The history of the Franconian tone contrast

by Paul Boersma; identical to published version, November 2017*

Abstract. The aim of this paper is to show that a sequence of typologically not unusual sound changes has led to three conspicuous properties of the dialects in a large connected area of Low and Central Franconian. First, these dialects have a binary contrast between acute and circumflex tones. Second, the majority of these dialects ("group A") show *length reversal*, in the sense that originally short non-high vowels have become longer than the corresponding originally long vowels. Third, the remaining dialects ("group B") show *tone reversal*, in the sense that where group A retains the original acutes, group B has circumflexes, and the reverse (at least in declarative intonation). This paper proposes a history consisting of a series of synchronic states connected by speakers' gradual phonetic shifts and listeners' discrete phonological reinterpretations. Each of the proposed elements is shown to have parallels elsewhere: the retraction of stress to the first mora, the lengthening of vowels in open syllables with retention of the linkage between syllables and tones, the inaudibility of tone on voiceless consonants, the drop of final schwa, the pronunciation of final voiced obstruents, the audibility of tone on voiced consonants, the devoicing of final obstruents, degemination, schwa insertion, and the effects of a markedness constraint that correlates tones and duration.

1. The tone contrast as a focus alignment contrast

In continental West-Germanic, a contiguous group of Low Franconian dialects (Limburgian or "Southern Low Franconian") and Central Franconian dialects (Ripuaric and Moselle Franconian) exhibit a binary lexical tone contrast on long vowels and diphthongs, which interacts with the intonation contour of the sentence. While this state of affairs has been known since Nörrenberg (1884), recent research by Carlos Gussenhoven and colleagues has taken large steps towards distentangling the influence of the lexical tone from the influence of the intonation contour. As a result, we can now more clearly see what the fundamental nature of the contrast is. The following real-life examples from Geleen Limburgian show that if the word **kniin** 'rabbit-PL' is in focus sentence-internally, it is pronounced with a rising or falling tone contour, depending on whether the sentence is declarative or interrogative (L = low, H = high):

(1) a. L L L L HL L L
vər ɦœbə də kniin ɣəʃlɑx
'we have slaughtered the rabbits'

b. L L L LH H HL
ɦœptxər də kniin ɣəʃlɑx
'have you-PL slaughtered the rabbits?'

* This paper is an extended version of a talk entitled "De volgorde der gebeurtenissen in de geschiedenis van het Limburgse tooncontrast" presented at a workshop by the Nederlandse Vereniging voor Fonetische Wetenschappen on "Segmentele toonverschijnselen, vooral in Limburg" in Antwerp on 23 May 2002. Thanks go to the attendants of those talks, including Michiel de Vaan and Carlos Gussenhoven, and to Wolfgang Kehrein, Björn Köhnlein, Ben Hermans and Marc van Oostendorp for their many discussions on the 2006 version of this paper.

A minimally contrasting word is **kniin** 'rabbit-SG'. If it is in focus sentence-internally, it is pronounced with a level high or low tone, depending on whether the sentence is declarative or interrogative:

(2) a. L L L L HH L L
vər fiœbə də kniin ɣəʃlɑx
'we have slaughtered the rabbit'

b. L L L LL H HL
fiœptxər də kniin ɣəʃlɑx
'have you-PL slaughtered the rabbit?'

Following Gussenhoven (2000b) for Roermond, the intonation contour of the declarative sentence in (1a) can be characterized underlyingly as an L tone at the left boundary, an L tone at the right boundary, and an H* focus tone; on the surface, H* becomes aligned with the first mora of the stressed syllable (**kniin**), and the two L tones spread from the edges in toward this focus tone, yielding (1a). The intonation contour of the interrogative sentence in (1b) is then characterized underlyingly as an L tone at the left boundary, an HL contour at the right boundary, and an L* focus tone; on the surface, L* becomes aligned with the first mora of **kniin**, and the two boundary tones spread inwards toward this focus tone, yielding (1b). Gussenhoven analyses the examples in (2) as having an additional underlying lexical tone, which surfaces on the second mora of **kniin**, as H in (2a) and as L in (2b).

For the examples in (1) and (2), a different analysis than a tonal one is possible, namely an analysis in terms of *focus tone alignment*. Both in (1a) and in (2a), there is a high focus tone on **kniin**, but the transition to the L boundary tone occurs after the first mora in (1a) but after the second mora in (2a). Likewise, both (1b) and (2b) have a low focus tone on **kniin**, but the transition to the HL boundary tone occurs after the first mora in (1b) but after the second mora in (2b). The present paper proposes, now, that the Franconian tone contrast arose historically as precisely this distinction between early and late alignment of the focus tone.[1]

In the present paper I denote the moving tones (HL or LH) of (1) by the term *acute accent*, following traditional terminology for Lithuanian and Ancient Greek (Kiparsky 1973 claims the acute is realized as HL in Lithuanian and as LH in Greek); other terms in use are *Stoßton* ('abrupt tone'; also used for Lithuanian), *Schärfung*, *correption*, and (since Schmidt 1986) *accent 1*. I denote the level tones (HH or LL) of (2) by *circumflex accent*, again following traditional terminology (Kiparsky 1973 identifies the circumflex as HH in both Lithuanian and Greek); other terms in use are *Schleifton* ('slurring tone'; also used for Lithuanian), *Trägheitsakzent*, and *accent 2*.

The fundamental insight of the present paper can be told in a couple of sentences. Around the year 1100, Franconian had no tone contrast, but it did have a short–long vowel contrast in open syllables: **/.ma.kən./** 'make' versus **/.slɑɑ.pən./** 'sleep' (where "**.**" stands for a syllable boundary). In sentence-internal focus position, these words were all accented on their first mora, i.e. realized as [mákə̀n, slɑ́ɑ̀pə̀n] in declarative sentences or [mə̀kə́n, slə̀ɑ́pə́n] in interrogative sentences. Subsequently, the

[1] The question whether the early-versus-late-alignment hypothesis can be maintained for the present-day Franconian dialects, i.e. whether it can also explain what happens nowadays in positions *outside* focus or whether an analysis of the data outside focus instead requires tone, has not been answered with certainty yet. Gussenhoven's synchronic analyses in terms of tone are countered by several synchronic analyses in terms of metrical structure (Kehrein this volume, Köhnlein this volume, Hermans 2012, Oostendorp this volume). This debate is not the subject of the present paper.

common process of *open syllable lengthening* lengthened the vowel, ***and lengthened the tone with it***, yielding the present-day contrast, which is [máákən] versus [slɑ́ɑ̀pən] in declarative sentences and [mààkɑ́n] versus [slɑ̀ɑ́pɑ́n] in interrogative sentences. Many subsequent developments have complicated this original correlation, but I will argue in detail that the lexical tone contrast originated in open syllable lengthening rather than in the later processes of schwa drop (as claimed by De Vaan 1999) or analogical lengthening (as claimed by Gussenhoven 2000c).

2. Historical correlations of the present-day tones

The present-day tone contrast occurs on syllables that contain two (or more) sonorant moras, i.e. on long vowels and on diphthongs, where any sequence of a short vowel and a sonorant consonant counts as a diphthong. I denote the acute tone with an acute accent sign above the first sonorant mora: **ʃlɔ́ɔpən** 'sleep', **klɛ́in** 'small', **lɑ́ŋ** 'long-FEM'. The circumflex tone is depicted with a macron above the second sonorant mora: **ʃpeēlən** 'play', **daāk** 'roof', **brɛīt** 'broad', **lɑɑ̄ŋk**.[2]

The rules that relate the current Franconian tones to original West-Germanic vowels have been known for some time (Nörrenberg 1884, Engelmann 1910). I summarize them in (3); they are described in more detail in the following sections.

(3) *The historical correlations of the Franconian acute and circumflex*

a. "Spontaneous acute accent": originally long non-high vowels (***aa**, ***ɛɛ**, ***ɔɔ**, ***ee**, ***oo**) currently have acute accent.

b. "Combinatory acute accent": originally long high vowels (***ii**, ***yy**, ***uu**), original diphthongs (***ɛi**, ***ɔu**), and lengthened originally short vowels have circumflex accent, except under the following region-dependent condition, in which case they have acute accent:

"Rule A": in most of Ripuaric and Moselle Franconian (Welter 1933), these vowels have acute accent if they were originally followed by an voiced consonant followed by schwa.

"Rule A2": certainly in most of Limburgian and other Low Franconian areas (Nörrenberg 1884: 409, Maurmann 1898, Grootaers 1909, Dols 1944), but also perhaps in some parts of Moselle Franconian (Vianden: Engelmann 1910: 390; Trier: Reitz 1986: 6), these vowels have acute accent if they were originally followed by a voiced consonant followed by schwa, but *only if this schwa was deleted*.

[2] The citation forms of the acute words can be pronounced with an early falling pitch contour: [ʃlɔ́ɔ̀pə̀n, klɛ́ì̀n, lɑ́ŋ̀], those of the circumflex words with a late falling pitch contour on the disyllables ([ʃpéélə̀n]) or (as a special result of the interaction between the lexical tone and the declarative intonation; see Gussenhoven 2000b) with a high-mid-high contour on the monosyllables ([dɑ́ɑ̌k, brɛ́ɛ̌t, lɑ́ŋ̌k]); other often-mentioned phonetic differences between the citation forms are that the circumflex is longer than the acute, which is especially audible in the second mora, and that the acute has strong intensity in the first mora. General remarks on the pronunciation of Limburgian consonants are: voiceless plosives are unaspirated, voiced plosives are fully voiced, /t, d, n/ are dental stops, /c, ɟ, ɲ/ are palatalized palatoalveolar stops, /r/ is a uvular trill, tap, fricative or approximant, /l/ is 'light' in all positions, /w/ is a bilabial approximant, and /ɦ/ is a voiced glottal fricative.

In (3), "originally" refers to the state of Low and Central Franconian around the year 1100, when the "Old" stage of these languages turned into the "Middle" stage. Table (3) makes references to at least three sound changes: vowel lengthening (in open syllable), schwa deletion (word-final), and consonant devoicing (also word-final). A strict chronology that works for rule A2 is:

(4) *A simplified strict chronology for Franconian tonogenesis, rule A2*

a. First, long non-high vowels receive the acute accent, and long high vowels and diphthongs receive the circumflex accent. The tone contrast is "just phonetic" (§4).

b. Subsequently, short vowels lengthen in open syllables and some other places, receiving the circumflex accent. This leads to a lexical tone contrast between originally long non-high vowels and the new lengthened vowels (§5).

c. Next, final schwa is deleted, causing a preceding circumflex to change to acute if the intervening consonants are voiced. This leads to a higher functional load of the tone contrast for words that end in a sonorant (§7).

d. Finally, final obstruents are devoiced. This leads to a higher functional load of the tone contrast for words that end in a fricative (§11).

Rule A2 was formulated later than Rule A and is perhaps for that reason sometimes regarded as an exceptional variant, or even a later development, of Rule A (e.g. Tans 1938: 21, 160, 216). This view cannot be correct, because the Rule A2 areas retain a contrast that was lost in the Rule A areas: in rule A2, disyllabic words with a voiced medial consonant have an acute for originally long non-high vowels (**kéezən** 'choose') but a circumflex for lengthened vowels (**lɛɛ̄zən** 'read'); in rule A, both have an acute (**kéezən**, **léɛzən**), thus neutralizing an original etymological distinction. If anything, therefore, the Rule A2 areas represent a more archaic situation, and the situation in the Rule A area is a later neutralizing development (a circumflex changed into an acute if followed by another syllable after intervening voiced consonants, e.g. **lɛɛ̄zən** > **léɛzən**). See §8 for details.

Finally, there are the Rule B areas. In one variant (Bach 1921), the tones are the reverse from those in area A, i.e. wherever area A has an acute, area B has a circumflex, and the reverse. Köhnlein (2011) shows that this is true only for declarative intonation. See §9 for details.

The order in (4) is nearly dictated by the correlations in (3). Nörrenberg (1884:408,410), for instance, argues that logically (4b) must precede (4c) and that very likely (4a) precedes (4c); to make this more precise, the present paper argues on the basis of non-neutralization that (4a) must precede (4b). Yet, some recent publications on the subject propose different orders. For instance, De Vaan (1999) proposes that the lexical contrast arose with schwa drop, i.e. with (4c) rather than with (4b), and Gussenhoven (2000c) proposes that it arose with a very late case of "analogical lengthening", namely during the schwa drop of (4c), rather than with the earlier lengthening in open syllable. Also Schmidt (2002) maintains that schwa drop is the cause of the lexical contrast. In §13 I show that these alternative proposals cannot

account for all the facts, and that the order in (4) must be correct, including the establishment of the lexical contrast in stage (4b).

The remainder of this paper gives a detailed account of the development of the Franconian tone system in terms of the phonology, the phonetics, the speakers, the listeners, and the learners of this language throughout the preceding millennium.

3. The dialects in space and time

This section presents the dialect described in this paper, and its relation with earlier stages of the language and with neighbouring dialects.

3.1. Why East Limburgian, and why Geleen?

The facts presented in §2 require that among the Franconian tone languages we investigate the dialect group that retains the original contrasts best, i.e. the dialects that follow Rule A2. This limits our choice to the Limburgian tone dialects. Within Low Franconian, these tonal dialects occupy an area in the Southeast (in Belgium, the Netherlands and Germany), separated from more Western and Northern varieties of Low Franconian by a thick bundle of isoglosses. Together, the Low Franconian tonal dialects are sometimes called Southern Low Franconian (*Südniederfränkisch*), but in this paper I simply call them Limburgian, which is the name given by its speakers in Belgium and the Netherlands, and by Goossens (1965) even for the German parts.

Fortunately, it is also within Limburgian that we find the most conservative vowel systems, which will allow us to easily project back and forth between the earliest forms and the modern forms. The most conservative group is what Goossens (1965) calls East (mainly Dutch) Limburgian, a group that comprises the cities of Venlo, Roermond, Sittard and Geleen. Within this group, Geleen has the most conservative vowel system, with the fewest mergers and a nearly one-to-one correspondence with the common West Germanic and Old High German vowel systems. As our source of present-day language data, therefore, I take the group of East Limburgian dialects, and I will write the present-day forms in a spelling most suitable for the Geleen variety.

3.2. The dialects in time

Table (5) shows the systems of long and short vowels in four stages of the language. Historical correspondences can be read off horizontally.

(5) *Four consecutive vowel systems*

Proto West Germanic	Old (Eastern) Low Franconian	Early Middle Limburgian	New East Limburgian
**î* **iu* **û*	*î* *ûi* *û*	***ii** ***yy** ***uu**	**iī** **yȳ** **uū**
**ê₂*, **eo* **ô*	*ie* *uo*	***ee** ***øø** ***oo**	**ée** **øø** **óo**
**ai* **au*	*ei* *ou*	***ɛi** ***œy** ***ɔu**	**ɛī** **œī** **ɔū**
	ê *ô*	***ɛɛ** ***œœ** ***ɔɔ**	**éa** **øa** **óa**
**â*	*â*	***ææ** ***ɑɑ**	**éa** **ɔ́ɔ**
**i* **u*	*i* *u*	***e** ***ø** ***o**	**e**,**eē** **ø**,**øø̄** **o**,**oō**
**e* **o*	*e* *o*	***æ**,**ɛ** ***ɔ**	**ɛ**,**ɛɛ̄** **ɔ**,**ɔɔ̄**
**a*	*a*	***a**	**æ** **ɑ**,**aā**

In table (5), "Proto West Germanic" is the traditionally reconstructed language that can be regarded as the common ancestor of all current West Germanic languages, including English, German, and Dutch.

"Old (Eastern) Low Franconian" (OLF) in table (5) is the language of the Wachtendonk psalms, usually thought to have been written around the year 900. The OLF vowel system is virtually identical to that of most known Old High German dialects. This language is generally considered a predecessor of current Limburgian (Cosijn 1872–1873; Van Helten 1896, 1902; Cowan 1959, 1969; Sanders 1968–1969); an introduction can be found in Robinson (1992), but the present paper uses the publication by Kyes (1969) and follows the historical connections proposed or assumed in the etymological dictionaries (Kluge 1915, Van Wijk 1912/1929).

"Early Middle Limburgian" (Early MLb) is the language that we need to reconstruct here as the first stage that is relevant for understanding the rise of the Franconian tone system. It is characterized by a general change of vowels in unstressed vowels to schwa, and must have been spoken around the year 1100. The Early MLb vowel system is reconstructed in §3.3, its accent system in §4.

"New East Limburgian" (NLb) is the language currently spoken by about 300,000 people in the central part of the Dutch province of Limburg, as exemplified by the Geleen dialect. Table (5) shows the situation after step (4b), i.e. after open syllable lengthening; the effects of step (4c), i.e. the drop of final schwa, have not been included in the table.

3.3. Reconstruction of the Early Middle Limburgian vowel system

This section discusses how the various Early MLb vowels can be reconstructed. The reconstruction requires projections from earlier as well as from later varieties. The projection from OLF to MLb requires an interpretation of the written OLF source.

Many historical correspondences have been suggested in table (5) by writing the vowels of various stages at the same height. Thus, PWG long **û* is still long **uū** in NLb (where not changed by schwa drop), and **ô* is still **óo**. PWG short **u* developed into NLb **ʊ** or **oō**, and short **o* developed into **ɔ** or **ɔɔ̄**; in these examples, the long circumflex vowels originated by open-syllable lengthening. Through an ambiguous vertical alignment, the table also shows the split of PWG **au*, which developed into **ɔū** and **óa** in NLb. The thing that happened around "***æ**,**ɛ**" in the table is a complicated twist: OLF *e* lowered to MLb ***æ**, whereas some instances of OLF *a* raised all the way to MLb ***ɛ**. The following subsections discuss the OLF and Early MLb vowels in detail.

The long high vowels. The OLF documents contain no length marks for vowels, but many of the vowels that were written <i> and <u> continue the Common Germanic long high vowels and must have been long in OLF since they are still long in the later dialects. Therefore, the traditional transliteration, which we will follow, is *î* and *û*. These vowels occur in words like *wîn* 'wine', *rîki* 'empire', *hûs* 'house'. There is no reason to assume that these vowels were pronounced very differently from [iː, uː] in any of the four periods under consideration, so we can write ***wiin**, ***riikə**, ***huus** for Early MLb (but see §4.6 for some important structural and phonetic detail). The vowel *ûi* (<ui>) continues the West Germanic diphthong /iu/ and is found in *fûir* 'fire'. This

spelling, with *i* as a syllable-internal marker for fronting, shows that it was already pronounced as [yː] in the OLF period, so we can reconstruct MLb *__vyyr__.

