INVESTIGATING SYLLABIC STRUCTURE AND ITS VARIATION IN SPEECH FROM FRENCH RADIO INTERVIEWS

Martine Adda-Decker, Philippe Boula de Mareüil, Gilles Adda & Lori Lamel

LIMSI-CNRS, BP 133
91403 Orsay cedex, FRANCE
{madda,mareuil,gadda,lamel}@limsi.fr

ABSTRACT

In this paper, we investigate syllabic structure and its variation in a corpus of French radio interview speech. The aim of this study is to relate sequential pronunciation variants, i.e. variants with different numbers of phonemes to syllabic restructuring. In French *schwa* and *liaison* are two well-known phenomena which allow for a variable number of phonemes. We first aim to quantify syllabic restructuring due to these phenomena. Our second aim is to identify other syllabic restructuring phenomena due to omitted vowels (i.e. syllable nuclei) or even omitted syllables.

The radio speech corpus is comprised of 30 1-hour shows of interviews mostly involving one professional anchor speaker and an artist or a politician. The speech style is fluent, spontaneous and only partially prepared. Syllable distributions computed from a word level representation are compared to those emerging from speech. Results confirm that the optional schwa vowel contributes to a large amount of variation in syllabic structure. Less well-described phenomena have been observed: other vowels than schwa, such as /u/, /e/ and /Œ/ appear to be optional in unstressed contexts. A substantial percentage of occurrences of word-final syllables may completely disappear.

1. INTRODUCTION

Speech recognition has made tremendous progress this past decade, with significant reductions in recognition word error rates. However current system performance remains significantly worse than human performance [16] on similar tasks. Present challenges concern improved language modeling and pronunciation modeling. Especially for spontaneous speech the problem of variant modeling appears to be crucial. In this type of speech a major problem is related to reduced pronunciations, i.e. pronunciations where the number of uttered phonemes is less than the number of theoretically expected phonemes. Spontaneous speech studies have shown that syllable composition and their number can be severely altered within and across words.

Recent progress in speech modeling provides the opportunity of using a speech recognizer to analyse large acoustic corpora in order to study pronunciation variants. This is the primary goal of this contribution. Instead of limiting the linguistic representations to the word and phoneme levels, as often is the case, we introduce an intermediate syllable level, which is described in the following sections. This intermediate syllable level allows us to examine observed variants with respect to expected syllabic structures. Other researchers have used the syllable level to investigate speech corpora and recognition errors [1].

In the next section syllabification rules of spoken French are described. We emphasize the effects of schwa and liaison on the syllabic structure. Section 3 briefly describes the speech corpus. In section 3 the general methodology of aligning transcribed speech on spoken syllables via written syllables is outlined. Section 4 explains the link between the word and an intermediate written language-based syllable level. Section 5 focuses on the spoken syllable level as most commonly adopted by linguists, and Section 6 presents some results on syllable restructuring.

2. SYLLABIFICATION RULES IN FRENCH

In psycholinguistics, syllables are often considered as the information processing units of perceptual mechanisms, for acoustic-phonetic decoding. Yet, syllabification, that is the segmentation of the spoken string into syllables, differs from one language to another, and a universal phonological theory does not exist. Syllabification also depends on the linguistic communities' conventions. In English for example some researchers do not agree on the number of syllables in words such as *communism*, *hour*, *real*. Since Saussure[14], a hundred years ago, various theories have been proposed to account for the tendency of some consonant sequences to be split. According to the so-called Sonority Sequencing Principle, clusters of more than two consonants containing an /s/ undergo a syllabic break after the latter (e.g. *obstruer*, /ɔps-tʁɥe/). The tautosyllabicity of this /s/ with regards to the following consonant is controversial, since a word may begin with what Italian grammar calls "impure *s*" (e.g. *sport*), without being dissyllabic. Languages exhibit different syllable structures. A study by Delattre[13] found a rate of 0.8 consonants per syllable in French and Spanish, 1.0 in English and 1.2 in German.

