaison à Montréal. In P. Thibault (ed.), *Le français parlé. Etudes sociolinguistiques*. Edmonton: Linguistic Research.
Trudgill, P. (1983). *On Dialect: Social and Geographical Perspectives*. Oxford: Blackwell.