Another source of MLb *__yy__ is *i-umlaut* of OLF *û*. The transition from the Old to the Middle stages of the West Germanic languages is traditionally defined by the loss of vowel distinctions in unstressed syllables. For Eastern Low Franconian, this meant that OLF *gevan* 'give' would become MLb *__ɣævən__, and OLF *namo* 'name' would become MLb *__namə__. However, if the vowel in the final syllable was *i* (or if it contained *j*), this vowel tended to leave a trace: the vowel of the first syllable was fronted (at least if it was back; if it already was front, it was raised). Thus, a hypothetical OLF *ûwila* 'owl' corresponds to MLb *__yylə__. The easiest way to think of the order of events is that *i*-umlaut preceded the loss of the vowel quality, i.e., that the sequence of forms was [uːwɪla] → [yːwɪla] → [yːwələ].

The closing diphthongs. Table (5) shows a phonemic split as a result of the monophthongization of West Germanic **ai* before /h, w, r/ and in final position, and of West Germanic **au* before /h/, before the dentals /d, t, θ, ð, s, z, l, r/, and in final position. Thus, OLF has *stein* 'stone', *leidon*/*leiden* 'lead', *loupon* 'walk', but *sêo* 'sea' (< **saiw*), *sêla* 'soul' (< **saiwala*), *-hôr-* 'hear', *dôt*[3] 'death'. This innovation equally affected OLF and most Old High German dialects, including the predecessors of Ripuarian and Moselle Franconian; it did not affect Low Saxon and Western Low Franconian (the predecessor of Dutch), both of which monophthongized nearly all of these diphthongs to /eː/ and /oː/. Because of the later reflexes, we can assume that *ei* and *ou* stood for /ɛi/ and /ɔu/. Mechanically, we expect the following correspondences between OLF and MLb: *stein* > *__stɛin__, *leidon* > *__lɛidən__, *loupon* > *__lɔupən__. From the spelling of the *short* vowels /ɛ/ and /ɔ/ as <e> and <o> (see below), and from the spelling of the diphthongs *ei* and *ou* as <ei> and <ou>, we can hypothesize that the scribe used the symbols <e> and <o> to denote lower mid vowels, so that we can reconstruct the monophthongs *ê* (<e>) and *ô* (<o>) as /ɛː/ and /ɔː/. This would be consistent with the likelihood that the relationships between *ê* and *ei* and between *ô* and *ou* had still been allophonic in an earlier OLF period: the underlying glides in /ɛi/ and /ɔu/ were pronounced only in certain positions, deleted with compensatory lengthening in others (e.g. *[stɛjn] vs. *[sɛːwala]). Also, the later diphthongal reflexes of *ê* and *ô* (i.e. __éa__ and __óa__) point to a MLb pronunciation as lower mid vowels, as does the fact that MLb *must* have had and OLF *may* have had a separate series of higher mid vowels (see next subsection). We thus project from OLF to MLb as follows: *sêo* > *__zɛɛə__ (also with voicing of initial fricatives), *sêla* > *__zɛɛlə__, *dôt* > *__dɔɔt__. It is possible that Early MLb already slightly raised the first part: *__zɐɛlə__, *__ɣrɐɔt__.

The source of MLb *__œy__ is *i*-umlaut of OLF *ou*. We can assume that OLF **rouvâri* 'robber' (literally 'reaver') was *__rœyvæ(æ)rə__ in Early MLb. The source of MLb *__œœ__ is *i*-umlaut of OLF *ô*. Thus, *hôren* became *__ɦœœrən__ in Early MLb.

At some time during MLb, the monophthongs *__ɛɛ__, *__œœ__, and *__ɔɔ__ must have "broken" into *__eɛ__, *__øœ__, and *__oɔ__. After this, the two parts of the diphthongs dissimilated. For the back vowel, the stage __oɑ__, with a close-mid first part, is still present in Geleen and the villages to the West of it (__ɣróɑt__ 'large'); this vowel has a

[3] I write voiceless final consonants that alternate with voiced consonants (*dôdi*) by underlining them, both in OLF and in MLb.

corresponding morphological umlaut (**pøœt**, plural of **pɔɑt** 'paw'). In these areas, dissimilation is complete for the original front vowel (**zíæl** 'soul') and for the primary umlaut (**ɦýœrə** 'hear'). In most of the remaining parts of East Limburgian (e.g. Maasbracht, and the more peripheral dialect of Venlo to the North), these vowels have simply turned into **iə**, **yə**, **uə**: **zíəl**, **ɦýərə**, **ɣrúət**, **púət**, **pýət**; these sounds are also heard in some of the neighbouring areas such as Weert (West) Limburgian to the West and Kerkrade Ripuarian to the South-East. I will transcribe these vowels as **ea**, **øa**, and **oa** when referring to the Geleen region, and as **iə**, **yə**, and **uə** when referring to the other regions. In the East Limburgian city of Roermond, however, these vowels merge with the close mid vowels and are pronounced **ée**, **øø**, and **óo**. Since Roermond is important in our story, I will generally cite Roermond forms quite often, but mention one of the more conservative dialects when such forms contain any of these three vowels.

The long higher mid vowels. The OLF vowel written <ie> continues West-Germanic **eo* (i.e. **iu* with *a*-umlaut) when it occurs in *liegon* 'tell a lie', and **$\hat{e}_2$* when it occurs in *hiera* 'here' and the preterites *riep* 'called out', *farliet* 'left'. The vowel written <uo> continues West-Germanic **ô* and occurs in *fuot* 'foot', *ruopen* 'call out', and with final devoicing in *guot* 'good', *muot* 'mood', *fluot* 'flood' (dative *fluode*). These vowels may have been diphthongs, but since later dialects show evidence of the original /eː/ and /oː/, there are three possibilities, all problematic: the psalms were from different dialects than the later tone dialects, or /ie/ and /uo/ later changed back to /eː/ and /oː/, or written *ie* and *uo* actually represent contemporary /eː/ and /oː/. For Early MLb, I simply write ***leeɣən** 'tell a lie' and ***voot** 'foot'.

The source of MLb ***øø** is *i*-umlaut of OLF ***oo**. We can assume that OLF **gruoni* 'green' was ***ɣrøønə** in Early MLb.

The long low vowel. The OLF long vowel written <a> continues West-Germanic *â* (Germanic **$\hat{e}_1$*). The traditional notation is *â*: *slâpan* 'sleep', *jâr* 'year', *gâvon* 'they gave'. Considering its later reflexes, it was probably pronounced as the back vowel [ɑː] in the Franconian areas. For early MLb, I write **slɑɑpən**.

All of the later dialects can have *i*-umlaut of this vowel, so we can safely assume *kâsi* > ***kææzɪ** > ***kææzə**. None of the later dialects makes a difference between this ***ææ** from umlauted PWG **â* and ***ɛɛ** from monophthongized PWG **ai* (which was discussed above). Thus, we have Geleen **kíæs**, Maasbracht **kíəs**, and Roermond **kées**. It is likely, therefore, that the two vowels must have fallen together some time during MLb, perhaps early. Since the MHG handbooks distinguish the two sounds, however, writing them as <æ> and <ê>, I will distinguish ***ææ** and ***ɛɛ** for Early MLb.

The short high vowels. The OLF short vowels written <i> and <u> continue West Germanic **i* and **u*: *ist* 'is', *tunga* 'tongue', *ruggi* 'back (ridge)'. If we take into consideration their reflexes in any modern West Germanic language, the vowels written *i* and *u* in older West Germanic scribal traditions must usually be regarded as not entirely high, i.e. as **ɪ**, **ʊ**, and (with umlaut) **ʏ**. This must have been especially true of OLF, since these vowels would later become **ee**, **oo** and **øø** after Open Syllable Lengthening, as the table indicates. Following most work on dialects in these regions, I write the vowels as **e**, **o** and **ø** (MLb ***es**, ***toŋgə**, ***røggə**).

The short low vowel. The OLF short vowel written <a> continues West Germanic **a* without *i*-umlaut: *sal* 'shall', *craft* 'strength'.

The short front mid vowel. The OLF orthographic symbol <e>, when referring to a short vowel, continues two different West Germanic sources: **a* with *i*-umlaut, and **e*. Since the later East Limburgian dialects (although not Venlo) distinguish between **æ** as a reflex of PWG **e* without *i*-umlaut and **ɛ** as a reflex of PWG **e* and **a* with *i*-umlaut, the distinction must have been available in OLF if this language is to be the predecessor of those dialects. The distinction was not always available in the form of an *i* or *j* in the next syllable. Thus, the attested OLF forms *tellon* 'tell' and *settan* 'set' have an original **a* with *i*-umlaut (PWG **taljan*, **satjan*) without having an *i* in the second syllable, whereas present-day East Limburgian has **tɛlən** 'count' and **zɛtən** with overt reflexes of *i*-umlaut. If all OLF *e* had been pronounced in the same way, the modern reflexes of *tellon* and *settan* would have been **tælən** and **zætən**, with the same vowel as in **væxtən** 'fight' from *fehtan*. It seems likely, therefore, that the OLF symbol *e* stood for two different vowels, a higher one in the case of *i*-umlaut of **a* or **e*, and a lower one in the case of **e* without *i*-umlaut. This situation is the same as with the scribal traditions of most Old High German dialects, in which the contrast was not written but for which a contrast is nowadays generally assumed on the basis of some later dialects (Middle High German grammars write *ę* for **ɛ** and *ë* for **æ**).[4] Thus, we can assume MLb **tɛllən**, **zɛttən**, **bɛddə** 'bed', **wɛggə** 'loaf', **bɛst** 'best' (< **batist*).

The short back mid vowel. The OLF short vowel written <o> continues West Germanic **o* (*a*-umlaut of Germanic *u*): *folkon* 'people', *thorna* 'thorn'. It was probably [ɔ]. Since OLF *o* arises from *a*-umlaut, we do not expect many opportunities for *i*-umlaut; the attested form *ovir* 'over' could be the basis of MLb ***œvər** and NLb **œɵ̄vər**, a local variant of **øø̄vər** (< ***øvər** < **uvir*).

3.4. Reconstruction of the Early Middle Limburgian consonant system

Three changes in consonants are relevant to the conditioning of tone. First, obstruents devoiced in final position. This *First Final Devoicing* is already apparent in the attested OLF, and marked in this paper by underlining (e.g. *go<u>t</u>* 'God-NOMSG' versus *godis* 'God-GENSG'). Second, non-geminate fricatives became voiced in intervocalic position (*kâsi* > **kɛɛzə**). Third, vowels in final unstressed syllables were neutralized to schwa. Three changes irrelevant to tonogenesis were: the voicing of most fricatives (*fûir* > **vyyr**, *sêla* > **zɛɛlə**), followed by the change of non-intervocalic **ð** (from *th*) to **d**; and *i*-umlaut, which had created the new vowels **øø**, **œy**, **œœ**, and **ʏ**, merged the vowels of *ûwila* and *kâsi* with those of *fûir* and *sêla*, and caused the split between **æ** and **ɛ**. Eastern Low Franconian shares all of these six changes with most High German dialects, including Ripuarian and Moselle Franconian. Examples of MLb words in this stage are **wiin**, **vyyr**, **yylə**, **ɦuu<u>s</u>**, **dee<u>f</u>**, **ɣrøønə**, **book**, **klɛinə**, **brɛit**, **rœyvæərə**, **lɔupən**, **zɛɛlə**, **kæəzə**, **ɦœœrən**, **dɔɔ<u>t</u>**, **slɑɑpən**. The Central Franconian dialects differed from their Low Franconian neighbours only in their voiceless obstruents (**boox**, **lɔufən**, **slɑɑfən**).

[4] Van Wijk (1912/1929) uses the notations *e* and *ë* for OLF (e.g. *tellon*, *settan*, *gëvan*, *fëhtan*).

3.5. The dialects in space

This paper mainly refer to those dialects that have no tone-related segmental changes, namely Geleen and the towns and villages to the West of it, and the villages between Roermond and Sittard. The Roermond dialect will be referred to because there is reliable information about its tones in a dictionary (Kats 1985), but the opening diphthongs /eə, øə, oə/ have monophthongized and merged with /ee, øø, oo/. The (older) Sittard dialect will be referred to, but here /ée, ǿø, óo/ have diphthongized to /ɛ́i, œ́i, ɔ́u/ (Dols 1944), causing a merger with schwa-drop-caused /ɛ́i, œ́i, ɔ́u/, while /eə, øə, oə/ have turned into /ée, ǿø, óo/ as in Roermond. The city of Venlo, in the Northern periphery of East Limburgian, is very different and will only very occasionally be mentioned. The Central and West Limburgian dialects of Maastricht, Weert, and the whole Belgian part of Limburg are too different to discuss. In the German part of East Limburgian (e.g. Dülken), the dialects are different as well.

For the present-day forms transcribed in this paper without location the segmental part will be based on the Geleen dialect, but the tones will be based only on absolutely reliable sources, namely on the Roermond data by Kats (1985), the Sittard data by Dols (1944), and occasional Maasbracht data by Ben Hermans (p.c.). If these reliable sources reveal regional variation in tone, this will be noted with markings such as "Rm." and "Sitt." where needed. If all reliable sources mention the same tone, and this tone does not contradict the present author's less reliable knowledge of Geleen, I will generalize this tone on the otherwise Geleen-based transcription without further comment. This transcription strategy is not 100 percent error-proof, but is currently optimal for combining maximally contrastive segmental transcriptions with maximally reliable tone transcriptions.

4. Reconstruction of the Early Middle Limburgian accent

The first pattern to explain is why the Early MLb long non-high vowels (*__aa__, *__ɛɛ__, *__ɔɔ__, *__ee__, *__øø__, *__oo__) later on came to behave differently from the long high vowels (*__ii__, *__yy__, *__uu__), the diphthongs (*__ɛi__, *__ɔu__), and the short vowels (*__ɑ__, *__æ__, *__ɛ__, *__ɔ__, *__e__, *__ø__, *__o__). The basic proposal will be that the six long non-high vowels were bimoraic, and that all the others were monomoraic. The following sections discuss in detail the reconstructed synchronic representations of all these sounds.

For each sound we have to distinguish three kinds of representations. First, there is the underlying form. The absence or presence (in the underlying form) of accent placement or tones will tell us what kinds of lexical contrasts the language has. The second relevant representation is the phonological surface structure, which is a prosodic structure containing metrical material (moras, syllables, feet) and tones (H, L). This is the structure that the speaker's grammar has to create out of the underlying form, and it is also the structure that the listener has to create out of overt auditory-phonetic information. The third relevant representation, then, is this overt auditory-phonetic pronunciation. Since the case of language evolution presented in this paper is proposed to reside for a large part in acquisition, and the acquisition processes proposed in this paper depend on the idea that the child has to invent phonological structures on the basis of overt primary language data, the phonetic representation under scrutiny here must be of an auditory rather than of an articulatory nature.

4.1. Underlying forms of non-high vowels in open syllables: a moraicity contrast

As is evident from table (5), Early MLb non-high vowels were contrastive for length. If we disregard some possible slight detail in vowel quality, we can see several short–long pairs, for instance *e ~ *ee, *ɔ ~ *ɔɔ, and *a ~ *ɑɑ. If length was contrastive, it had to be indicated in the underlying form. I follow here the usual solution of writing short vowels as monomoraic, long vowels as bimoraic. The underlying timing structures for short and long non-high vowels, then, are those presented in (6).

(6) *Early MLb underlying forms for short non-high vowels*

μ μ	μ μ	μ μ	μ μ	μ μ	μ μ μ	μ μ
| |	| |	| |	| |	| |	| | |	| |
n a m ə	**w æ k ə**	**b ɛ k ə**	**ɔ p ə n**	**ɦ e m ə l**	**k ø k ə n ə**	**s t o v ə**
‘name’	‘week’	‘brook’	‘open’	‘heaven’	‘kitchen’	‘stove’

(7) *Early MLb underlying forms for long non-high vowels*

μ μ μ	μ μ μ	μ μ μ	μ μ μ	μ μ μ
| | |	| | | |	| | |	| | |	| | |
s l ɑ ɑ p + ə n	**k æ æ z ə**	**z ɛ ɛ l ə**	**ɦ œ œ r + ə n**	**b ɔ ɔ n ə**
‘sleep–INF’	‘cheese’	‘soul’	‘hear–INF’	‘bean’

μ μ μ	μ μ μ	μ μ μ
| | |	| | |	| | |
s p e e ɣ ə l	**ɣ r ø ø n ə**	**k o o k ə**
‘mirror’	‘green’	‘cake’

I could have written each vowel in (7) with a single vowel symbol connected to two moras. This is often done in the notation of metrical structures. I chose instead to use the slightly redundant two-symbol notation here for the sake of readability, i.e. in order to have segmental representations that are similar to the linearized abbreviations of these underlying forms, which are **|slɑɑpən|**, **|kææzə|**, **|zɛɛlə|**, and so on.

The underlying forms in (7) contain no marks for stress. This is because MLb did not have lexical stress. All stress was assigned by the grammar, as shown in the next section.

4.2. Surface structures for non-high vowels in open syllables: moraic stress

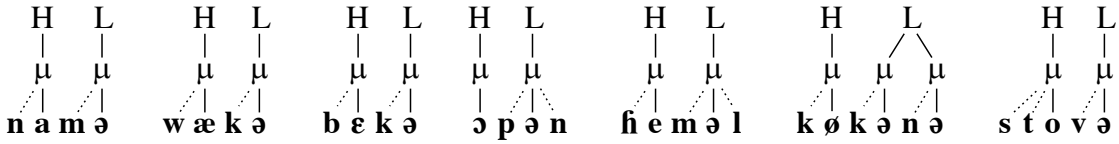
Stress in OLF quite likely fell on the first syllable of a word, disregarding a limited number of unstressed prefixes such as *ge-* and *far-*. As we see in table (5), the first syllable of a word could contain any of ten long vowels and diphthongs and any of five short vowels. The last syllable of a word could contain only a short vowel, mostly *-a*, *-e*, *-i* and *-o*. Because these final vowels historically derive from Proto-Indo-European long vowels, the reduced inventory of final vowels must be a Germanic innovation, and scholars agree that the cause of this reduction is the Germanic shift from a lexical accent on the first or second syllable to a grammatical accent on the first syllable. This accent shift caused initial syllables to become increasingly longer, and final syllables to become increasingly shorter.

In MLb, the unstressed vowel inventory had shrunk even more. There may not have been any other final unstressed vowels than *ə, and full unstressed vowels only in derivational suffixes such as **-liik** ‘-ly’ and **-əkiin** ‘(diminutive)’. We can assume, then, that Early MLb, like OLF, had a grammatical accent that fell within the first syllable

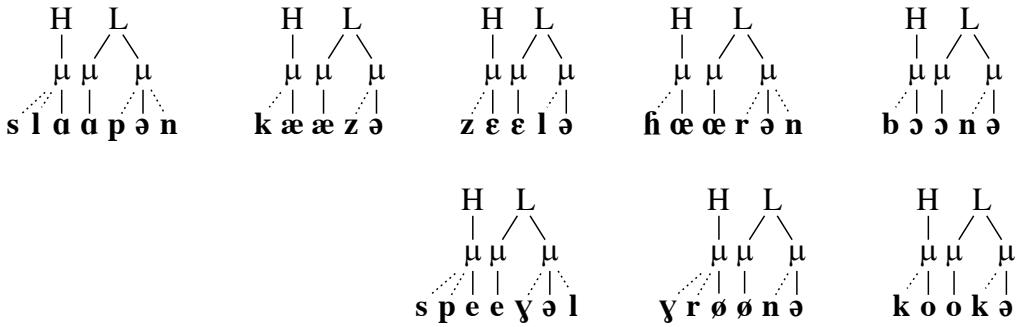
(again disregarding some unstressed prefixes), even in words that derive from Latin, such as *__køkənə__ 'kitchen' < *coquîna*.