In French each syllable contains one and only one vowel and a consonant alone cannot constitute a syllable. Irrespective of the grammatical and orthographical word tokenization, in French, each consonant belongs to the same syllable as the vowel immediately following: in particular, a syllabic break falls before an intervocalic consonant.

So far, we have spoken about phonological syllables, which do not always correspond to graphic syllables (typographical hyphenation), partly because of the schwa (/ə/). The schwa, which may or may not be spoken (thus influencing the number of syllables), is one of the most intricate aspects of French phonology[4, 12]. Consider the word "*amener*" (to bring). It has three graphic syllables "*a-me-ner*", corresponding to three or two phonological syllables /a-mə-ne/ or /am-ne/ depending whether the ə is maintained or not. Even if it enables a phonological opposition between words – e.g.

pelage (/pəlaʒə/, "coat") vs *plage* (/plaʒə/, "beach") –, the schwa vowel is most often optional.

Liaison is another phenomenon which complicates the syllabification process. Liaisons consists in the realization of a normally mute final consonant in the context of a following word which begins with a vowel, thus avoiding a hiatus. For example the words *les îles* ("the islands") pronounced /le/ and /il/ in isolation are pronounced /lezil/ together. Liaison generally results in cross-word syllabification: /le-zil/. How and when is liaison made? We are here in a ticklish field [7, 10, 11], and there is no consensus to answer this question which goes beyond the scope of this paper. Liaison should not be confused with chaining. Liaisons without chaining (i.e. no cross-word syllabification : /lez-il/) can be heard, particularly in political debates[8]. A limited number of consonants are used for liaison: /z/, /t/, /n/, /ʁ/, /p/ – the rank order is from highest to lowest for the frequency of occurrences. Cross-word syllabification makes the word boundary recognition and thus lexical access perceptually more difficult.

2.1. Syllabification procedure

The adopted syllabification procedure is part of our grapheme-to-phoneme converter GRAPHON+[3]. Syllabification can be optionally carried out after the grapheme-to-phoneme conversion proper. The syllabic break is obtained by using the first applicable rule among the list of rules given in Table 1. The same syllabification procedure (without grapheme-to-phoneme conversion) is applied to the aligned phone sequence (see Section 5).

Sequence		syllabified	example	pronounced
əC{0;4}V	→	ə-C{0;4}V	r**e - froi**di	ʁə. frwɑdi
VV	→	V-V	r**é - a**lise	ʁe . aliz
V . V	→	V - . V	**i - ma**ge	i . mɑʒ
V . GV	→	V - . GV	st**u - dio**	sty . dɥo
VOLV	→	V - OLV	p**u - bli**c	py . blik
V . . V	→	V . - . V	**ob - je**t	ɔb . ʒɛ
VOLGV	→	V - OLGV	**em - ploi**	ɑ̃ . plwɑ
V . . GV	→	V . - . GV	v**ic - toi**re	vik . twɑʁ
V . OLV	→	V . - OLV	**es - pri**t	ɛs . pri
V . . . V	→	V . . - . V	**ex - pe**rt	ɛks . pɛʁ
V . OLGV	→	V . - OLGV	**al - trui**ste	ɑl . trɥist
V V	→	V . . - . . V	**ex - pie**r	ɛks . pje
V . . . . GV	→	V..-..GV	**ex - ploi**t	ɛks . plwɑ

Table 1. Syllabification rules for French. The syllabic break is noted by a dash in the right part of the rules. C{0;4} stands for 0 to 4 consonants, ə=maintained schwa, V={vowels}, L={liquids}, G={glides} O={obstruents: plosives, fricatives or nasals}, · = any phoneme.

2.2. Syllabic structures of standard French

The syllabic structures of standard French, resulting from a manual syllabification on a corpus of spoken utterances, are reported in Table 2. The French language appears to prefer free (or "open") syllables, accounting for 80% of all syllables. With about 55% of occurrences, the CV type, which is the least marked syllable, is the most frequent. Liaisons, chaining as well as phenomena such as the use of *cet* for *ce* ("this"), *mon* for *ma* ("my") before vowel contribute to this trend (increasing the number of free syllables and decreasing the number of syllables with empty onset).

syllables	example	pronunciation	resyllabification
CV	*veau*	vo	**CV**
CCV	*gré*	gʁe	**CCV**
CVC	*masse*	mɑs{ə}	CV - CV
V	*eau*	o	**V**
CCVC	*grade*	gʁɑd{ə}	CCV - CV
CVCC	*test*	tɛst{ə}	**CVC** - CV
VC	*hâte*	ɑt{ə}	V - CV
CCCV	*strie*	stʁi	**CCCV**
CCVCC	*Brest*	bʁɛst{ə}	**CCVC** - CV
CCCVC	*strate*	stʁɑt{ə}	CCCV - CV
VCC	*ogre*	ɔgʁ{ə}	V - CCV
CVCCC	*filtre*	filtʁ{ə}	CVC - CCV
CCCVCC	*strict*	stʁikt{ə}	**CCCVC** - CV
CVCCCC	*dextre*	dɛkstʁ{ə}	**CVCC** - CCV
#distinct : 14			#distinct : 8

Table 2. Syllabic structures of standard French using C (consonants) and V (vowels) classes. The last column shows resyllabification if a schwa-vowel is produced. The total number of syllable types is reduced from 14 to 8.

Nevertheless, colloquial French forms such as *d'jà* for *déjà* ("already"), *déj'ner* for *déjeuner* ("lunch") and *m'sieur* for *monsieur* ("Sir") may be observed, where the drop of an unstressed vowel leads to a re-syllabification. The consequences on the syllable structure distribution and possible problems for speech recognition are investigated in this paper. At the same time, a speech recognition system can be used to automatically label very large speech corpora in order to carry out further linguistic analyses. By aligning the data with acoustic word models which allow for pronunciation variation, the observed alignments provide frequencies for the variants involved in the corpus. Explanations for the observed variants can be proposed at a linguistic level, by the speech data characteristics, or at a speech engineering level, by the properties of the acoustic models.

In the following sections, the speech corpora and methodology used in this study are described, and results are presented.

3. SPEECH CORPUS AND APPROACH

The speech corpus used in this study contains 30 1-hour shows of interviews involving most often one professional anchor speaker and one artist or politician. Some shows include more speakers. Speech is studio quality and most speakers are native. The speech style is fluent, spontaneous, and only partially prepared. All shows have been manually transcribed. The corpus contains a total of approximately 245k word occurrences, with 13.5k distinct lexical entries.

Figure 1 shows a generic syllable representation with (optional) consonantal onset and nucleus, the latter having a mandatory unique vowel and an (optional) consonantal coda. The right part gives the structure of the simplest syllable: the V-type syllable.

Syllabic structure is first defined at two levels: written language word level and then at spoken language phrase level.

For the word level we want to keep word boundary information and syllables are first defined on an isolated word basis. A syllable formalism is defined using full and partial syllables (no vowel). The need for partial syllables arises from the presence of short function words in French which are reduced to a single

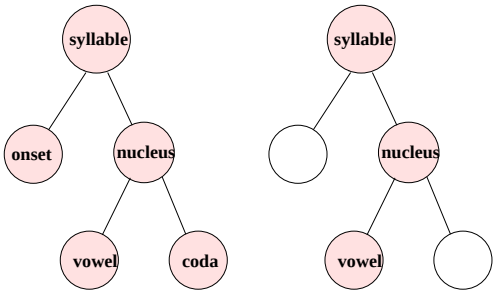

Fig. 1. **Left** : General syllable structure with optional consonantal onset, syllable nucleus, which contains always a unique vowel in French and optional consonantal coda. **Right** : example of V-syllable (onset and coda are empty).