I propose that the MLb accent was not carried by the initial syllable but by the initial mora (for a possible explanation, see the following section). The difference is illustrated by considering the alignment of sentence-internal focus intonation contours, which I assume to have been H*L for declarative focus (where H* is the high focus tone and L the transition to the low boundary tone) and L*H for interrogative focus (where L* is the low focus tone and H the transition to a high or high-low boundary tone). In present-day German, Dutch and English we can assume that the tone bearer is the syllable, and that it can host two tones, so that the stressed syllable (i.e. the accented syllable of the stressed word) carries the whole HL (or LH) movement. In MLb, I assume that the tone bearer was the mora, and that it could host only a single tone. The H* (or L*) was then connected only to the first mora, while the boundary L (or H) was connected to the following moras. This leads to the surface structures in (8) and (9).

(8) *Early MLb surface structures for short non-high vowels, aligned to H*L*



(9) *Early MLb surface structures for long non-high vowel, aligned to H*L*



The recipe for creating these structures can thus be formulated as follows:

(10) *Recipe for creating tone structures inside MLb sentences*

a. the declarative focus intonation is H*L, the interrogative focus intonation is L*H;

b. the first mora of a word carries the accent;

c. moras can host one tone (H or L);

d. tones are aligned from left to right, starting with the accented mora.

Present-day reflexes. For the long non-high vowels in (9), the structures in (9) have survived to the present day, at least for those words that did not lose the second syllable. These divide into four types, all of which can be seen in (11) and (12): words that used to consist of three or more syllables, disyllabic words whose second syllable used to end in -**ər**, -**əl**, or -**əm**, or -**ən** (e.g. the infinitive and first and third persons plural present and past of most verbs, the past participle of strong verbs, the plural of most nouns

whose singular ended in -**ə**, the first constituent of many compounds, and the masculine singular of adjectives, which derives from the MLb accusative case), some disyllabic words whose second syllable ended in -**ə** (namely the plural and feminine singular of adjectives ending in a MLb plosive or voiceless consonant cluster), and a couple of monosyllabic words that ended in a vowel. For such words, the present-day reflexes of long non-high vowels in open syllables usually follow the structures in (9), i.e. they carry the acute accent.

In non-voiced cases, i.e. where the following consonant was and is voiceless, or where there was no following consonant, the predictability of the present-day accent from the structures in (9) is nearly perfect, as can be seen in (11). The only exceptions are formed by two specific morphological categories, namely some comparative forms of adjectives and the first and third person plural of the past tense of a number of strong verbs.

(11) *Modern forms with originally long non-high vowels in open syllables, non-voiced case: nearly always acute*

Early MLb ***ɑɑ** > Gel. Rm. Sitt. **ɔ́ɔ**: **ʃlɔ́ɔpən**[5] 'sleep-INF/1,3PL' < ***slɑɑpən**, **lɔ́ɔtən** 'let-INF/1,3PL', **ɣəʃlɔ́ɔpən** 'slept (p.p.)' < ***ɣəʃlɑɑpən**, **ɣɛlɔ́ɔtən** 'let (p.p.)', **ʃtrɔ́ɔtən** 'streets', **ɦɔ́ɔkən** 'hooks', **ʃɔ́ɔpəvlɛīʃ** 'mutton', **ʃtrɔ́ɔtəndræk** 'street refuse', **mɔ́ɔtən** f.pl. 'measures'

Early MLb *ææ > Gel. **éa** (Rm. Sitt. **ée**): **ɣəspréakələk** 'talkative' (cf. NHG *Gespräch* 'talk') < ***ɣəspræækəliik**, Sitt. **éekər** 'bucket'

Early MLb *ɛɛ > Gel. **éa** (Rm. Sitt. **ée**): **zéa** 'sea'

Early MLb *œœ > Gel. **ǿa** (Rm. Sitt. **ǿø**): (no cases known)

Early MLb *ɔɔ > Gel. **óa** (Rm. Sitt. **óo**): **ɣróatə** 'large-PL/FEMSG' < ***ɣrɔɔtə**, **ɣróatən**[6] 'large-MASCSG' < ***ɣrɔɔtən** '(acc.)', **ʃtóatən** 'thrust-INF/1,3PL', **ɣəʃtóatən** 'thrust (p.p.)', Rm. Sitt. **ɣroōtər** 'larger'

Early MLb *ee > Gel. Rm. **ée** (Sitt. **éi**): **déepə** 'deep-PL/FEMSG', **déepən** 'deep-MASCSG', Rm. **deēpər** 'deeper', **ɣənéetən** 'enjoy', **ʃéetən** 'shoot', Sitt. **ʃléitən** 'close', **zéekə** 'infected-PL/FEMSG', **zéekən** 'infected-MASCSG', Rm. **ʃléepən /** Sitt. **ʃlɛīpən**[7] 'slept-1,3PL', Rm. **léetən /** Sitt. **lɛītən** 'let-PAST1,3PL', Rm. **réepən /** Sitt. **rɛīpən** 'called-1,3PL', **knée** 'knee' (Sitt. **knéi**; Rm. **kníj**)

---

[5] The final -**n** of the infinitive, past participle, and most other forms written here with -**ən** shows up only before words that start with a vowel (i.e. a glottal stop) or **ɦ**, thereby usually replacing the glottal stop or **ɦ**. Because the citation form drops **n**, few authors on Limburgian include **n** in their transcriptions. I do write -**ən** in this paper in order to distinguish these NLb forms unambiguously from MLb forms that ended in a -**ə** that was later lost. Thus, MLb had ***strɑɑtə** ~ ***strɑɑtən** 'street ~ streets', NLb has **ʃtrɔ́ɔt** ~ **ʃtrɔ́ɔtə(n)**, the last of which is written here as **ʃtrɔ́ɔtən**. This **n** is similar to 'linking *r*' in English in that it is also used in some non-etymological cases such as after the enclitic -**zə** 'you-SG'. However, it cannot be used 'intrusively' after adjectival plurals and feminine singulars such as **ɣróatə** and **déepə**, i.e. those words that exceptionally retained MLb final -**ə**.

[6] The final -**n** of the masculine singular of adjectives shows up before words that start with a vowel, **ɦ**, **d**, **t**, or **n**; the glottal stop or **ɦ** need not be deleted. In other positions, the -**n** is deleted (in Venlo, it turns into -**m** before **b**).

[7] According to Dols (1944), the diphthong in these three Sittard past tenses betrays an earlier acute.

Early MLb ***øø** > Gel. Rm. **ǿø** (Sitt. **œi**): **zǿøkən** ‘seek’, **zǿøtə** ‘sweet-PL/FEMSG’, **zǿøtən** ‘sweet-MASCSG’

Early MLb *oo > Gel. Rm. **óo** (Sitt. **ɔ́u**): **róopən** ‘call-INF/1,3PL’, **ɣəróopən** ‘called (p.p.)’, **kóo** ‘cow’ (Sitt. **kɔ́u**; Rm. **ku**)

In voiced cases, the great majority of present-day polysyllables has an acute as well, at least in Sittard, as can be seen in (12). A non-negligible minority, though, is seen to have an unexpected circumflex, especially in Roermond (as observed by Dols 1944). An explanation for these exceptions was given both by Maurmann (1898) and by Dols and is discussed in §7.3.

(12) *Modern forms with originally long non-high vowels in open syllables, voiced case: usually acute*

Early MLb *ɑɑ > Gel. Rm. Sitt. **ɔ́ɔ**: **jɔ́ɔmər** ‘a pity’, **ʃwɔ́ɔɣər** ‘brother-in-law’, Rm. **vrɔɔ̄ɣən** / Sitt. **vrɔ́ɔɣən** ‘ask-INF/1,3PL’, Rm. **rɔɔ̄jən** / Sitt. **rɔ́ɔən** ‘guess’ < *rɑɑðən, Rm. **brɔɔ̄jən** / Sitt. **brɔ́ɔən** ‘fry’ < *brɑɑðən, **blɔɔ̄zən** ‘blow’, **ɣəblɔɔ̄zən** ‘blown’

Early MLb *ææ > Gel. **éa** (Rm. Sitt. **ée**): **jéaməran** ‘lament’, **kréamər** ‘salesman’, **léaɣən** ‘low-MASCSG’, Rm. Sitt. **leēɣər** ‘lower’, **ɣəbéarən** ‘strive’, Rm. **keēzən** m.pl. ‘cheeses’, **ɦéareŋ** ‘herring’ < *ɦæærenk

Early MLb *ɛɛ > Gel. **éa** (Rm. Sitt. **ée**): **kéarən** ‘turn-INF/1,3PL’ < *kɛɛrən < OLF *kêran*

Early MLb *œœ > Gel. **ǿa** (Rm. Sitt. **ǿø**): **ɦǿarən** ‘hear’

Early MLb *ɔɔ > Gel. **óa** (Rm. Sitt. **óo**): **bóanən** f.pl. ‘beans’

Early MLb *ee > Gel. Rm. **ée** (Sitt. **éi**): **kéezən** ‘choose’, **vərléezən** ‘lose’ (Rm. **vərléerən**), **déenən** ‘serve’, **léeɣən** ‘tell a lie’, **vléeɣən** ‘fly’, **bədréeɣən** ‘cheat’, Rm. **ʃpeēɣəl** / Sitt. **ʃpéiɣəl** m. ‘mirror’

Early MLb *øø > Gel. Rm. **ǿø** (Sitt. **œi**): **vǿørən** ‘lead’, **rǿørən** ‘stir’, **vǿølən** ‘feel’, **ʃpǿølən** ‘wash the dishes’, **bədrǿøvən** ‘disappoint’, **vǿøɣən** ‘cement, add’

Early MLb *oo > Gel. Rm. **óo** (Sitt. **ɔ́u**): Rm. **ɦóovən** / Mbr. **ɦoōvən** / Sitt. **ɦɔ́uvən**~**ɦɔūvən** ‘need’, Rm. **bloōmən** / Sitt. **blɔūmən** ‘flowers’

4.3. Pronunciation of non-high vowels in open syllables

Duration of vowels in OLF: short moras. The predecessor of MLb, Old Low Franconian, still had some full (although short) vowels in unstressed syllables, from which we can conclude that the Germanic shift to initial stress had not run to completion yet. In order to accommodate a final vowel contrast, the final syllable in OLF must have taken up an appreciable part of the time needed to pronounce the whole word. In such a situation, where final syllables are relatively long (as compared with later stages of the language), duration is probably not yet the main auditory cue to phonological stress. As a result, duration was a cue to phonological vowel length only: the short vowels may still have been pronounced really short, say 100 ms, and the long vowels twice that, i.e. 200 ms. Thus, we transcribe [namɔ] but [slɑːpan].

Duration of short vowels in MLb open syllables: long moras. Vowel durations changed in Early Middle Limburgian. The only allowed final vowel was -**ə**, and this must have been allowed to be pronounced with a very short duration, so that more of the duration could shift to the initial (i.e. stressed) syllable. In such a situation, duration can become an important auditory cue to stress. We can assume that language users match their own and others' comprehension and production, so speakers will avoid producing stressed vowels with a duration of only 100 milliseconds.

The avoidance of very short rhymes must have affected most those syllables whose rhymes consisted of only a short vowel. The idea is that all other kinds of stressed syllables had something to contribute to length: long vowels may have been 200 ms long, and in closed syllables with a short vowel the coda consonant could contribute to the duration of the syllable, making it perhaps 160 ms. Open syllables with short vowels were now problematic if they lasted only 100 ms, which was too short for a listener to perceive them as stressed. Speakers will thus have lengthened specifically these vowels, perhaps to, say, 130 ms. The vowels in (8) were thus pronounced as half-long, e.g. *__namə__ was pronounced [naˑmə].

That this lengthening occurred is uncontroversial, because all these vowels would ultimately end up as long after a process of Open Syllable Lengthening (OSL), in Limburgian as well as in many other West Germanic languages, among which present-day standard English ([neim]), German ([naːmə]) and Dutch ([naːm]). This process is generally thought to have taken several centuries to complete, and I propose that its consequences made themselves heard already in Early MLb.

MLb long vowels: incomplete push chain effect and short moras. The lengthening of the short non-high vowels may have caused the long non-high vowels to lengthen, but not to double the duration of a short vowel, i.e. not to 260 ms. We know this because in the West Germanic languages where OSL took place, the two kinds of vowels ended up being equally long (e.g. the lengthened vowel in Dutch *maken* is nowadays the same as the originally long vowel in Dutch *slapen*). Perhaps they just lengthened to 220 ms in MLb. I still simply write [slɑːpən]. If these 220 ms count as two moras, these moras must each have been shorter (namely 110 ms) than the lengthened mora of the short vowel (130 ms).

The alignment of the tone movement. In order to get the first "phonetic" tone distribution started, we need one assumption about the phonetic implementation of the surface structures in (8) and (9). The proposed phonetic implementation rule is simply: a high tone is audible on every voiced segment that is linked to a H in the surface structure, and a low tone is audible on every voiced segment linked to an L. The auditory forms of the words in (8) and (9) are therefore [ńáˑm̀ə], [ẃǽˑkə], [b́éˑkə], [ɔˑpə̀n], [ń́éˑm̀ə̀l], [kǿˑkə̀nə], [stóˑv̀ə], [sĺûːpə̀n], [kæːz̀ə], [źɛːl̀ə], [ń́œːr̀ə̀n], [b́ôːǹə], [spêːɣ̀ə̀l], [ɣ́r̂øːǹə], [kôːkə].

The monomoraic circumflex. The short vowels in open syllable, like [áˑ], carry relatively long high tones: if counted from the mora head, they last 130 ms, as opposed to the high tones in words with long vowels, which last only 110 ms (i.e. one half of 220 ms). We could call this opposition *monomoraic circumflex* versus *bimoraic acute*, and perhaps write the above words in a mixed phonological–phonetic orthography (phonological length; phonetic accent) as *__nāmə__, *__wǽkə__, *__bɛ̄kə__, *__ɔ̄pən__, *__ɦēməl__, *__kǿkənə__, *__stōvə__, *__slɑ́ɑpən__, *__kǽæzə__, *__zɛ́ɛlə__, *__ɦœ́ərən__, *__bɔ́ɔnə__, *__spéeɣəl__, *__ɣrǿønə__, and *__kóokə__.

Foreshadowings. In the auditory forms we already see two factors that contribute to the later tone contrast and its conditioning. First, the enhanced duration of the monomoraic high tones will later lead to their reinterpretation as linked to two moras rather than one (§5.1). Second, tone is audible on voiced consonants, not on unvoiced consonants. This difference will later lead (twice) to the cross-generational preservation of an association line to L for voiced consonants but not for unvoiced consonants (§5.5, §7.2, §7.3).

4.4. Short and long non-high vowels in closed syllables

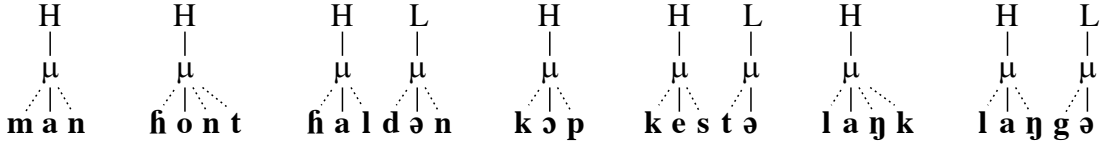
If a short V was monomoraic and a long VV was bimoraic, how many moras went into a VC rhyme? In general, languages differ as to whether they count coda consonants as moraic. For MLb, I propose that codas in monosyllabic words were not moraic. The main reason for this is the phonetic duration of VC rhymes as compared to V and VV rhymes.

There are two reasons to think that the vowel in VC sequences was phonetically short, i.e. that these rhymes were pronounced [VC], not [V·C]. The first reason is theoretical and has been explained before: the C contributed to the duration of the rhyme, so that the vowel did not have to be lengthened in order to provide a duration cue for foot headship. The second reason is observational and involves back projection: the present-day languages that underwent OSL retain a length distinction between original [VC] and [V:C] rhymes (e.g. NLb **mɑn̄** vs. **ʃɔ́ɔp**, NHG *Mann* vs. *Schaf*), whereas the original length distinction between [V] and [V:] rhymes has been lost (e.g. NLb **draāɣən** vs. **vrɔ́ɔɣən**; a full merger is seen in NHG *tragen* versus *fragen*, both with the same long [a:]).

Imagine now a child who has to learn Middle Limburgian and has to decide on the number of moras in a [VC] rhyme. If such a rhyme is ambiguously monomoraic or bimoraic in languages with a [V] ~ [V:] opposition, it must be biased towards monomoraic in languages with a [V·] ~ [V:] opposition, because the duration of [VC] is closer to that of a monomoraic V in languages that pronounce this as a lengthened [V·] than in languages that pronounce this as a really short [V]. Thus, I propose that MLb children interpreted the [VC] rhyme as monomoraic. Some surface structures are shown in (13) and (14).

In (13), we see that closed syllables with short vowels consisted of one mora. The vowel heads the mora, and all consonants that follow the vowel are included as satellites into this same mora. In the case of monosyllabic words, no L is included in the structure; since these structures reflect a sentence-internal alignment with H*L, there is room for the L to be realized on the first syllable of the next word, analogously to what happens in (2).[8]

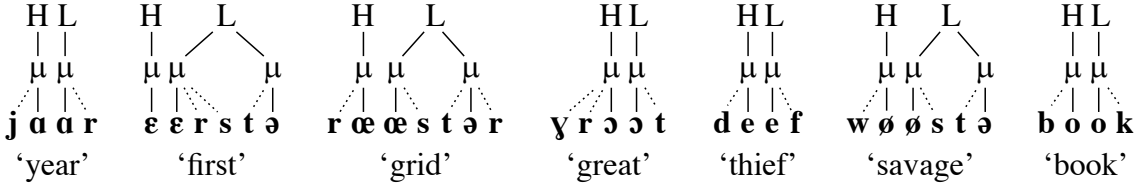
(13) *Early MLb structures for short vowels in closed syllables*



[8] Sentence-finally, the L would either have to be deleted or (more interestingly) incorporated somehow into the last mora (a case of *tone crowding*).

In (14) we see that closed syllables with long vowels consist of two moras. There is always room to host both the H and the L tone.