consonant. Beyond the word boundary motivation spontaneous speech can give rise to word fragments and hesitations limited to consonantal speech segments which also have to be modelled by a partial or degenerate syllable (without vowel). Partial syllables can then be merged across word boundaries to form full syllables. A canonical phonemic transcription of each word is split into syllables using the GRAPHON+ syllabification rules. Concerning syllabification at word boundaries two situations are obviously problematic: partial syllables and liaison consonants. These are further investigated in Section 4.

At the spoken language phrase level we use also the GRAPHON+ syllabification rules to derive syllable boundaries, but without knowledge of the word boundaries. This is done on both the canonical phonemic transcriptions and on the aligned phone sequence. Details are described below.

4. WRITTEN LANGUAGE SYLLABLES

4.1. Word-level syllables

Each lexical entry is phonemically transcribed into a maximal-length canonical (MLC) phonemic string. By maximal-length we mean that all possible phonemes are supposed to be pronounced, in particular schwa-vowels. For example the word *développer ("to develop")* has the pronunciation /devəlɔpe/. This MLC string is then split into **written language syllables** or W-syllables using the syllabification algorithm described in Section 2 (see Table 3).

lexical entry	MLC string	W-syllables
une	ynə	y nə
développer	devəlɔpe	de və lɔ pe

Table 3. Examples of lexical entries, transcribed into Maximum-Length Canonical phonemic string, which is then split into Word-syllables according to the syllabification algorithm.

The most frequent words are short function words which are often monosyllabic: *de ("of"), est ("is"), je ("I"), que ("that"), et ("and"), vous ("you"), la ("the")* . . ..

Some lexical entries can only be considered as **partial syllables** as there is no vowel nucleus: *c' ("this"), l' ("the"), j' ("I"), n' ("not"), d' ("of"), qu' ("that"), m' ("me"), s' ("he")*. There are roughly ten of these entries which are reduced to one consonant: they have to be combined to a neighbouring vowel (end of a preceding word, or more typically start of a following word) to form a full (admissible) speech syllable. Whereas limited in number these entries are frequent in the language, accounting for about 5% of the corpus.

There is a total of 354k W-syllable occurrences in the corpus. There are 1050 distinct syllables which account for 99.8% of the corpus. The full syllable list (100% coverage) contains 1570 syllables. This list includes a large number of rare events : arising from foreign proper names, from word fragments (truncated words of spontaneous speech) and from errors (mainly transcription spelling errors and subsequent grapheme-to-phoneme conversion problems). Figures concerning the occurrences in the corpus of the different syllable types are given in Table 4. The observed syllable structures are identical to those presented in Section 2, with an additional C class corresponding to the partial syllables (see *W-syll isol* column).

Syllable type	*W-syll isol*	*W-syll cont*	*W-syll + Liais*
CV	57.6	63.2	68.2
V	14.6	12.0	9.7
CCV	9.8	10.5	10.7
CVC	9.2	10.3	7.9
C	4.3	-	-
VC	2.6	2.0	1.1
CCVC	1.0	1.1	0.8
CCCV	0.5	0.5	0.5
CVCC	0.3	0.3	0.3
VCC	0.2	0.1	0.1
CCVCC	ϵ	ϵ	ϵ
CVCCC	ϵ	ϵ	ϵ
CCCVC	ϵ	ϵ	ϵ
VCCC	ϵ	ϵ	ϵ
CCCVCC	ϵ	ϵ	ϵ

Table 4. Different W-syllable types observed in our corpus with their percentage of occurrence. The partial syllable C is mainly due to syllabification carried out on isolated words (*W-syll isol* column). The *W-syll cont* column gives corrected full syllable percentages, where partial syllables are glued to the following syllables. The *W-syll + Liaison* column corrects liaison syllabification. ϵ means that the percentage is < 0.05.