(14) *Early MLb surface structures for long non-high vowels in closed syllables*



If, as usual, the intonational cue to stress was implemented as a tonal movement directly *after* the first mora, **deev* must have been pronounced as [d́éèf] or [dèéf] depending on the intonation contour of the sentence.

Present-day reflexes. The structures in (14) have survived in modern Limburgian, if a final schwa was not lost later. The forms are either MLb monosyllables (e.g. most masculine and neuter singular nouns, the imperative singular of most strong verbs, and the independent forms of most adjectives) or MLb polysyllables with the first syllable ending in -**s** or -**r**. The predictability is perfect: all of these forms have an acute accent and I am not aware of any exceptions in Geleen, Roermond, or Sittard.[9]

(15) *Modern forms with originally long non-high vowels in closed syllable: acute*

Early MLb ***ɑɑ** > **ɔ́ɔ**: **ʃlɔ́ɔp** 'sleep-IMPSG' < ***slɑɑp**, **lɔ́ɔt** 'let-IMPSG', **ʃlɔ́ɔp** m. 'sleep' (OLF *slâp*), **ʃɔ́ɔp** n. 'sheep' < ***sxɑɑp**, **jɔ́ɔr** n. 'year' (OLF *iâr*), **wɔ́ɔr** 'true', **ʃwɔ́ɔr** 'heavy', **ɦɔ́ɔr** m. 'hair', **dɔ́ɔt** m. 'deed' (OLF *misdât* 'crime'), **rɔ́ɔt** m. 'advice', **drɔ́ɔt** m. 'thread', **pɔ́ɔl** m. 'pole', **mɔ́ɔltiit** m. 'meal', **krɔ́ɔm** 'shop', **pɔ́ɔʃən** 'Easter' < ***pɑɑsxən** 'Easter-DATPL'

Early MLb ***ææ** > Gel. **éa**, Rm. Sitt. **ée**: (no instances known)

Early MLb ***ɛɛ** > Gel. **éa**, Rm. Sitt. **ée**: **ɦéal** 'whole' (cf. non-monophtongized OHG *heil*), **éaʃtə** (Rm. **éerstə**) 'first-PL/FEMSG' < ***ɛɛrstə** (MHG *êrste*), **éaʃtən** (Rm. **éerstən**) 'first-MASCSG'

Early MLb ***œœ** > Gel. **ǿa**, Rm. Sitt. **ǿø**: **rǿastər** m. 'grid'

Early MLb ***ɔɔ** > Gel. **óa**, Rm. Sitt. **óo**: **ɣróat** 'great', **dóat** m. 'death' (OLF *dôt*), **dóat** 'dead', **róat** 'red', **lóan** n. 'salary', **ʃtóat** m. 'thrust', **tóan** m. 'tone', **ʃtóattóan** 'acute accent', **ɦóax** 'high' < ***ɦɔɔx** (i.e. not ***ɦɔɔx**)

Early MLb ***ee** > Gel. Rm. **ée**, Sitt. **éi**: **déef** m. 'thief', **léef** 'nice, cute', **léex(t)** 'light' (OLF *lioht*), **déep** 'deep', **ʃéet** 'shoot-IMPSG', **ɣənéet** 'enjoy yourself'

Early MLb ***øø** > Gel. Rm. **ǿø**, Sitt. **œi**: **wǿøstən** 'savage-MASCSG'

Early MLb ***oo** > Gel. Rm. **óo**, Sitt. **ɔ́u**: **bóok** n. 'book', **dóok** m. 'cloth' (NHG *Tuch* n.), **ɦóot** m. 'hat', **ɦóok** m. 'corner', **ɦóostən** 'cough', **ɣóot** 'good', **blóot** 'blood', **móot** 'mood, courage', **ɦóon** 'chicken', **ʃtóol** 'chair'

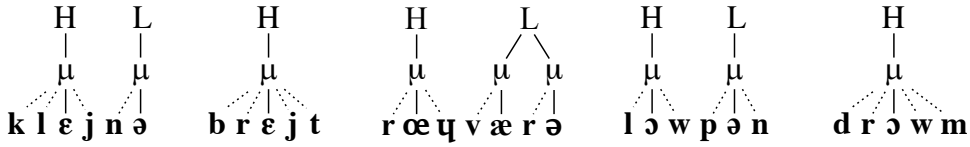
[9] Some imperative singulars have a circumflex in the Low Franconian – Ripuaric transition dialects of Moresnet (Jongen 1972: 42): **ʃlɔɔ̃p**.

4.5. Diphthongs

I propose that the MLb diphthongs ***ɛi***, ****œy***, and ****ɔu*** were monomoraic. When short vowels in open syllables lengthen to half-long, their duration will approach that of the diphthongs, which do not have to lengthen (their duration is already compatible with foot headship). As the duration of short vowels in open syllable increases, a child learning MLb becomes more likely to perceive a diphthong as a single mora, as long as the comparably long lengthened vowel is perceived as a single mora.

Structurally, the diphthongs ***ɛi***, ****œy***, and ****ɔu*** can be interpreted as monomoraic if we regard them as the underlying VC sequences /ɛj/, /œɥ/ and /ɔw/, and if coda consonants are still non-moraic. Conversely, if the diphthongs are VC sequences, they side with other VC sequences in not requiring lengthening of the first part in open syllables, and in being interpreted by learner as equally long as the lengthened vowels in open syllables, i.e. as monomoraic. We see some example words in (16), which assumes, as before, default left-to-right tone assignment on moras.

(16) *Early MLb structures for diphthongs*



If Middle Limburgian had the default phonetic interpretation of the link between segments and tones (§4.3), ****lɔupən*** must have been pronounced as [Ĺɔ́ẃpə̀ǹ] or [Ì\u0300ɔ̀ẁpə́ń] depending on the intonation contour of the sentence. Aligned with H*L, the words in (16) sounded like [kĹɛ́jǹə̀], [bŕɛ́jt], [ŕœ́ɥ̀væ̀ŕə̀], [Ĺɔ́ẃpə̀ǹ], and [d́ŕɔ́ẃḿ], with the usual restriction of not sounding tone on voiceless segments.

The new situation. If we compare the data in (16) with those of the previous section, we see that Middle Limburgian had moving tones on long vowels, and level tones on diphthongs. The same diversion between monophthongs and diphthongs occurred in the early history of Lithuanian tonogenesis. Garde (1976) gives the form /sé:dé:teī/, in which the two long vowels receive an acute, and the diphthong a circumflex.[10]

4.6. The long high vowels: monomoraic

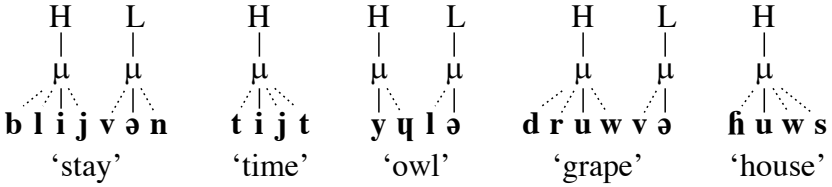
The cross-linguistic weight distinction between CVV and CVC (§4.4) and the parallel development found in Lithuanian (§4.5) are enough for me to judge the proposed early MLb structures as typologically feasible. We do still have to explain why the high vowels (****ii***, ****yy***, ****uu***) side with the diphthongs and short vowels rather than with the other long vowels, as they do in Lithuanian. In Lithuanian, the long high vowels receive an acute accent, suggesting that they were bimoraic in Lithuanian just like the non-high vowels. By contrast, I propose that in early MLb the long high vowels were monomoraic. Unlike Lithuanian /i:/ and /u:/, the MLb long high vowels did not have

[10] Garde's notation indicates that the diphthong was bimoraic, as in a later stage of MLb. It may have started its life as monomoraic, though.

to contrast with short vowels (all MLb short vowels were non-high), so that they were not pushed towards the duration of 220 ms that I proposed for the non-high long vowels. Also, long high vowels tend to be universally shorter than long non-high vowels, so that their duration may have been shorter than the old duration of the non-high vowels, which was 200 ms. Perhaps it was 170 ms, close to the duration of the lengthened short vowels. It was easy for a learner, then, to interpret these high vowels as monomoraic.

If the MLb long high vowels were monomoraic, there are still two possibilities for their phonological surface representation: they were phonologically simple short vowels (/i/, /y/, /u/), or they were, like the diphthongs, VC structures (/ij/, /yɥ/, /uw/). It is unlikely that *__ii__ and *__uu__ were short, since their later reflexes before /st/ are long (Rm. __liīstən__ 'lists', __vuūstən__ 'fists'), whereas the later reflexes of the short vowels before /st/ are still short (Rm. __kɑstən__ 'cupboards', __kɔstən__ 'cost', __kestən__ 'chests', __kœstən__ 'coasts'). So it is likely that *__ii__ and *__uu__ were diphthongal VC structures (with non-moraic C). Liberman (1999) actually argues that *__ii__ and *__uu__ were the diphthongs /ij/ and /uw/ in the whole West-Germanic area. This would explain the fact that when later the lengthening of vowels in open syllables started to crowd the vowel space, the high vowels changed towards [aj] and [aw] independently in at least three unconnected areas (England, Brabant, and Bavaria, ultimately leading to the diphthongs in 'wine' and 'house' in standard English, Dutch, and German). The fact that none of the Scandinavian languages has developed this kind of diphthongs, despite the same lengthening of vowels in open syllables, should then be attributed to the idea that the Scandinavian high vowels were genuinely bimoraic /iː/ and /uː/. Thus, an explanation for the difference between Lithuanian and Limburgian is given in the structures in (17).

(17) *Early MLb structures for long high vowels*



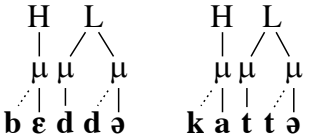
Foreshadowings. Given the diversion between non-high monophthongs on the one hand, and diphthongs and high vowels on the other hand, we can now already see the contours of an explanation for the behaviour of the words in §1 and §2: modern __ʃlɔ́ɔpən__ 'large' has a moving tone because it has an original long vowel (MLb *__slɑɑpən__), while __brɛ̄it__ 'broad' and __lɑ̄ŋk__ 'long' have level tones because they have an original diphthong or short vowel. But we can also see that several issues still have to be settled: the apparent bimoraicity of modern __brɛ̄it__ and __lɑ̄ŋk__ is related to a MLb process of open syllable lengthening (§5), and the fact that the moving tone in __knɪ́in__ 'rabbits' is different from the level tone in __knīin__ 'rabbit' is related to a later process of final schwa drop (§7).

4.7. Geminate consonants: moraic

There is an exception to the non-moraicness of coda consonants. If the coda consonant is the first half of a geminate, it has to be represented lexically by a separate mora, at least according to the usual view of the singleton–geminate distinction in universalist

phonology. Thus, both long vowels (§4.1) and long consonants have an additional underlying mora. Surface structures will reflect this. Examples are in (18).

(18) *Early MLb structures for geminates*



I propose, therefore, that geminates project a mora, whereas syllable-final sonorants do not, e.g. that *kattə 'cat' has three moras but *kantə 'side' (§4.4) only two.

Under the usual assumptions, the articulations must have been [ɓɛ́dd̀ə̀]$_{Art}$ and [ḱátt̀ə̀]$_{Art}$. The auditory forms must just have been [ɓɛ́dd̀ə̀]$_{Aud}$ and [káttə̀]$_{Aud}$, because tones can only be heard on voiced consonants. In other words, an intrasyllabic fall was heard on *bɛddə, but not on *kattə.

4.8. The predictable accent contrast and its phonetic correlates

The previous sections illustrate that Early MLb and other Franconian dialects had a predictable (i.e. non-lexical) phonological accent contrast. The accent could be predicted by the number of moras in the syllable, which was always unambiguously derivable from the phonetic duration.

First, there was the predictable phonetic acute accent: when aligned to H*L, this was realized as a short H mora followed by L. It audibly occurred in the long non-high vowels [áà], [ǽæ], [éɛ̀], [œœ̀], [ɔɔ̀], [éè], [øø̀], and [óò], and in the voiced geminates such as [énː] and [éd̀ː]. There was also a predictable phonetic circumflex accent: when aligned to H*L, this was realized as a long H mora. It audibly occurred in the lengthened vowels [á·], [ǽ·], [é·], [ɔ́·], [é·], [ǿ·], and [ó·], in the long non-high vowels [íj], [ýu̯], and [úw̯], in the diphthongs [éj], [œ́u̯], and [ɔ́w̯], and in sequences of a short vowel and a voiced non-geminate coda (which was always a sonorant) such as [éń]. The remaining rhymes were sequences of a short vowel and a voiceless coda. The geminate case was articulatorily acute, i.e. had an early fall ([és̀s̀]$_{Art}$), while the non-geminate case was articulatorily circumflex, i.e. had a late fall ([éś]$_{Art}$), but auditorily both were ambiguous ([és]$_{Aud}$, [éss]$_{Aud}$), because tone cannot be heard on voiceless segments.

5. The Late Middle Limburgian bimoraic reanalysis

It is a surprisingly small step from the predictable phonetic accent contrast described in §4.8 to a lexical accent contrast. Since the predictability in §4.8 depends on knowing the number of moras in the first syllable, all that is needed for the switch to occur is that the lexical length contrast, which relies on having different numbers of moras in a syllable, is suspended. Precisely that is what happened when Open Syllable Lengthening proceeded in Late Middle Limburgian and its Central Franconian neighbours.

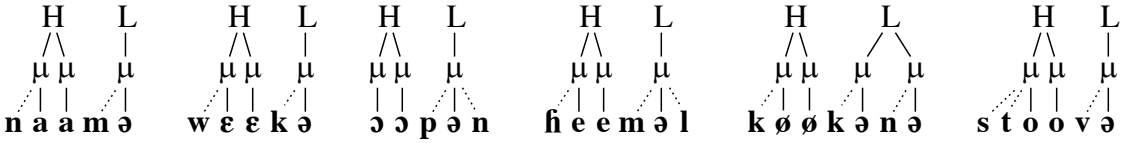
5.1. Even more lengthening of vowels in open syllables

In §4.3 we saw that as a result of the West-Germanic initial stress shift, phonologically short (i.e. monomoraic) vowels in open syllables were phonetically half-long. Thus, ***makən** 'make' was pronounced [maˑkən].

When this lengthening progressed over the generations, the phonetic durations of the phonologically short vowels in open syllable came very close to the durations of the phonologically long vowels. As a result, some learners decided to ignore the duration distinction between the two groups and to analyse the lengthened vowels as bimoraic. We know that this happened in the histories of many current Germanic languages, and our Franconian dialects formed no exception. But whereas in most languages the lengthening vowels completely coalesced with the originally long vowels, the Franconian dialects kept the distinction alive, albeit not in the mora count.

The location of the pitch drop in Early MLb was *after* the first syllable in these words, i.e. [ḿáˑkə̀ǹ] and [ńáˑm̀ə̀], and it is this syllable alignment that the learners honoured in their newly constructed phonological representations. If the vowels in [ḿáˑkə̀ǹ] and [ńáˑm̀ə̀] are reinterpreted as bimoraic, and the tone movement is interpreted as occurring after the first syllable, the learner has no option but to interpret the H tone as including both the first and the second mora. The structures in (8) now become those in (19). For the lengthened vowels, the distinction between ***æ** and ***ɛ** was lost, either at this point or later; I will write both as **ɛɛ** from now on.

(19) *Late MLb surface structures for lengthened non-high vowels*



Present-day reflexes. For the lengthened non-high vowels in (19), the structures in (19) have survived to the present day, at least for those words that did not lose the second syllable (phonetically, the present-day circumflexes are longer than the acutes, a length reversal that I return to in §9). Examples from present-day Limburgian are listed in (20) and (21). All of these forms necessarily have two or more syllables, since no short vowels occurred word-finally in Early MLb monosyllables.

(20) *Modern forms with lengthened vowels in open syllables, non-voiced case: always circumflex*

Early MLb ***a** > **aā**: **maākən** 'make', **naātə** 'wet-PL/FEMSG', **aāpən** m.pl. 'apes'

Early MLb ***æ** > **ɛɛ̄**: **wɛɛ̄kən** f.pl. 'weeks', **brɛɛ̄kən** 'break', **ʃprɛɛ̄kən** 'speak', **ʃtɛɛ̄kən** 'sting', **ɛɛ̄tən** 'eat', **ɣɛɛ̄tən** (Rm. **ɣəɣɛɛ̄tən**) 'eaten', **bɛɛ̄tər** 'better' < ***bɛtərə** (OLF *betera*)

Early MLb ***ɛ** > **ɛɛ̄**: **bɛɛ̄kən** f.pl. 'brooks', **pɛɛ̄pər** m. 'pepper'

Early MLb ***ɔ** > **ɔɔ̄**: **ɔɔ̄pən** 'open (adj.)' < ***ɔpən**, **ɣəzɔɔ̄pən** 'boozed', **ɣəbrɔɔ̄kən** 'broken', **ɣəʃprɔɔ̄kən** 'spoken', **ɣəʃtɔɔ̄kən** 'stung', **ɣənɔɔ̄tən** 'enjoyed', **ɣərɔɔ̄kən** 'smelled', **ɣəʃlɔɔ̄tən** 'closed'

Early MLb ***e** > **eē**: **ɣəweēkən** ‘give way’, **ɣəleēkən** ‘seemed’, **ɣəbeētən** ‘bitten’, **leēkən** ‘leak’

Early MLb *ø > **øø̄**: **køø̄kən** f. ‘kitchen’ < ***køkənə** (MHG *küchene*), **ʃløø̄təl** m. ‘key’

Early MLb *o > **oō**: (no examples)

(21) *Modern forms with lengthened vowels in open syllables, voiced case: usually circumflex*

Early MLb *a > **aā**: **draāɣən** ‘carry’, **jaāɣən** ‘hunt’, **klaāɣən** ‘complain’, **vaārən** ‘to drive’, **bətaālən** ‘pay’, **naāmən** m.pl. ‘names’, **ʃmáalən** ‘meagre-MASCSG’

Early MLb *æ > **ɛɛ̄**: **ɣɛɛ̄vən** ‘give’, **lɛɛ̄vən** ‘live’, **lɛɛ̄zən** ‘read’, **rɛɛ̄ɣən** ‘rain’, **tɛɛ̄ɣən** ‘against’ < ***təɣɛɣən**, **ɛɛ̄zəl** ‘donkey’, **lɛɛ̄vər** ‘liver’, Rm. **véɛɣən** ‘wipe’, Rm. **bəwéɛɣən** ‘move’

Early MLb *ɛ > **ɛɛ̄**: **zɛɛ̄ɣən** ‘saws’, Rm. **zéɛɣən** ‘saw’ (denominative verb)

Early MLb *ɔ > **ɔɔ̄**: **bɔɔ̄vən** ‘above’, **ɔɔ̄vən** m. ‘oven’, **bəlɔɔ̄vən** ‘promise’, **ɦɔɔ̄lən** ‘fetch’ (NHG *holen*; note that Du. *halen* is a false friend), **ɣəʃtɔɔ̄lən** ‘stolen’, **ɣəlɔɔ̄ɣən** ‘lied’, **ɣənɔɔ̄mən** ‘taken’, **ɣəbɔɔ̄jən** ‘bid (p.p.)’ < ***ɣəbɔðən**, **ɣəbɔɔ̄rən** ‘born’, **vərlɔɔ̄rən** ‘lost’, **ɣəkɔɔ̄zən** ‘chosen’ < ***ɣəkɔrən**[11], **ɣrɔ́ɔvən** ‘course-MASCSG’

Early MLb *e > **eē**: **ʃpeēlən** ‘play’, **ɦeēməl** m. ‘heaven’, **zeēvən** ‘seven’, **ɣəbleēvən** ‘stayed’, **reēɣəl** m. ‘ruler’ (NHG *Riegel*)

Early MLb *ø > **øø̄**: **øø̄vər** ‘over’ < ***øvər** (NHG *über*), **bøø̄ɣəl** m. ‘braces’ (NHG *Bügel*), **møø̄lən** f. ‘mill’ < ***møləna**, **ɣəbøø̄rən** ‘happen’, **løø̄ɣən** f. ‘lie’ (NHG *Lüge*)

Early MLb *o > **oō**: **zoōmər** m. ‘summer’, **woōnən** ‘dwell’, **voōɣəl** m. ‘bird’

The new situation. For the originally long non-high vowels, the structures will still be those in (9). The alignment of H*L with moras is no longer predictable from the number of moras in the syllable. It is possible, of course, that the H*L alignment was predictable from vowel quality, i.e. the lengthened vowels might have been ‘lax’, the originally long vowels ‘tense’, or some such contrast. However, at some point in time the qualities of the lengthened vowels have fallen together with those of the originally long vowels, in all languages involved. Present-day East Limburgian constitutes no exception, as we can see when comparing (11), (12) and (15) with (20) and (21) for present-day **ɔɔ**, **ee**, **oo**, and **øø**.

If Early MLb ***ɑɑ** had already shifted to ***ɔɔ** in Late MLb (and Early MLb ***ɔɔ** had shifted to ***oa**), then it is possible that the four originally long MLb vowels ***ɔ́ɔ**, ***ée**, ***óo**, and ***øǿ** already coalesced with the lengthened vowels ***ɔɔ̄**, ***eē**, ***oō**, and ***øø̄** in Late MLb. In that case, Late MLb had a lexical mora-accent contrast for these four vowels.

[11] With a **z** introduced from the infinitive **kéezən**.

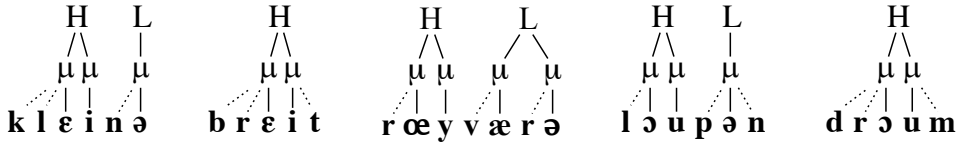
We do not have sufficient information yet as to whether and where the vowel groups coalesced. Some relevant examples, though, can be found. From 14th century Sittard, forms like *daighen* 'days-DAT' (< ***daɣən**), *saicken* 'things' (< ***zakən**), and *aepen/oepen* 'open' (< ***ɔpən**) have been transmitted (Dols 1944: 200–202), with *i* and *e* used as length symbols, indicating that the vowel lengthening had reached phonological (i.e. bimoraic) status; this is confirmed by writing an original ***ɑɑ** with a single vowel symbol, as in *beswaren* 'burden' and *na* 'after', which indicates coalescence with the lengthened vowel.

If such written evidence is valid, a lexical mora-accent contrast must have emerged for the non-high long vowels already in Late MLb. But even if Late MLb vowels were also distinguishable on the basis of tenseness or some such thing, the mora alignment distinction must have been a major auditory cue to the new distinction between ***ɔ́ɔ/ée/ǿø/óo** on the one hand and ***ɔɔ̄/eē/øø̄/oō** on the other. In such a case, one could just as well say that vowel quality was predictable from mora alignment as that mora alignment was predictable from vowel quality.

5.2. Bimoraic reanalysis of diphthongs

Now that 'short' vowels had been reanalysed as bimoraic, the diphthongs had to follow suit. That is, children had to interpreted these as bimoraic *a fortiori*. We can therefore use VV notations (/ɛi/, /œy/, /ɔu/) again. The structures in (16) now become those in (22).

(22) *Early MLb structures for diphthongs*



Present-day reflexes. In words that did not lose their final syllable, the diachronic predictability of the accent is good. As usual, some morphological classes and disyllabic voiced cases diverge.

(23) *Modern forms with diphthongs in open syllables: non-voiced cases: always circumflex*

Early MLb ***ɛi** > **ɛī**: **ɦɛītə** 'hot-PL/FEMSG', **ɦɛītən** 'hot-MASCSG', **wɛīkə** 'weak-PL/FEMSG', **wɛīkən** 'weak-MASCSG', **blɛīkə** 'pale-PL/FEMSG', **blɛīkən** 'pale-MASCSG', Rm. **ɦɛītən** 'be called', **ʃwɛītən** 'sweat', **tɛīkən** n. 'sign (*token*)', **ʃlɛīpən** 'drag'

Early MLb ***œy** > **œī**: **dœīpən** 'baptise' (NHG *taufen* has lost the umlaut)

Early MLb ***ɔu** > **ɔū**: **lɔūpən** 'run, walk INF', **kɔūpən** 'buy-INF'

(24) *Modern forms with diphthongs in open syllables: voiced cases: usually circumflex*

Early MLb ***ɛi** > **ɛī**: **ɛīɣən** 'own (adj.)', **klɛ́inən** 'small-MASCSG', **ʃɛ́ivən** 'skewed-MASCSG', **ɛ́inən** 'one-MASCSG'

Early MLb ***œy** > **œī**: **drœīmən** ‘dream’, **ɣlœīvən** ‘believe’ (NHG *glauben* has lost the umlaut), **bœīɣən** ‘bend, bow’ (note that NHG *beugen*; Du. *buigen* [bœyɣə] is a false friend)

Early MLb ***ɔu** > **ɔū**: **ɔūɣən** n.pl. ‘eyes’, **dɔ́uvən** ‘deaf-MASCSG’

(25) *Modern forms with diphthongs in closed syllables: always circumflex*

Early MLb ***ɛi** > **ɛī**: **brɛīt** ‘broad’, **ʃɛīf** ‘skewed’ (Du. *scheef*, NHG *scheib*), **wɛīk** ‘weak’, **blɛīk** ‘pale’, **ɦɛīt** ‘hot’, **lɛīt** n. ‘suffering’, **rɛīp** m. ‘bar’, **ʃtɛīn** m. ‘stone’, **bɛīn** n. ‘bone, leg’, **ɦɛīm** n. ‘home’, Gel. Sitt. **ɦɛīʃən** ‘be called’ < ***(ɦ)ɛisxən** ‘ask’ (NHG *heischen* ‘strive’), **ɣɛīs(t)** ‘ghost’, **mɛīstər** ‘master’, **ʃlɛīptóan** ‘circumflex accent’

Early MLb ***œy** > **œī**: (probably no examples)

Early MLb ***ɔu** > **ɔū**: **lɔūp** ‘run, walk (*leap*) IMPSG’, **knɔūp** m. ‘button’, **rɔūk** m. ‘smoke’, **lɔūk** n. ‘leek’, **ɦɔūp** m. ‘heap’ (note that NHG *Haufen* < MHG *hûfe* is a false friend), **ɔūx** (Venlo **ɔūk**) ‘also’ < ***ɔuk** (**k** > **x** testifies of being in the Rhenish fan), **bɔūm** m. ‘tree (*beam*)’, **zɔūm** m. ‘seam’, **drɔūm** m. ‘dream’, **tɔūm** m. ‘bridle (*team*)’, **dɔūf** ‘deaf’

5.3. Bimoraic reanalysis of long high vowels

High vowels were now interpreted as bimoraic as well. The forms in (17) were reinterpreted as those in (26), with two high-toned moras (in sentence-internal H*L alignment) each carries by a full vowel.

(26) *Late MLb structures for long high vowels*

Present-day reflexes. The structures in (26) have survived in NLb, at least if the final syllable was not lost.

(27) *Modern forms with high vowels in open syllables: non-voiced case: always circumflex*

Early MLb ***ii** > **iī**: **kiīkən** ‘look’, **liīkən** ‘look like’, **ʃtriīkən** ‘spread’, **biītən** ‘bite’, **ʃiītən** ‘shit’, **riīkə** ‘rich-PL/FEMSG’, **riīkən** ‘rich-MASCSG’, **vriȷ̄** ‘free’, **biȷ̄** ‘by’, **ɦiȷ̄** ‘here’

Early MLb ***yy** > **yȳ**: **kyȳkən** n. ‘chicken’ (< *kiukîn*), **ryȳkən** ‘smell’

Early MLb ***uu** > **uū**: **bruūkən** ‘need’, **ɣəbruūkən** ‘use’, **ʃluūtən** ‘close’, **nuw̄** ‘now’, **ʃuw̄** ‘shy’

(28) *Modern forms with high vowels in open syllables: voiced case: usually circumflex*

Early MLb *ii > **iī**: **ʃriīvən** ‘write-INF’, **bliīvən** ‘stay-INF’, **wiīzən** ‘point-INF’, **ʃtiīvən** ‘starch-INF’, **wíizən** ‘wise-MASCSG’, **ɣríizən** ‘grey-MASCSG’, **ʃtíivən** ‘stiff-MASCSG’

Early MLb *yy > **yȳ**: **zyȳvər** ‘clean’, **yȳlən** m.pl. ‘owls’, **ɦyȳrən** ‘hire’, **ʃtyȳrən** ‘send, steer’

Early MLb *uu > **uū**: **druūvən** f.pl. ‘grapes’, **duūvən** f.pl. ‘doves’, **ʃruūvən** f.pl. ‘screws’, **duūmən** m.pl. ‘thumbs’, **duūzənt** ‘thousand’

(29) *Modern forms with high vowels in closed syllables: always circumflex*

Early MLb *ii > **iī**: **tiīt** m. ‘time’, **wiīt** ‘wide’, **iīs** n. ‘ice’, **wiīs** ‘wise’, **ɣriīs** ‘grey’, **ʃriīf** ‘write-IMPSG’, **bliīf** ‘stay-IMPSg’, **wiīn** m. ‘wine’, **liīstən** f.pl. ‘lists’

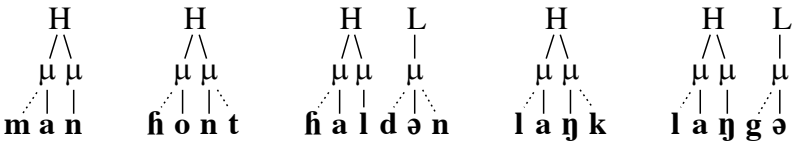
Early MLb *yy > **yȳ**: **vyȳr** ‘fire’

Early MLb *uu > **uū**: **ɦuūs** ‘house’, **uūt** ‘out’, **muūs** ‘mouse’, **luūs** ‘louse’, **bruūt** [12] ‘bride’, **vuūs(t)** ‘fist’, **ʃuūm** ‘foam’, **zuūr** ‘sour’

5.4. Bimoraic reanalysis of vowel-sonorant sequences

The long H tone was audible on all sequences of short vowel and tautosyllabic sonorant coda. As in the case of the diphthongs and high vowels, learner will have interpreted this sequence as a bimoraic circumflex. The relevant cases in (13) will be reanalysed as the structures in (30).

(30) *Early MLb structures for short vowels in closed syllables*



Present-day reflexes. The structures in (30) have been preserved in NLb, at least in words where no syllable was later lost (by final schwa drop, see §7) or gained (by schwa insertion into sequences like **-lf** and -**rx**, see §12.2) and otherwise no segments dropped (by lenition and degemination, see §7.7 and §12.1). The present forms are listed in (31) and (32). As usual, forms where there is a *following* voiced consonant are listed separately; all these are from Venlo, because lenition and degemination later deleted a segment in the other dialects, causing these forms to go into a different accent class (§7.7).

(31) *Modern forms with tautosyllabic vowel-sonorant sequences: non-voiced case: always circumflex (for nonhomorganic cluster cases see §12.2)*

Early MLb *am > ɑm̄: **dɑm̄p** m. ‘vapour’

[12] The underlining of the **t** is based on the MLb reconstructed form, which must have had **d** (e.g. NHG *Braut*). Present-day Roermond has the unexpected plural **bruūtən**.

Early MLb ***om** > **om̄**: **dom̄p**[13] ‘dumb’, **kom̄p** m. ‘bowl’ (MHG *kumpf*)

Early MLb ***an** > **ɑn̄**: **mɑn̄** ‘man’ < *__man__, Venlo **kɑn̄** ‘can-1,3SG’, **vɑn̄** ‘of, from’, **ɦɑɲc̲** f. ‘hand’ < *__ɦant__, **lɑɲc̲** n. ‘land’, **zɑɲc̲** n. ‘sand’, **wɑɲc̲** f. ‘wall’, **kɑɲcən** f.pl. ‘sides’ < *__kantən__

Early MLb ***en** > **en̄**: **en̄** ‘in’, Rm. **kiɲc̲**, Gel. **kɛɲc̲** n. ‘child’, Rm. **weɲcər**, Gel. **wɛɲcər** m. ‘winter’

Early MLb *__on__ > **on̄**: **ɦoɲc̲** m. ‘dog’, **ɣroɲc̲** m. ‘ground’, **moɲc̲** m. ‘mouth’, **ɣəzoɲc̲** ‘healthy’, **bloɲc̲** ‘blonde’

Early MLb *__aŋ__ > **ɑŋ̄**: **lɑŋ̄k̲** ‘long’, **dɑŋ̄k** m. ‘thanks’, **ɣɑŋ̄k̲** m. ‘gait, passage’

Early MLb *__ɛŋ__ > **ɛŋ̄**: **dɛŋ̄kən** ‘think’

Early MLb *__eŋ__ > **eŋ̄**: Rm. **deŋ̄k̲** , Gel. **dɛŋ̄k̲** n. ‘thing’, Rm. **reŋ̄k̲** , Gel. **rɛŋ̄k̲** m. ‘ring’

Early MLb *__al__ > **ɑl̄**: **bɑl̄** m. ‘ball’, **ɑl̄** ‘all’, **vɑl̄** m. ‘fall’, **kɑl̄** m. ‘chat’, **zɑl̄** ‘shall-1,3SG’, **wɑl̄** m. ‘wall’, Venlo[14] **ɑl̄t̲** ‘old’, Venlo **kɑl̄t̲** ‘cold’, Venlo **wɑl̄t̲** n. ‘wood’, Venlo **zɑl̄t** n. ‘salt’

Early MLb *__æl__ > **æl̄**: **ɣæʎ̄c̲** n. ‘money’, **ɦæl̄** ‘hard’

Early MLb *__ɔl__ > **ɔl̄**: **vɔl̄** ‘full’

Early MLb *__el__ > **el̄**: **wel̄** ‘want-1SG’, Rm. **weʎ̄c̲**, Gel. **wɛʎ̄c̲** ‘wild’

(32) *Modern forms with tautosyllabic vowel-sonorant sequences: voiced case: usually circumflex*

Early MLb *__an__ > **ɑn̄**: Venlo **ɑn̄dər** ‘other’

Early MLb *__en__ > **en̄**: Venlo **ben̄dən** ‘bind’

Early MLb *__on__ > **on̄**: Venlo **on̄dər** ‘under’

Early MLb *__al__ > **ɑl̄**: Venlo[15] **ɦɑl̄dən** ‘hold’

5.5. Reanalysis of voiced and voiceless geminates

The next question to answer is how the learning child interpreted the auditory forms [bə́də̀də̀]$_{\text{Aud}}$ and [kə́ttə̀]$_{\text{Aud}}$ (§4.7). Now that the language allows acute as well as circumflex syllables, it is no longer evident that the child should interpret both forms as acute, as their parents were forced to do. In fact, we know that the present-day dialects tend to distinguish these forms (as least their MLb plurals *__bɛddən__ ‘beds’ and *__kattən__ ‘cats’ which did not undergo final schwa drop later). In the modern language, where the difference between the two groups has become larger, ‘beds’ has a clearly

[13] The underlined final clusters **mp**, **ɲc**, **ŋk**, and **ʎc** derive from the MLb ‘underlyingly voiced’ **mp**, **nt**, **ŋk**, and **lt**, which became **mb**, **nd**, **ŋg**, and **ld** before a vowel. Late MLb deleted the plosives in prevocalic position, leaving just the present-day **m**, **ɲ**, **ŋ**, and **ʎ** in prevocalic position (as well as word-finally in cases of schwa drop). See §7.7 for more detail.

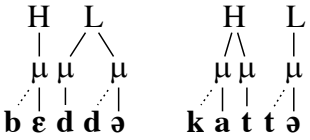
[14] These Venlo forms end in -**aāt** or -**aūt** etc. in the other dialects. According to Van der Meer (1949: 299), Venlo mainly had -**aat** in the 14th century as well, and -**alt**-like forms were reintroduced in the 15th century.

[15] This form is **ɦaājən** in the other dialects. The same comment as in the previous footnote applies here.

audible acute whereas ‘cats’ has a clearly audible circumflex (Ben Hermans, p.c.; see §12.1 for details).

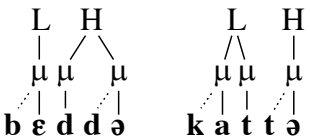
Although the acute for ***bɛddə** and the circumflex for ***kattə** could have arisen later, it is tempting to seek the cause of the distinction in the bimoraic reanalysis of Late MLb, when children had the first chance to create it. Thus, I propose that the structures of (18) turned into the contrasting pair in (33).

(33) *Late MLb structures for geminates, aligned to H*L*



Taking the structures in (33) at face value, the first thought of a phonologist is that the explanation for the split must lie in the commonly attested correlation of tone with voice, namely that H tones go together with voicelessness and L tones with voicedness. This correlation is visible in (33), and in fact a similar explanation has been provided by Hermans (2006) for a similar but different case in Limburgian. However, the voice-tone correlation exemplified so beautifully in (33) in fact owes its plausibility entirely to the presence of the H*L (declarative) contour that I used to illustrate sentence-internal surface structure. I could just as well have used the L*H (interrogative) contour. In that case, the surface structures would have been those in (34).

(34) *Late MLb structures for geminates, aligned to L*H*



Both the structures in (33) and those in (34) must have coexisted in Late MLb. Unless the declarative contour H*L was *much* more frequent than the interrogative (or continuative) contour L*H, no universal correlation between voice and tone can explain why *bɛ́ddə received an acute accent and *kat̄tə a circumflex.