The partial syllable C is mainly due to syllabification carried out on isolated word syllables. If these partial C syllables are merged with the onset of the following syllables, figures slightly change, with a reduction of the V structure and a grow of the CV structure (see *W-syll cont* column). The CV syllable represents roughly 60% of the corpus, the V syllable around 13%, CCV and CVC occurring about 10% each. There are 46k W-syllables (13% of the corpus) which can produce a liaison consonant. Only 24% of them (11k) are in a right vowel context (the most favourable to produce liaison). Column *W-syll + Liaison* displays percentages when liaison consonants are added to the onset of the following syllable.

Ignoring the partial C syllable, the CV, V, CCV, CVC, VC and CCVC syllables account for 99% of the corpus.

4.2. W-syllable pronunciation dictionary

The lexical transcription of the corpus is augmented with a W-syllable transcription (see example in Table 5).

lex.	*je*	*pense*	*que*	*c' est*	*de*	*la*	*sagesse*	
W-syll.	ʒə	pɑ̃	sə	kə	s	ɛt*	də	la	sa	ʒɛ	sə

Table 5. Example of a sentence start (lexical form and W-syll form). The W-syllable ɛt* , corresponding to the word *est* indicates a possible /t/ liaison.

In order to align the W-syllables to the acoustic signal we introduce a pronunciation dictionary. For each W-syllable (MLC transcription) the pronunciation dictionary allows for pronunciation variants. As we are mainly interested in reduction phenomena (with a smaller number of phonemes as theoretically possible) any shorter phone sequence included in the MLC form is allowed (see Table 6).

W-syll.	pronunciations
sa	sa s a
tʁo	tʁo tʁ to ʁo t ʁ o
kʁwa	kʁwa kʁw kʁa kwa ʁwa kʁ kw ka ...

Table 6. Excerpt of the W-syllable pronunciation dictionary: the left side corresponds to the W-syllable and the right part to the optional smaller length pronunciations.

4.3. Optional W-syllables

In the W-syllable pronunciation dictionary each entry can be shortened down to one phoneme. Beyond these reductions, we want W-syllables to be optional: if a W-syllable has not been uttered it should be possible to skip it.

Alignments are carried out using a W-syllable graph corresponding to the W-syllable transcription, where every other syllable may become optional as shown in Figure 2.

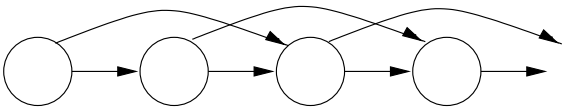

Fig. 2. Speech alignment is carried out using a W-syllable graph where every other syllable is optional.

5. SPOKEN LANGUAGE SYLLABLES

This section concerns the description and analysis on a spoken syllable basis. Here we don't consider word boundaries during the syllabification process. The MLC phonemic string as well as the aligned string can be syllabified using the GRAPHON+ rules (described in section 2), thus producing S-syllables (referred to a S-s and S-a respectively). Table 7 shows an example sentence start with corresponding W and S-syllables. Syllabification on the MLC phonemic string produces about 340k S-syllables. This is roughly 4% less than the W-syllable syllabification (explained by the absence of partial syllables). Using the aligned phonemic string, a total of 290k S-syllables is measured. This corresponds to a 15% deletion rate. The number of distinct syllables is increased from 1560 (W-syllables) to 2250 S-syllables (S-s). The additional syllables are due to cross-word syllabification and result for a large amount in CVC and CCV syllables. Cross-word syllabification on the aligned sequences (S-a), introduces approximately 6000 distinct syllables. The possible alignment of non-standard phone sequences (as compared to standard MLC sequences) is the explanation for the larger number of distinct syllables here. 2900 syllables occur more than once (coverage of 98.5%) and 900 more than 10 times (coverage of 94.9%). In comparison for W-syllables we have 1500 syllables occuring more than once (coverage of 99.7%) and 800 more than 10 times (coverage of 98.4%). Many of the observed S-syllables are in common with the W-syllables. But cross-word resyllabification allows to create new syllables not occurring in isolated French words. For example there are 27 syllable starts with /ʒl/ corresponding to a resyllabification of word sequences like *je le ... (I ... so)*. For example the syllable /ʒlɛs/ may arise from the word sequences *je l' espère (I hope so), je laisse (I let),* ....