I propose instead that the split was caused by a difference in the audibility of the tones. The auditory form [bɛ́ɛ̀dɛ̀ə]$_{\text{Aud}}$ contains much more information about the tone on the second mora than the auditory form [kátːə]$_{\text{Aud}}$ does. It is not surprising that [bɛ́ɛ̀dɛ̀ə]$_{\text{Aud}}$, with its overtly low first [d], was interpreted as acute. The form [kátːə]$_{\text{Aud}}$, by contrast, had a voiceless second mora (the first [t]), which was ambiguous with regard to tone, so we know that it was phonetically more cicumflex than [bɛ́ɛ̀dɛ̀ə]$_{\text{Aud}}$. But if [kátːə]$_{\text{Aud}}$ was ambiguous, why did children decide to analyse it as circumflex? Wouldn’t an acute analysis have been equally likely?

One explanation could involve a directional on-line perceptual-phonological bias. If you hear a H tone, you are waiting for the next L tone, and you will decide that the tone has changed to L only when you hear positive evidence of it, i.e. a low pitch on a voiced segment. If a voiceless stretch intervenes between the H vowel and the L vowel, this stretch contains no information on the fall, so it is better for you not to assume yet that it has happened. After all, the next vowel could still be high-pitched, in which case you should perceive a continuous H tone shared by both syllables. If you wait until you get

positive evidence, you will perceive both a disyllabic H tone and a disyllabic HL contour correctly. So the second, voiceless mora of [kátt**ə**]$_{Aud}$ is perceived with the same tone as the preceding [a].

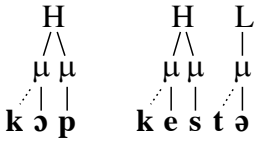
Another explanation could lie in the non-directional competition between mora-based and syllable-based cues. Perceiving the second mora of [kátt**ə**]$_{Aud}$ as having a H tone (i.e. perceiving the first syllable as circumflex) violates the cue constraint that says that "a mora without phonetic high pitch should not be perceived as a phonologically high-toned mora". Perceiving the second mora of [kátt**ə**]$_{Aud}$ instead as having a L tone (i.e. perceiving the first syllable as acute) would violate the cue constraint that says that "a syllable without a phonetic pitch movement should not be perceived as an acute". If the syllable-based constraint outranks the mora-based constraint, i.e. if the ranking follows the prosodic hierarchy, the interpretation of [kátt**ə**]$_{Aud}$ as circumflex is explained.

A third explanation could lie in a language-specific structural constraint. A comparison of the number of forms in (11), (12) and (15) on the one hand, and (20), (21), (23), (24), (25), (27), (28), (29), (31) and (32) on the other suggests that in Late MLb the circumflexes must have outnumbered the acutes by a factor of two or so.[16] The MLb-specific markedness constraint "syllables are circumflex" could have caused the interpretation of [kátt**ə**]$_{Aud}$ as the default, i.e. as circumflex.

5.6. Reanalysis of voiceless non-geminate codas

Two of the forms in (13) have not been discussed yet. It is those with a voiceless coda that is not part of a geminate, namely ***kɔp** and ***kestə**. The stressed syllables of all forms that we have seen so far, including those ending only in a half-long vowel, were interpreted by Late MLb learners as bimoraic, and there is no reason to suppose that it would be otherwise for these forms. Following the same reasoning as in §5.5 (with the same three possible explanations), the accent must have been analysed as circumflex *a fortiori*, because the parent articulation [ḱéśt̀ə]$_{Art}$ is even less likely than [ḱátt̀ə]$_{Art}$ to have had a small cue on the first mora to the impending pitch drop. The structures that the child builds are those in (35).

(35) *Early MLb structures for short vowels in closed syllables*



Present-day reflexes. For words that did not lose final schwa (as e.g. ***kestə** itself), present-day Limburgian retains the structures in (35). The current pronunciation of these forms is such that the movement to L occurs well after the stressed syllable, as in (2), so that the circumflex on e.g. **kestən** is no longer inaudible.

[16] One counts all the forms, even the ones that have a different accent in present-day Limburgian (because the accent changes are of a later date). One should also count the Late MLb forms of the words introduced in later chapters, but these will not change the factor of two for Late MLb. For present-day Limburgian, the factor is much closer to 1, because of the later changes described from §7 on.

(36) *Modern forms with sequences of a short vowel and a voiceless consonant: circumflex*

kɔp̄ m. ‘head’ < *__kɔp__, **krɑx̄(t)** f. ‘strength’ < *__kraft__, **ex̄** (Venlo **ek**) ‘I’, **mex̄** ‘me’, **dex̄** ‘you-SG’, **zex̄** ‘herself, himself, themselves’, **ræx̄(t)** ‘right’ (OLF *reht*), **ʃlæx̄(t)** ‘bad’, **kēstən** f.pl. ‘chests’ < *__kestən__ (NHG *Kisten*)

5.7. The influence of voicing

The tonogenesis described in the previous section provides the first opportunity for the split between Rule A and Rule A2.

In general, the tone contrast is more difficult to hear on disyllables with voiced consonants than on disyllables with voiceless consonants. In [ɔ́ɔ́pə̀ǹ], the low tone starts one vowel plus one consonant later than in [slɔ́ɔ̀pə̀ǹ], whereas in [lǿǿɣ̀ə̀ǹ], the low tone starts only one vowel later than in [vǿǿɣ̀ə̀ǹ]. It is possible that some reanalysing children were capable of positing a tone opposition only on the voiceless cases, while merging the voiced cases into acutes just as the non-tonal Germans, Dutch and English must have done.

5.8. The new situation: a lexical contrast for non-high vowels in polysyllabic words

In Late Middle Limburgian, all stressed syllables were phonologically long, i.e. bimoraic (as in present-day Swedish, for instance). On the surface, every stressed syllable either had the acute accent (where each mora had a different tone) or the circumflex accent (where each mora had the same tone).

In many cases, the Late MLb accent was predictable by grammatical rule. High vowels, closing diphthongs, *__aa__, *__ɛɛ__, monosyllabic words with short vowels, and rhymes with voiceless codas or with non-geminate voiced codas were always circumflex. The vowels **ea**, **øa**, and **oa** and closed syllables with non-high long vowels were always acute. The only VC rhymes that were predictably acute were the geminate voiced codas like *__bɛ́ddə__. The explanation of this exception involved both a phonological and a phonetic observation: phonologically, geminates were the only VC rhymes that were forced to be bimoraic, hence had to host a L tone, and phonetically, voiced codas were the only codas on which an L tone could be audible.

In a minority of cases, namely for the vowels *__ɔɔ__, *__ee__, *__øø__, and *__oo__ in an open syllable in polysyllabic words, the accented mora had to be specified in the lexicon (e.g. by a star on the first mora for acutes, and on the second mora for circumflexes; this lexical star then aligns on the surface with the star in H*L or L*H). This is what made Late Middle Limburgian a lexical mora-accent language. Hermans (1984) argues that it still is.

6. Analogical lengthening

After lengthening of vowels in open syllable, some verbal, nominal, and adjectival paradigms contained alternations between short and long vowels. In these cases the paradigm was levelled towards the long version, so that many short vowels in closed syllables lengthened as well. Since the model form, with a vowel lengthened in the previous step, necessarily had a circumflex tone, the analogically lengthened vowel was circumflex as well.

In the verbal paradigm, not much happened. In weak verbs, all forms had a theme vowel (schwa), so all forms had been lengthened already (e.g. **lɛ̄və** ‘I live; live-

IMPSG', **lɛēvən** 'to live; we/they live', **lɛēvət** 'lives; you-PL live; live-IMPPL', **lɛēvəs** 'you-SG live'). In strong verbs, the infinitive was lengthened (**ɣɛēvən** 'to give'), as were several forms of the present indicative, namely the first person singular (**ɣɛēvə** 'I give'), the second person plural (**ɣɛēvət** 'you-PL give'), and the first and third person plural (**ɣɛēvən** 'we/they give'). The only consonant-final form in strong verbs was the imperative singular. This stayed short (cf. present-day **ɣef** 'give!'), perhaps because the second and third persons singular of the present indicative stayed short as well (cf. present-day **ɣøfs** 'you-SG give' and **ɣøf** 'gives').

In the nominal paradigm, the nominative/accusative singular of (mainly masculine) nouns with a plural ending in -**ə**, genitive singular in -**əs**, and dative singular in -**ə**, was lengthened: **dax** 'day' became **daāx** analogously to **daāɣə** 'days'. For neuter nouns with zero plural endings and for nouns without plurals, the source of analogy was restricted to the genitive or dative singular:[17] **ɦɔf** 'garden' became **ɦɔɔ̄f** analogously to **endən ɦɔɔ̄və** 'in the garden-DATSG', and **ɣras** 'grass' became **ɣraās** 'grass' analogously to **optəŋ ɣraāzə** 'on the grass-DATSG'.

In the adjective paradigm, the predicative form, e.g. **nat** 'wet' was lengthened to **naāt** in analogy with attributive forms such as the feminine and plural nominative singular (**naātə**) and the masculine accusative singular (**naātən**).

It will turn out (§6.2) to be relevant that analogical lengthening created circumflex tones independently of whether the following consonant (or the following consonant elsewhere in the paradigm) was voiced or voiceless. Thus, analogical lengthening is the source of the following present-day Limburgian forms with an original voiceless consonant:

(37) *Analogical lengthening in voiceless paradigms: circumflex*

Early MLb *a > aā: **daāk** f. 'roof', **ɣəmaāk** n. 'ease', **vaāt** n. 'barrel', **naāt** 'wet', **vlaāk** 'flat', **ʃwaāk** 'weak'

Early MLb *æ >> ɛē: (no cases known)

Early MLb *ɔ >> ɔɔ̄: **ʃlɔɔ̄t** n. 'lock', **ʃpɔɔ̄r** f./n. 'track'

Early MLb *e >> eē: **ɣəbeēt** n. 'teeth', **ʃeēp** n. 'ship', **bleēk** n. 'tinplate'

Early MLb *o >> oō: perhaps **noōt** 'nut' < ***not(ə)** (MHG *nuz*, MDu *note*)

It is also the source of the following forms that end in a voiceless consonant that corresponds to a voiced consonant elsewhere in the paradigm:

(38) *Analogical lengthening in voiced obstruent paradigms: circumflex*

Early MLb *a > aā: **daāx** m. 'day', **ɣraās** n. 'grass', **ʃlaāx** m. 'stroke', **paāt** n. 'path', **raāt** n. 'wheel', (for Rm. **káaf** n. 'chaff' see §7.3)

Early MLb *æ >> ɛē: **wɛēx** m. 'way', **ɣəbɛēt** n. 'prayer'

Early MLb *ɔ >> ɔɔ̄: **ɦɔɔ̄f** m. 'court', **ɣəbɔɔ̄t** n. 'commandment', **ɣrɔɔ̄f** 'coarse, rude' < ***ɣərɔf**

Early MLb *e >> eē: **ʃmeēt** m. 'smith' (pl. **ʃméej**), **ɣəleēt** n. 'member'

[17] Or sometimes perhaps the form that was the first part of a nominal compound.

Early MLb *o >> oō: (no cases known)

And it is the source of the following forms that end in a voiced consonant:

(39) *Analogical lengthening in sonorant paradigms: circumflex*

Early MLb *a > aā: **laām** 'lame', **ʃmaāl** 'narrow', **taām** 'tame'

Early MLb *æ >> ɛɛ̄: (no cases known)

Early MLb *ɔ >> ɔɔ̄: **ɦɔɔ̄l** 'hollow'

Early MLb *e >> eē: **ʃpeēl** n. 'game', **teēn** n. 'tin'

Early MLb *o >> oō: (no cases known)

A possible fourth group consists of four adjectives that are reported as ending in a consonant in Middle High German, but in a schwa in Middle Dutch. These are words that had *-*û* in West Germanic:

(40) *Possible analogical lengthening for û words: circumflex*

ɣɛɛ̄l 'yellow' (MHG *gël*, MDu *gele*)
ʃɛɛ̄l 'squinting' (MHG *schël*, MDu *schele*)
kaāl 'bald (*callow*)' (MHG *kal*, MDu *kale*)
vaāl 'pale (*fallow*)' (MHG *fal*, MDu *vale*)

If these words ended in a vowel in Early MLb (*ɣæl, *ʃæl, *kal, *val), these must be cases of analogical lengthening. If these words had a schwa in early MLb (*ɣælə , *ʃælə , *kalə , *valə), their long vowels must be due to Open Syllable Lengthening (*ɣɛɛ̄lə , *ʃɛɛ̄lə , *kaālə , *vaālə), in which case the present circumflex accent suggests that the final schwa dropt early (see §7.1).

Analogical lengthening did not happen in Limburgian and Central Franconian alone. It generally happened in 'standard' Middle High German as well, although it was restricted there to cases in which the consonant was voiced in the oblique forms, i.e. in the cases of (38), (39), and (40) (NHG *Tag* 'day', *Weg* 'way', *Gras* 'grass', *schmal* 'narrow'), but not in the voiceless cases of (37) (NHG *Dach* 'roof', *Schloss* 'lock'). In Dutch, no analogical lengthening occurred (except perhaps in *noot* 'nut'): it still has a **dax** ~ **daaɣən** alternation. The Franconian dialects are thus the only ones with lengthened vowels in **naāt** and **daāk**. And since Dutch adjectives were levelled towards the short vowel (**naatə** 'wet-PL' turned into **nɑtə** analogously to the predicative form **nɑt** 'wet'), the Franconian dialects are also the only ones with lengthening in Limburgian **naātə** 'wet-PL' or Ripuaric **naāsə** 'wet-PL'.

The fact that lengthening in consonant-final syllables was analogical is proved by the existence of an unlengthened vowel in the cases where there was no paradigm or where the paradigm did not contain lengthened vowels. Examples of unlengthened vowels outside the paradigms can be found in prepositions. Limburgian still has short vowels in **eñ** 'in', **vɑñ** 'of', **met** 'with', whereas we find a regularly lengthened vowel in **doōr** 'through' (< *thuru*). Other examples are **dɑt** 'that' and especially the content word **wɛx** 'away', which was apparently felt to be outside the paradigm of **wɛɛ̄x** 'way'

(this word has a short vowel in NHG *weg* as well). Examples of unlengthened vowels inside paradigms are those that were originally followed by geminate consonants, as in **mɑn** 'man', whose plural was **mɛnnə** 'men' (these examples are less convincing than the former, though, since it cannot be completely excluded that the pronunciation was **mɑnn**).

6.2. The new situation: lexical contrast on monosyllables with long non-high vowels

Analogical lengthening led to a large reduction in the correlations between tonal structure, syllable structure, and segmental structure. Before the change, the tone in monosyllables with long vowels could be predicted from vowel quality: high vowels and diphthongs had a circumflex, non-high vowels an acute. After the change, tone has become contrastive in monosyllables. Thus, the new form **ʃlɔɔ̄t** 'lock' contrasts with the old form **rɔ́ɔt̲** 'advice', and **ʃmeē̲t̲** 'smith' contrasts with **léet̲** 'song'.

6.3. Implications for dating the Rule A/A2 split

In the Rule A dialects, all forms in (24), (28) and (32) have an acute rather than the circumflex that is characteristic of the Rule A2 dialects under discussion here. Many authors have tried to explain the history of Rule A without assuming an intermediate Rule A2 stage. But if Rule A developed directly as a result of bimoraic reanalysis, the disyllabic forms on which the monosyllabic forms in (38) and (39) are based must have had an acute accent in the Rule A area at the moment when analogical lengthening occurred. In that case, the forms in (38) and (39) should have an acute accent in the Rule A area. They have not. They currently all have a circumflex, even in the Rule A area. This is evidence against an early split between between Rule A2 and Rule A, and evidence in favour of the theory (also assumed by De Vaan 1999) that the present Rule A dialects went through an earlier A2 stage.

6.4. Dating analogical lengthening

Analogical lengthening most probably started some time after Open Syllable Lengthening. In Van der Meer (1949) one can see that in the earliest records of Venlo MLb (the year 1320) long vowels in open syllables can be written with a single vowel symbol. Thus, *jaer* 'year-NOM' but *jare* 'year-DAT'. The writing therefore coalesces with that of originally short vowels in open syllables, e.g. *maken* 'make'. If we assume that the vowel in *jare* was long, the vowel in *maken* must have been long as well. We could thus place Open Syllable Lengthening in the 13th century at the last, i.e. not long after the start of Middle Limburgian, which was characterized by the coalescence of unstressed short vowels into schwa.

Analogical Lengthening in Venlo MLb happened in the 15th century, with some early forms in *haoff* (i.e. **ɦɔɔf**) and *daegh* (i.e. **daax**).

7. Schwa drop

There have been several rounds of the drop of schwa in final unstressed syllables, and these rounds had different effects on the tone system.

7.1. Early schwa drop

The earliest schwa drop was interconsonantal, and it occurred even before Open Syllable Lengthening. It happened in front of the final **t** of the second and third person

singular in the present indicative of strong verbs. Thus *__brekəs__ > *__breks__ 'break-2SG' and *__brekət__ > *__brekt__ 'break-3SG'.[18] This earliest schwa drop prevented the application of Open Syllable Lengthening, which affected most other forms in the paradigm such as the first person singular (*__brɛkə__ > *__brɛ̄kə__), the first and third person plural (*__brɛkən__ > *__brɛ̄kən__), or indeed the second person plural (*__brɛkət__ > *__brɛ̄kət__). This gives us the present-day paradigm __brɛ̄k breks brek brɛ̄kən brɛ̄k(t) brɛ̄kən__.

The earliest schwa drop just discussed was restricted to the second and third person singular of strong verbs ending in obstruents. The second schwa drop, which is more important to the story of the origin of tone in Limburgian, applied after Open Syllable Lengthening, and occurred when the schwa was preceded by a long vowel and a sonorant consonant but followed by a coronal obstruent (__s__, __z__, __d__, or __t__). The conspicuous characteristic of this second schwa drop (and what distinguishes it from the later general schwa drop discussed in §7.3) is that it had no effect on tone.

Unlike the morphologically limited first schwa drop, the second schwa drop occurred across the board, for instance in the second and third person singular of weak as well as strong verbs and in the past participle of weak verbs (the past participle of strong verbs did not end in -__t__). Thus, we crucially see that the cicrumflex on lengthened vowels remained in the following present-day Limburgian forms, the infinitives of which were mentioned in (21): __ɦɔ̄ls__ 'fetch-2SG', __ɦɔ̄lt__ 'fetch-3SG' and __ɣəɦɔ̄lt__ 'fetched' from the weak verb __ɦɔ̄lən__, and in __vøø̄rs__ 'drive-2SG' and __vøø̄rt__ 'drive-3SG' from the strong verb __vaārən__. Other examples are the forms of the weak verbs __woōnən__ 'dwell' (__woōns__, __woōnt__, __ɣəwoōnt__) and __ɣəbøø̄rən__ 'happen' (__ɣəbøø̄rt__, __ɣəbøø̄rt__). Presumably (although the present-day tone cannot evidence it, as will become clear in §7.3), this second schwa drop also occurred in the acute verbs mentioned in (12), for instance __vǿørs__ 'lead-2Sg', __vǿørt__ 'lead-2Sg' and __ɣəvǿørt__ 'led' from the weak verb __vǿørən__.