lex.	je	pense		que	c'	est	de	la	...
W-s	ʒə	pɑ̃	sə	kə	s	ɛt*	də	la	...
W-a	ʒə	pɑ̃	s	kə	s	ɛ	d	la	...
S-s	ʒə	pɑ̃	sə	kə	sɛ		də	la	...
S-a	ʒə	pɑ̃s		kə	sɛ		dla		...

Table 7. Example of a sentence start (lex.), corresponding W-syllables (W-s), and aligned phones (W-a) , and finally S-syllables (S-s: from MLC phonemic sequence, S-a : from aligned sequence) using GRAPHON+ resyllabification rules.

Table 8 shows syllable structures of S-syllables (both S-s and S-a). We can note that the more complex syllables (CCVC, CVCC) are significantly more frequent for the aligned S-a syllables than for the standard S-s syllables. The overall percentages measured for the main syllable types remain nonetheless similar to the percentages measured for the W-syllables (see table 4). The relatively high V-syllable rate (14.6%) obtained for isolated W-syllables is reduced here to about 12%. A smaller rate of V-syllables could have been expected, given the cross-word context and measured vowel deletions. Investigating the automatic alignment and syllabification results, we could observe that simple vowels (often the schwa vowel) are sometimes aligned with unclearly uttered syllables (e.g. repetitions of word fragments). Such alignments produce V syllables.

The most frequent CV syllable structure represents 60% of the corpus (see Table 8). Open syllables (CV, V, CCV, CCCV) account for 82.5% of the corpus. The most frequent closed syllable structure is CVC (11.6%).

6. RESULTS

Syllabic restructuring can be observed by deleted syllables and by changes in the syllabic structure distributions. Figure 3 illustrates a typical resyllabification of two consecutive CV syllables into a CVC syllable (e.g. *toute forme* /tu təfɔRm/ →) tut fɔRm). This kind of resyllabification is rather common, as indicated by the CVC figures in Table 8

syll. type	%S-s	%S-a
CV	67.3	60.4
V	11.8	12.5
CCV	10.5	9.2
CVC	7.6	11.6
VC	1.1	1.6
CCVC	0.6	1.4
CVCC	0.4	1.4
CCCV	0.4	0.4

Table 8. Most frequent spoken syllable types in French. The S-s column derives from the syllabified MLC sequence, the S-a columns from the syllabified aligned phone sequence.

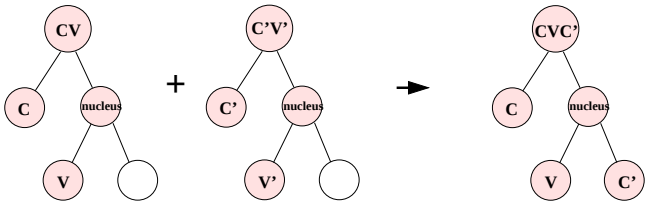

Fig. 3. Resyllabification producing closed CVC' syllables due to V' vowel omission.

6.1. Deleted syllables

Using the W-syllables alignment with the optional syllable graphs 6% of W-syllables were skipped (deletions). A small part of these missing syllables must be attributed to transcription errors: within spontaneous speech there are segments which are often difficult to transcribe at a word level because of unclear articulation, hesitation and repetition of word fragments. An important part however correspond to well-known linguistic phenomena.