Another set of forms that probably show this early second schwa drop are those with (usually sonorant) consonants flanked by two originally high vowels:

(41) *Early drop of a schwa that used to be a high vowel after a high vowel*

__zoōn__ 'son' (OHG *sunu*)
__døø̄r__ 'door' (OLF *duri*), cf. the NHG variant *Türe*

---

[18] The divergently high vowel in __breks__ and __brekt__ is either due to an earlier *i*-umlaut (**brekis* > **brikis*), or to the non-application (because of the following high vowel) of the even earlier West-Germanic *a*-umlaut ({ **brikis*, **brikan* } > { **brikis*, **brekan* }), or to both of these. Not surprisingly, this vowel is often identical to the vowel in the non-lengthened imperative singular (§6), e.g. in __et__ 'eat-IMPSG' ~ __ets__ 'eat-2SG', although the two can diverge as well, as in our earlier example __ɣef__ 'give-IMPSG' ~ __ɣøfs__ 'give-2SG'. The 2SG and 3SG forms have been subject to an extensive morphologization of phonological alternations. One alternation is due to an original *i*-umlaut and explains the phonologically regular alternation of __ʃlɔ́ɔpən__ 'sleep-INF' (< *__slɑɑpən__) ~ __ʃléaps__ 'sleep-2Sg' (< *__slææpɪs__). Another alternation is the regular long-short alternation of __brɛɛkən__ ~ __breks__, which was analogically extended to originally long vowels, as in __lɔūpən__ 'walk-INF' (< *__lɔupən__) ~ __løps__ 'walk-2Sg', which occurs instead of the perhaps regular __løips__. The third alternation is the regular rounding alternation of __zeēn__ 'see-INF' (< WGm **seo-* < **siu-* with *a*-umlaut) ~ __zyȳs__ 'see-2Sg' (< WGm **siu-* without *a*-umlaut), which was analogically extended to forms like __ɣøfs__ and __ʃlǿaps__ 'sleep-2SG' (an alternate form of __ʃléaps__). Thus, the present-day strong-verb 2SG and 3SG morphemes are synchronically characterized to various extents by umlaut, shortening, and rounding, instances of which can be seen in several examples in the text.

doōr ‘through’ (OLF *thuro*; Old Saxon *thuru*)
vøø̄r ‘for, before’ (OLF *furi*)

Some of these forms could be dismissed as cases of schwa drop before Open Syllable Lengthening, followed by analogical lengthening (e.g. **sunu* > ***zon* > **zoōn**). Quite apart from the need to explain such an earlier schwa drop, it would not account for the case of **doōr**, which has no source for analogy. De Vaan (1999) mentions some of the words in (41) separately as well as possible cases of an early schwa drop (namely ‘son’ and ‘door’, but also ‘sow’, which is **zoōx** in some dialects).[19]

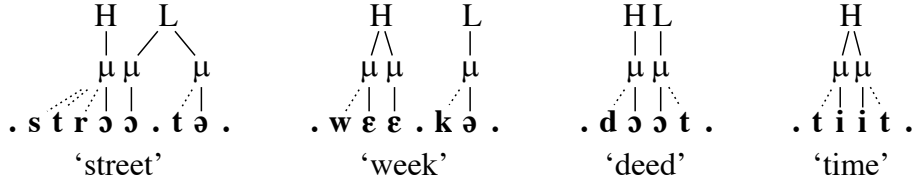
The idea that schwa dropped rather early between sonorants and coronals obstruents is confirmed with the data from Middle Dutch as transcribed in the dictionary by Verdam (1961), who allows forms like *geweget* ‘aimed’, *gevoedet* ‘fed’, and *gevollet* ‘filled’ (beside *geweecht*, *gevoet*, and *gevolt*) but not *gewonet* ‘resident’ or *gemalet* ‘painted’ (only *gewoont* and *gemaelt*).

7.2. Drop of final schwa: voiceless version

All words that ended in a schwa, or that ended in a schwa followed by an obstruent, eventually dropped it. The main exception is the plural and feminine singular of adjectives ending in a plosive or voiceless consonant cluster. Thus present-day Limburgian still has **naātə** ‘wet-FEMSG’, **riīkə** ‘rich-FEMSG’ and **ɣɑn̄sə** ‘whole-FEMSG’, but in nearly all other cases final schwa was lost.[20]

We know that the loss of schwa ultimately led to a change in phonological structures across generations: original **strɔ́ɔtə** ‘street’ and **dɔ́ɔt̲** ‘deed’ consisted of three and two moras, respectively, as shown in (7) and (14), while present-day **ʃtrɔ́ɔt** and **dɔ́ɔt̲** rhyme perfectly, i.e., they must consist of the same number of moras, which I will assume to be two. As in the account of Open Syllabe Lengthening, the account of Schwa Drop will involve three elements: the phonological structure maintained by the first generation (cf. (8)), the variation in pronunciation allowed by this structure, and the reinterpretation of selected overt forms as a new phonological structure by the next generation (cf. (19)). The lexical tonal representations for generation 1 of the acute form **strɔ́ɔtə** ‘street’ and the circumflex form **wɛɛ̄kə** ‘week’, aligned to an HL intonation contour, are given in (42), which also includes the structures for words without final schwa.

(42) *Surface structures of words with and without final schwa, voiceless version*



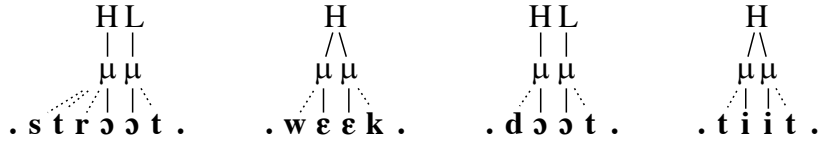
In these graphs syllable boundaries are represented as periods.

[19] Perhaps the word **aān** ‘at, on’ (< **anə**) lost its schwa in this round as well, perhaps in order to improve the distinction with ***ɑ́ɑnə** ‘without’ after Open Syllable Lengthening (***aānə**).
[20] Other scattered exceptions exist, for instance the word **néaʃə** ‘sewstress’ < ***nææjərsə**, cf. attested MLb (sermons) *minnerse* ‘female lover’.

Now that we have the surface structures of generation 1, the next question is how the structures were actually pronounced. For **dɔ̀ɔt** and **tiīt** we can assume the pronunciations [dɔ̀ɔ̀t] and [tíít] (in H*L-focus environments); in the absence of evidence to the contrary we can assume that the final consonants were fully released, as they are in present-day Limburgian. For **strɔ̀ɔtə** and **wɛ̄kə**, obviously possible pronunciations are [strɔ̀ɔ̀tə̀] and [wɛ́ɛ́kə̀], with a full final vowel. But in a situation of imminent schwa loss there must have been variant forms with a reduced final vowel, which after a voiceless consonant could well be voiceless: [strɔ̀ɔ̀tə̥̆] and [wɛ́ɛ́kə̥̆]. Note that in these forms the L tone is not audible on the final vowel, because this vowel is voiceless. Such situations, with vowels that are fully present in the phonological surface structure but are hardly audible in the overt form, can in some languages, like Japanese, apparently easily be maintained by speakers and transferred to the next generation.

But Limburgian is not Japanese. Learners of Japanese "only" have to know that word-final consonants are generally ruled out in their language, as is evidenced by the full vowels that have to occur after voiced obstruents and sonorants. Limburgian learners, however, had ample evidence that Limburgian did allow word-final consonants of all sorts. In a situation where the learner interprets [dɔ̀ɔ̀t] as the corresponding structure in (42), she also has to interpret [strɔ̀ɔ̀tə̥̆] in a way that makes sense in her emergent phonology. While some learners may indeed come up with the bisyllabic structure in (42), perhaps aided by some overt adult forms with full vowels, some learners will act differently. We have to understand that the transcription [strɔ̀ɔ̀tə̥̆] suggests a final vowel, but that a voiceless vowel, especially a short one, often cannot be distinguished from the glottal fricative [h], so that an almost equivalent transcription is [strɔ̀ɔ̀th]. Since utterance-final plosives tend to be slightly aspirated anyway (by the continuation of uninhibited expiration), so that **dɔ̀ɔt** will sound very much like [dɔ̀ɔ̀t^{h}], we have the ingredients for a merger, especially since word-final aspiration contrasts are cross-linguistically extremely rare. Some children, therefore, will perceive [strɔ̀ɔ̀tə̥̆] and [wɛ́ɛ́kə̥̆] as the monosyllabic structures in (43).

(43) *Surface structures after schwa drop, voiceless version*



For **strɔ̀ɔt** the change was well integrated in Limburgian phonotactics: the resulting rhyme was identical with the already existing rhyme in **dɔ̀ɔt**. For **wɛ̄k**, the change was also reasonably well integrated, because the vowel **ɛɛ̄** was already allowed in monosyllabic words deriving from analogical lengthening (note that the MLb ***ɛɛ** had become **éa**). Now if some children create the structures in (43), they may go on to learn at a later point that some of these words can have a schwa appended to them. In effect, the structures have changed across the two generations from bisyllabic with variable schwa drop to monosyllabic with variable schwa extension (perhaps like modern French **sɛ̃c** 'five'). It is likely that speakers with the structures in (43) produce forms with schwa less often than people who still have the bisyllabic structures. The third generation, then, will hear even fewer forms with schwa than the second and is thus

even more likely to posit a monosyllabic analysis, leading to even more schwa-less pronunciations. The merger, once started, appears irreversible, at least if schwa-less structures are fully licit in the language to begin with (this excludes the Japanese case).

If the consonant was voiceless, therefore, the lexical tone on the newly created monosyllabic word was identical to the lexical tone on the original bisyllabic word. Thus, originally long non-high vowels retained the acute. We can see this in many present-day Limburgian forms, which usually still end in *-e* in German. These forms include the first person singular of the present tense of verbs, the singular of many nouns (mostly feminine; if masculine, they usually end in *-en* in present-day German)[21], and the plural of some nouns. The following list presents up to four forms of each word mentioned: the current East Limburgian form, the reconstructed form just before schwa drop (but after open syllable lengthening), the reconstructed early MLb form, and sometimes the attested or reconstructed OLF form.

(44) *Schwa drop after acute plus voiceless consonant: acute*

Early MLb **ɑɑ**: **ʃtrɔ́ɔt** f. 'street' < ***strɑɑtə** (NHG *Straße*), **mɔ́ɔt** f. 'measure' (MHG *mâʒe*), **ɦɔ́ɔk** m. 'hook' (NHG *Haken*), **ʃlɔ́ɔp** 'sleep-1SG', **lɔ́ɔt** 'let-1SG'

Early MLb **ee**: **téek** f. 'cover' (< Lat. *thēca*; NHG *Zieche*)

Early MLb **øø**: **zǿøk** 'seek-1SG' < ***zøøkə** (NHG *suche*), **vǿøt** 'feet', Rm. **bøø̄k** 'beech', **zǿøt** 'sweet' (OLF *suoti*), **wǿøs** 'savage' < **wøøstə** (OLF *wuosti*)

Early MLb **oo**: **kóok** m. 'cake' < ***kookə** (NHG *Kuchen*)

And high vowels, diphthongs, and lengthened vowels retained the circumflex:

(45) *Schwa drop after circumflex plus voiceless consonant: circumflex (for nonhomorganic cluster cases see §12.2)*

Early MLb high vowels and diphthongs: **riik** n. 'empire' < ***riikə** (OLF *rîki*), **riik** 'rich', **liik** n. 'dead body' (NHG *Leiche*), **ruūt** 'diamond', **zɛīp** 'soap' (NHG *Seife*), **ɛīk** 'oak' (NHG *Eiche*)

Early MLb **an**: **kɑɲc** 'side' (NHG *Kante*)

Early MLb **om**: **klom̄p** 'lump' (NHG *Klumpen*), **ɦom̄p** 'chunk' (NHG *Humpen*), **lom̄p** 'rag' (NHG *Lumpen*)

Early MLb **ɔl**: **bɔūt** 'bolt' < ***bɔltə** (NHG *Bolzen*)

Early MLb **a**: **zaāk** f. 'business' < ***zaākə** < ***zakə** (OLF *saka*, NHG *Sache*), **aāp** m. 'ape' (NHG *Affe*), **ɦaāt** m. 'hate', **laāt** 'late'

Early MLb **ɛ**: **bɛɛ̄k** 'brook' < ***bɛɛ̄kə** < ***bɛkə** < **baki* (cf. umlautless NHG *Bach*)

[21] Both the feminine and masculine nouns had a nominative in *-e* and an accusative in *-en* in early MHG. The feminine nouns soon changed their accusative into *-e*, generalizing the nominative affix. Most masculine nouns later changed their nominative to *-en*, generalizing the accusative. Some masculine nouns (e.g. *Name* 'name', *Wille* 'will', and nouns denoting males like *Junge* 'boy' and *Affe* 'ape') retain the *-e*/*-en* alternation in present-day German.

Early MLb **æ**: **wɛɛ̄k** ‘week’ < ***wɛɛ̄kə** < ***wækə** < **weka*, **brɛɛ̄k** ‘break-1SG’ < ***brækə*** (NHG *breche*), **ʃprɛɛ̄k** ‘speak-1SG’

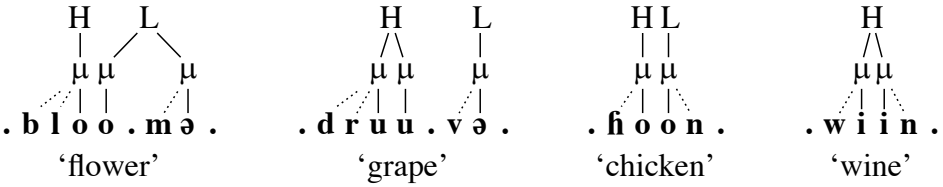
Early MLb **ɔ**: **knɔɔ̄k** m. ‘bone’ < ***knɔɔ̄kə** < ***knɔkə*** (NHG *Knochen*), **kɔɔ̄k** ‘cook-1Sg’ < ***kɔkə*** (NHG *koche*), **ɦɔɔ̄p** ‘hope-1Sg’ (NHG *hoffe*)

The new situation. Like analogical lengthening, schwa drop led to less predictability in the relation of the mora accent with segmental and syllable structure. Thus, the new form **ɦɔɔ̄p** ‘hope-N; hope-1SG’ contrasts with the old form **slɔ́ɔp** ‘sleep-N; sleep-IMPSG’ and with the new form **slɔ́ɔp** ‘sleep-1SG’.

7.3. Drop of final schwa: voiced version

Something slightly different happened when the intervening consonant was voiced rather than voiceless. The structures of generation 1 are shown in (46).

(46) *Surface structures of words with and without final schwa, voiced version*



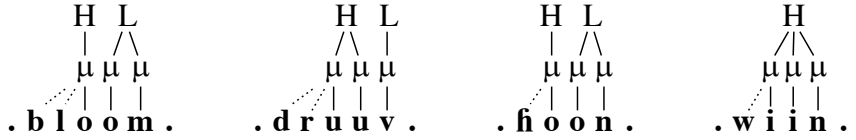
These structures are completely analogous to (42). Again, the tones are only associated with the heads of the tonal moras, and onset segments cannot head moras. As before, there cannot be more than two moras in a syllable, simply because no more than two tonal moras are needed to represent the surface structure. This means that the final **n** of **ɦóon** or **wiin** does not head a mora (this contrasts with cases like **mɑn̄** ‘man’, in which the final **n** is the second tonal mora of the syllable).

The next question is how these abstract structures were pronounced. I assume that every segment belongs to a mora, as in (46), and that every segment is realized with the tone that its mora head is linked to. Thus, the **n** of **ɦóon** or **wiin** can be regarded as a part of the second mora, hence will be pronounced with the tone linked to that mora: [ɦóòǹ] and [wííń]. For the full pronunciations of **blóomə** and **druūvə** I assume that the syllable boundary coincides with a mora boundary, so that the **m** and **v** can be regarded as a part of the mora in the second syllable, hence the pronunciations [blóòm̀ə] and [drúúv̀ə]. In a situation where final schwas are heavily reduced, some realizations must have sounded like [blóòm̀] and [drúúv̀]. The crucial difference with the voiceless forms of the previous section is that a low tone can now be heard on the final consonant. Forms like [drúúv̀] may look weird, with a low-toned final segment, but it is the same pronunciation that was judged as ‘trivial’ by Kiparsky (1973) for the Lithuanian word **dvaāras**, which came to be pronounced [dváár̀s] when the final vowel was dropt (for more details, see §7.4).

Some learners of generation 2 will base their surface structures on the reduced forms [blóòm̀], [drúúv̀], [ɦóòǹ] and [wííń]. A perception as disyllabic would violate a constraint against having a syllable that is not headed by a vowel. This structural constraint is cross-linguistically high ranked and we can assume that it is also high ranked for beginning learners. So the learner will assign monosyllabic structures to all

of the four forms (note that the four forms have analogous segmental patterns: [drúúv̀] looks different from the others since it ends in a fricative, but it stands in for a large number of forms that also include sonorant-final cases like the generation-1 word **duūmə** ‘thumb’, which has the reduced form [dúúm̀]). The forms show three overt tone patterns: an level high pattern in [wííń], a fall before the last segment in [drúúv̀], and a fall in the midst of the long vowel in [blóòm̀] and [ɦóòǹ]. The learner can make sense of this by positing a mora boundary at every change of tone. This requires that she posit the trimoraic analysis shown in (47).

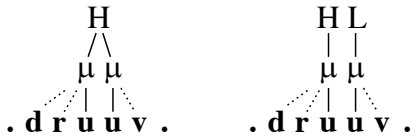
(47) *Trimoraic analysis of schwa loss, voiced version*



Some generation-2 learners may indeed come up with the structures in (47), thus violating a cross-linguistically high-ranked constraint against having more than two tonal moras in a syllable. The result is a triple contrast: the tone moves after the first, second, or third mora. Such a situation is uncommon in languages, like e.g. the triple contrast between oral, nasal, and half-nasal vowels, or the triple contrast between dentals, alveolars and retroflexes: not impossible, but uncommon, either for perceptual reasons (the auditory contrast is too small) or for structural reasons (here: three moras is too crowded). In Limburgian there were also systemic reasons that disfavoured these trimoraic forms. First, the triple contrast is specific to monosyllables and does not occur in disyllabic words. When an ending is added, the third mora is typically resyllabified as a non-moraic consonant, as in **.druū.vən**. ‘grapes’. These restrictions on the third mora are in stark contrast with the freedom of sonorant consonants as a *second* mora. Thus, **n** can be the second mora in the monosyllabic **mɑn̄** ‘man’, or when followed by a voiced plosive in MLb **lɑn̄t** ‘land’, or when in the disyllabic **.mɛn.nə.** ‘men’, where it stays in the first syllable (for the tone structure of this word, see the next section).