Among the observed deletions, 40% (i.e. 9k occurrences out of 24k) correspond to syllables containing a schwa (see Table 9). Another important part of the remaining deleted W-syllables are V syllables, mainly ɑ, ɛ, e, y, ɛ̃, i, ɔ, o, ɑ̃, u (see Table 10). They correspond to another 30% (7k occurrences out of 24k). Among the partial syllables, the most omitted are *n', l', d', qu', j'* (see Table 11). More complex syllable deletions concern the words *il, vous, la, les, fait, par, qui* which are of VC, CV and CVC types and word endings with CCV (see Table 12).

Using the S-syllables alignment (S-s, S-a) about 15% of syllables are omitted (54k occurrences). This shows that, among the large number of omitted syllables, more than the half are cross-word syllables: parts of words on word boundaries disappear more easily than regular W-syllables.

6.2. Missing vowels

As observed already before vowels in general are prone to deletion (not only the schwa vowel). Vowel deletion occurs in VV sequences within words, but more typically on word boundaries. Vowels are also prone to deletion in N (nasals) and L (liquids) contexts.

In the following are some examples of observed vowel deletions:
extraordinaire (/ɑɔ/ → /ɑ/)
mais enfin (/mɛɑ̃/ → mɑ̃/)
cinéma (/nem/ → /nm/)

word	W-s syllables	W-a syllables	#occ (%)
de	də	-	1465 (15.5%)
ne	nə	-	1051 (14.7%)
te	tə	-	1002 (18.3%)
le	lə	-	919 (11.9%)
que	kə	-	653 (8.5%)
je	ʒə	-	599 (8.5 %)

Table 9. Examples of CV əvowel W-syllable deletion. For each W-syllable function word the canonical syllable transcription (W-s syllable) is given along with the aligned form in speech (W-a syllable). #occ gives the number of deleted syllables in the corpus. The percentage of deletion for each syllable is shown in brackets. For each of these function words deletion rates vary between 8.5% and 18.3%.

carrier word	W-s syllables	W-a syllables	#occ (%)
a	a	-	2279 (14.9%)
ai	ɛ	-	1012 (17.1%)
une	y nə	- nə	371 (11.8%)

Table 10. Examples of syllable deletion not including the ə vowel in vowel function words.

comment (/kɔm/ → km/)
politique (/pɔl/ → pl/)
solitaire (/sɔl/ → sl/)
voulait (/vul/→ vl/)
personnel (/sɔn/ → sn/)
philosophie (/lɔz/ → lz/)
c'est à (/sɛtɑ/ → stɑ/)
c'est pas (/sɛpɑ/ → spɑ /)
je sais pas (/ʒəsɛpɑ/ → ʒspɑ/)

6.3. Reduced consonant clusters

The truncation of words such as *montre* ("watch" or "show") and *prendre* ("to take") resulting from the drop of the final liquid is a well-known phenomenon in spoken French. The analysis of our data confirms that, for words in *-tre* and *-dre* preceding a consonant, the pronunciations [t] and [d] (rather than [tʀ] and [dʀ] respectively) are preferred, in an average ration of 3 to 2: after the elision of the schwa in this context, the liquid falls in 240 occurrences, and is maintained together with the plosive in 170 occurrences. This way, too massive a violation of the three consonant law is avoided.

7. CONCLUSIONS & PERSPECTIVES

In this contribution we described a new methodology of carrying out corpus analysis on a syllable basis. The use of W-syllables allows to relate word level to spoken syllables. For the different word and speech syllables used we found relatively stable syllable structure distributions with the CV structure accounting for more than the half of the data. Open syllables (CV, V, CCV, CCCV) account for about 80% of the corpus. The most frequent closed syllable structure is CVC. Whereas French admits theoretically 14

carrier word	W-s syllables	W-a syllables	#occ (%)
n'	n	-	538 (24.8%)
s'	s	-	304 (8.2%)
l'	l	-	253 (8.4%)
d'	d	-	160 (7.9%)
qu'	k	-	160 (8.2%)

Table 11. Examples of **partial** W-syllable deletion. The deletion rate is particularly high for the negation *n'* (24.8%).

carrier word	W-s syllables	W-a syllables	#occ (%)
être	ɛtrə	ɛ-	214 (6.1%)
capable	ka pa blə	ka pa -	97 (11.1%)
exemple	ɛgzɑ̃ plə	ɛgzɑ̃ -	73 (16.3%)

Table 12. Examples **CCV** W-syllable deletion. These word-final syllables are particular prone to deletion or at least reduction.

different syllable structures (using C and V classes), the 6 structures CV, V, CCV, CVC, VC and CCVC syllables account for 99% of the corpus. Whereas syllable deletions are relatively frequent for S-syllables (15%) W-syllables have a much lower deletion rate (6%). Deletions mainly occur in cross-boundary positions.

In future work we intend to refine the analysis of alignment results, improve the present approach, especially the W-syllables and MLC pronunciations.
The perspectives of these studies are diverse: contribution to better quantifying well-known linguistic phenomena on large corpora; the developed framework can serve as a tool for manual transcription checking: omitted syllables point to either linguistic phenomena, or simply to transcription errors; syllable modeling for OOV words, syllable-based language modeling.

8. ACKNOWLEDGEMENT

The audio corpus used in this study was provided by the Radio Archives of INA (Institut national de l'audiovisuel) in the context of a collaboration with their research and experimentation department.

9. REFERENCES

[1] Greenberg S., Chang S., "*Linguistic dissection of Switchboard-Corpus Automatic Speech Recognition Systems*", ISCA-ITRW workshop, ASR-2000, Paris (pp. 195-202).

[2] Adda-Decker M. & Lamel L. (1999), "*Pronunciation variants across system configuration, language and speaking style*", Speech Communication 29 (pp. 83-98).

[3] Boula de Mareüil P. (1997), "*Étude linguistique appliquée à la synthèse de la parole à partir du texte*", PhD thesis, University of Paris XI, Orsay.

[4] Dauses A. (1973), "*Études sur l'e instable dans le français familier*", Niemeyer Verlag, Tübingen.

[5] Dell F. (1973), Les règles et les sons, Hermann, Paris.

[6] Durand J. & Laks B. (2000), "*Relire les phonolgues du français : Maurice Grammont et la loi des trois consonnes*", Langue française, vol. 126 (pp. 29-38).

[7] Eggs E. & Mordellet I. (1990), "*Phonétique et phonologie du français. Théorie et pratique*", Niemeyer Verlag, Tübingen.

[8] Encrevé P. (1988), "*La liaison avec et sans enchaînement. Phonologie tridimensionnelle et usages du français*", Éditions du Seuil, Paris.

[9] Fouché P. (1969), "*Traité de pronunciation française*", Klincksieck, Paris.

[10] Fougeron C., Goldman J.-P. Frauenfelder U.H. (2001), "*Liaison and schwa deletion in French: an effect of lexical frequency and competition*", Eurospeech, Aalborg (pp. 639-642).

[11] Lucci V. (1983), "Étude phonétique du français contemporain à travers la variation situationnelle", Publications de l'Université des Langues et Lettres de Grenoble, Grenoble.

[12] Verney Pleasants J. (1956), "Études sur l'e muet, timbre, durée, intensité, hauteur musicale", Klincksieck, Paris.

[13] Delattre P. (1965), "*Comparing the phonetic features of English, Spanish, German and French*", Julius Gross Verlag, Heidelberg.

[14] Saussure F. de (1915), "*Cours de linguistique générale*", Payot, Paris.

[15] Wioland F. (1991), "*Prononcer les mots du français. Des sons et des rythmes*", Hachette, Paris.

[16] Lippmann R. (1997), "*Speech recognition by machines and humans*", Speech Communication, **22**(1), pp.1-16.