With the problematicity of the trimoraic analysis, several generation-2 learners will directly create non-trimoraic analyses from the [blóòm̀], [drúúv̀], [ɦóòǹ] and [wííń] of generation 1. And generation-2 learners that do create the forms in (47) will beget generation-3 learners, some of whom will create non-trimoraic analyses from the [blóòm̀], [drúúv̀], [ɦóòǹ] and [wííń] of this generation-2 subgroup. All these non-trimoraic analysers will come up with a bimoraic analysis, and the two moras will be the two parts of the long vowels in (46) and (47). For the three forms [blóòm̀], [ɦóòǹ] and [wííń] the analysis is easy, since it can be the same as for [ɦóòǹ] and [wííń] in (46). For [drúúv̀], the choice is simply between the two forms in (48).

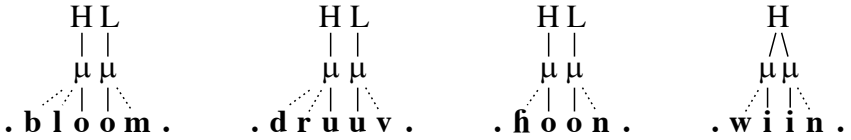
(48) *The two bimoraic options for the perception of* [drúúv̀]



The first of these forms honours the high tone audible on the second part of the vowel in [drúúv̀], but since the child must pronounce the first form in (48) as [drúúv́],

analogously to [wííń], it does not honour the low tone audible on the final consonant. The second form, on the other hand, does honour the audible low tone in [drúúv̀], but since it must be pronounced as [drúùv̀], analogously to [ɦóòǹ], it does so by severing the audible link between the high tone and the second part of the vowel in [drúúv̀]. The choice is between two evils, and what the Franconians did is clear: they chose the second option in (48), thus creating an acute accent on **drúuv** 'grape'. When we compare the new structure to the 'grape' versions in (46) and (47), we see that Limburgian has "preferred the stability of tones to the stability of segment-to-tone links". The square quotes are necessary here, because the same "stability preference" would predict the same circumflex-to-acute change in the case of the voiceless consonants of the previous section. Speaking in terms of "stability", i.e. speaking in terms of conversions between phonological structures, is of course incorrect for sound change, which must pass through the bottleneck of the overt-form-and-reanalysis scheme. The "voiceless" form **wɛɛ̄kə** has to pass through the overt form [wɛ́ɛ́k(ə̯)] and thereby loses the information about the low tone on the schwa in the surface structure. The four structures, then, are the ones shown in (49).

(49) *Surface structures after schwa drop, voiced version*



When comparing the voiceless and voiced forms, we see that only one of the eight forms changed the tonal contour on the first two moras: **drúuv**. Its structure has become indistinguishable from that of **ɦóon**, so that we must assume that the F0 contour in the pronunciation of **drúuv** as spoken by generation 3 fell in the middle of the vowel, so that we must conclude that there was a pronunication change from [drúúv̀] to [drúùv̀] within a couple of generations.

The seemingly weird accent alignment change is less exceptional than it may seem: for a comparable change in Lithuanian, see §7.4.

The change of the mora accent around voiced consonants has left its reflexes in present-day Limburgian. List (50) shows the present-day forms for original acutes, which stayed acute. The forms in the list do not take into account the effects of the second final devoicing (§11), so many forms end in a voiced obstruent in (50) whereas most of the present dialects have voiceless obstruents instead.

(50) *Modern forms with schwa drop after acute and voiced consonant: acute*

Early MLb **ɑɑ* > **ɔ́ɔ**: **lɔ́ɔɣ** f. 'layer' < ***lɑɑɣə** (NHG *Lage*), **mɔ́ɔn** f. 'moon' (OLF *mâno*), **ɣrɔ́ɔv** m. 'count'

Early MLb **ææ* > Gel. **éa**, Rm. Sitt. **ée**: **kéaz** f. 'cheese' < ***kææzə** (NHG *Käse*), **léaɣ** 'low' (OHG *lâgi*), **ʃéar** 'scissors' (NHG *Schere*)

Early MLb **ɛɛ* > Gel. **éa**, Rm. Sitt. **ée**: **zéal** 'soul' < ***zɛɛlə** < *sêla* < WGm **saiwala* (NHG *Seele*)

Early MLb **œœ* > Gel. **øa**, Rm. Sitt. **øø**: **drøaɣ** 'dry' < ***drœœɣə**

Early MLb ***ɔɔ** > Gel. **óa**, Rm. Sitt. **óo**: **bóan** f. 'bean' < ***bɔɔnə** (NHG *Bohne*), **ʃóan** 'beautiful', **óar** n. 'ear', **dóaz** f. 'box' (NHG *Dose*), **ʃóal** f. 'school'

Early MLb ***ee** > Gel. Rm. **ée**, Sitt. **éi**: **vléeɣ** f. 'fly' (NHG *Fliege*)

Early MLb ***øø** > Gel. Rm. **øø**, Sitt. **œi**: **ɣrøøn** 'green' < ***ɣrøønə** (OHG *gruoni*), **køøl** 'cool', **vrøøɣ** 'early'

Early MLb ***oo** > Gel. Rm. **óo**, Sitt. **ɔu**: **blóom** f. 'flower' < ***bloomə** (NHG *Blume*), **ɦóor** f. 'whore'

The **drúuv** case extends to all high vowels, diphthongs, and lengthened vowels. In all these cases, an original circumflex shows up as an acute in the present language if schwa was dropped after a voiced consonant:

(51) *Modern forms with schwa drop after circumflex and voiced consonant: acute (for nonhomorganic cases see §12.2)*

Early MLb ***ii** > ***iī** > **íi**: **wíiz** f. 'melody' (NHG *Weise*), **líin** f. 'line' (NHG *Leine*), **píin** 'pain' (Late Latin *pēna*), **ɣərdíin** 'curtain', **ʃíiv** f. 'disc' (NHG *Scheibe*)

Early MLb ***yy** > ***yȳ** > **ýy**: **krýyn** 'crown of head' < ***kryȳnə** < ***kryynə**

Early MLb ***uu** > ***uū** > **úu**: **kúul** 'mine' < ***kuūlə** 'pit' < ***kuulə**, **drúuv** f. 'grape' (NHG *Traube*), **dúuv** f. 'pigeon' (NHG *Taube*), **dúum** m. 'thumb' (NHG *Daumen*)

Early MLb ***ɛi** > ***ɛī** > **éi**: **klɛ́in** 'small' < ***klɛīnə**, **ɑlɛ́in** 'alone' (NHG *alleine*), **ɣəmɛ́in** 'nasty' (MHG *gemeine*)

Early MLb ***ɔu** > ***ɔū** > ***ɔ́u**: **ɔ́uɣ** n. 'eye' < ***ɔūɣə** < ***ɔuɣə** (NHG *Auge*), **ɣəlɔ́uv** n. 'belief' (NHG *Glaube*)

Early MLb ***a** > ***aā** > **áa**: **náam** m. 'name' < ***naāmə** < ***nāmə** < ***namə** < *namo*, **báan** f. 'road', **ʃwáan** f. 'swan' < ***swanə**, **ráam** m. 'frame' (NHG *Rahmen*), **náaz** f. 'nose', **ráav** m. 'raven' (NHG *Raben*), **máaɣ** m. 'stomach' < ***maāɣə** < ***maɣə** (NHG *Magen*), **máax(t)** f. 'virgin' < ***máaxt** < ***maāɣət** < ***maɣət** (NHG *Magd*), **ʃáaj** f. 'damage' < ***sxaðə** (NHG *Schade* f., *Schaden* m.), **kráaɣ** m. 'collar' (NHG *Kragen*), **dráaɣ** 'carry-1SG' (NHG *trage*), Rm. (also in Belgian Limburg; Goossens 1977) **káaf** n. 'chaff' < ***kavə**[22]

Early MLb ***æ** > ***ɛɛ̄** > **ɛ́ɛ**: **nɛ́ɛv** 'nephew, cousin' < ***nævə**, **wɛ́ɛm** 'whom' < ***wæmə**, **lɛ́ɛv** 'live-1SG' < ***lævə**

Early MLb ***ɛ** > ***ɛɛ̄** > **ɛ́ɛ**: **zɛ́ɛɣ** f. 'saw' < ***zɛɣə**

Early MLb ***ɔ** > ***ɔɔ̄** > **ɔ́ɔ**: **kɔ́ɔl** f. 'coal' < ***kɔɔ̄lə** < ***kɔlə** (NHG *Kohle*), **bɔɔ̄r** f. 'drill' (NHG *Bohre*), **bɔ́ɔɣ** m. 'bow' < ***bɔɣə** (NHG *Bogen*)

[22] Reported mediaeval forms from surrounding languages of the word meaning 'chaff' are *kaf* and *kave*. The present-day Dutch form *kaf*, with a short vowel, must come from *kaf*. As Goossens notices, the acute accent shows that the present-day Limburgian form comes from *kave* (if it had come from *kaf*, it would have become **kaāf** by analogical lengthening). Of course, this form cannot be an argument for the historical predictability of the present-day tone; rather, the tone disambiguates the mediaeval form in this case.

Early MLb *e > *eē > **ée** (Sitt. **éi**): **ʃméej** m.pl. ‘smiths’ < ***smeēðə** < ***smeðə** (NHG *Schmiede*), **ʃéem** m. ‘shade’ < ***sxemə** (Old Saxon *scimo*; different vowel NHG *Schemen*)

Early MLb *o > *oō > **óo** (Sitt. **ɔ́u**): **ʃtóov** f. ‘stove’ < ***stoōvə** < ***stovə** (NHG *Stube*)

The new phonology. The changes proposed in this section only work if the language can be regarded as a lexical mora-alignment language both before and after the change. Schwa drop after voiced consonants greatly extended the number of forms in which mora alignment could be contrastive. The acute accent could now appear on high vowels, on diphthongs (including short vowels followed by sonorant consonants), and on the long vowels **aa** and **ɛɛ**. It may be possible that schwa drop in this way saved the mora accent contrast in the Franconian dialects, whereas the neighbouring High-German dialects lost the contrast because they retained final schwa.

Reflexes in present-day phonology. Schwa drop led to a large number of tone alternations within paradigms where the consonant was voiced and the vowel originally had a circumflex tone. In weak verbs, the first singular could receive an acute, while the infinitive retained the circumflex:

lɛ́ɛv ‘live-1SG; live-IMPSG’ ~ **lɛɛ̄vən** ‘to live’

In strong verbs, where the imperative singular ended in a consonant, some minimal pairs were introduced:

bliif ‘stay-IMPSG’ ~ **blíiv** ‘stay-1SG’ ~ **bliivən** ‘to stay’, **draāx̲** ‘carry-IMPSG’ ~ **dráaɣ** ‘carry-1SG’ ~ **draāɣən** ‘to carry’

In adjectives, the plural and feminine singular could form a minimal tone pair with the predicative form:

wiis ‘wise’ ~ **wíiz** ‘wise-PL&FEMSG’ ~ **wíizən** ‘wise-MASCSG’

This example also shows that the masculine singular received an unexpected acute, perhaps in analogy with the feminine and plural form (De Vaan 1999).

In nouns, we have to distinguish various classes. Consonant-final stems (mostly masculine) lost the schwa in the plural:

daāx̲ ‘day’ ~ **dáaɣ** ‘days’, **kniīn** ‘rabbit’ ~ **kníin** ‘rabbits’

Schwa-final stems (mostly feminine) lost the schwa in the singular:

líin ‘line’ ~ **liīnən** ‘lines’, **dúuv** ‘dove’ ~ **duūvən** ‘doves’, **dúum** m. ‘thumb’ ~ **duūmən** ‘thumbs’, **ʃrúuv** ‘screw’ ~ **ʃruūvən** ‘screws’, **drúuv** ‘grape’ ~ **druūvən** ‘grapes’, **ýyl** ‘owl’ ~ **yȳlən** ‘owls’, **rɛ́iz** ‘travel’ ~ **rɛīzən** ‘travels’, **ɔ́uɣ** n. ‘eye’ ~ **ɔūɣən** ‘eyes’, **náam** m. ‘name’ ~ **naāmən** ‘names’, **báan** ‘job’ ~ **baānən** ‘jobs’, **kɔ́ɔl** ‘coal’ ~ **kɔɔ̄lən** ‘pieces of coal’

Later on, these alternations led to localized cases of analogy, with the normal exceptions in (12), such as Roermond **blóom** ~ **blóomən̲**, coming to follow the **drúuv** ~ **druūvən** pattern (Dols 1944). Another thing that arose locally was purely

morphological use of tone alternations, such as Moresnet **ʃlɔ̄ɔ̄p** ‘sleep-IMPSG’ ~ **ʃlɔ́ɔp** ‘sleep-1SG’, coming to follow the **bliīf** ~ **blíiv** pattern (footnote 9).

7.4. Lithuanian parallels of tone change by schwa drop

This seemingly weird change, which has attracted the ingenuity of many researchers (§13), has a parallel in Lithuanian, where the form [dváárs] that we saw before is interpreted as an acute accent: **dváars** (Kiparsky 1973). The following makes the parallel explicit:

(52) *Parallel between Lithuanian and Limburgian: schwa drop creates a fake acute*

dvaāras	→	[dváárs]	→	**dváars**
druūvə	→	[drúúv]	→	**drúuv**

Kiparsky regards **dváars** as a ‘fake acute’ because he claims that it is pronounced as [dváárs] rather than [dváàrs]. If the pronunciation is indeed [dváárs], Lithuanian would reflect a stage that has been intermediate in the development of Limburgian.

The parallel in (52) also resolves a different problem in Lithuanian that still eluded Kiparsky (1973: 830: fn. 22). In Lithuanian nominal inflection, all disyllabic endings used to have a circumflex in the first syllable, and this shifted later into a short accent on the second syllable, e.g. in the instrumental plural **-oōmis** > **-oomìs**. The dative plural form, which can be reconstructed as **-oōmus**, escaped this shift because the **u** was dropped at some point. However, instead of the expected **-oōms**, this form presently turns up as **-óoms**. This form is understandable as a perfect parallel to Limburgian schwa drop: in a language that already has a Limburgian/Lithuanian-type tone contrast, a circumflex will turn into an acute if a schwa drops from the following syllable, at least if the intervening consonant is voiced:

(53) *Parallel between Lithuanian and Limburgian: schwa drop creates a real acute*

druūvə	→	[drúúv]	→	**drúuv**	→	[drúùv]
oōmus	→	[óóm̀s]	→	**óoms**	→	[óòm̀s]

In (53) we see three steps: a speaker-induced drop of a final vowel, a reinterpretation by listeners as an acute (i.e. as a low tone connected to the second mora), and a new pronunciation based on the reinterpretation (i.e. a pronunciation with low F0 on the second part of the vowel).

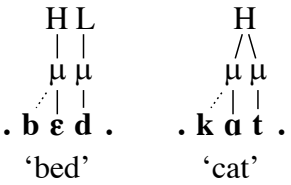
7.5. The behaviour of obstruent geminates in schwa drop

After a voiceless geminate consonant such as in ***kɑttə** ‘cat’, a dropping schwa would leave no trace of the low F0, and there is no reason why the resulting auditory form [kɑ́tt(ə̯)] should be perceived as anything else than two high-toned moras, just as the auditory form [stǿk] (I assume here that a trimoraic analysis is ruled out, so that the final consonant cannot be analysed as a geminate).[23] After a voiced geminate obstruent

[23] This opens up the question whether a structure like $((\mathbf{kɑt})_{\mu}(\mathbf{t})_{\mu})_{\sigma}$ is universally ruled out or not. If it is not, then bimoraicity is compatible with syllable-final geminates, and my argument for assigning a mora to the first **d** in ***bɛddə** (§4.7) would vanish as well, because the structure could then be $((\mathbf{bɛd})_{\mu})_{\sigma}((\mathbf{də})_{\mu})_{\sigma}$.

such as in ***bɛ́ddə*** ‘bed’, a dropping schwa would still leave a trace of the low F0 on the final voiced consonant, and given the analysis in §7.3 there seems to be no reason for the learner to interpret the new auditory form [bɛ́d̀] as anything else than an acute. The structures in (54) show the results.

(54) *Surface structures after schwa drop, obstruent version*



Present-day reflexes. In the areas that maintain a lexical tone contrast on short vowels before voiceless consonants, such as Moresnet (Jongen 1972), words like **kɑt̄** currently have a circumflex, i.e. the pitch movement occurs in the next syllable, and words like **bɛ́d** currently have an acute, i.e. the pitch movement occurs in the syllable itself (instead of between **ɛ** and **d**, the pitch movement nowadays falls within **ɛ** as a result of the second final devoicing and/or degemination, see §11). Some present-day forms are listed in (55) and (56).

(55) *Schwa drop after voiceless geminates*

> **nɑk̄** (Rm. **næk̄**) m. ‘neck’ < ***nɑk̄kə*** (NHG *Nacken*), **plæk̄** m. ‘spot’ < ***plæk̄kə*** (MDu. *plecke*), **ɦɑk̄** m. ‘heel’ < ***ɦɑk̄kə*** (NHG *Hacke* f.), **tɑp̄** m. ‘tap’ < ***tɑp̄pə*** (NHG *Zapfen*), **lɑp̄** m. ‘piece of cloth’ < ***lɑp̄pə*** (NHG *Lappen*)

(56) *Schwa drop after voiced geminate obstruents*

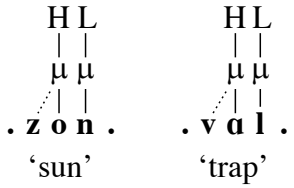
> **bɛ́d** n. ‘bed’ < ***bɛ́ddə***, **brœ́g** f. ‘bridge’ < ***brœ́ggə*** (NHG *Brücke*), **lɛ́g** ‘lie-1SG’ < ***lɛ́ggə*** (MDu *ligge*), **lɛ́g** ‘lay-1SG’ < ***lɛ́ggə*** (MDu *legge*), **zɛ́g** ‘say-1SG’ < ***zɛ́ggə*** (MDu *segge*), **wɛ́g** ‘loaf’ < ***wɛ́ggə*** (MHG *wecke*), **ɦɛ́g** ‘hedge’ < ***ɦɛ́ggə*** (NHG *Hecke*), **ɦœ́b** ‘have-1SG’ < ***ɦœ́bbə*** (MDu *hebbe*)[24]

7.6. The behaviour of sonorant geminates in schwa drop

After a geminate sonorant, schwa drop led to a new kind of acute syllables. Thus, the word ***zónnə*** ‘sun’ was pronounced as [zónǹə̀] before schwa drop (§4.7) and as [zóǹ] after schwa drop. In nearly all of the Limburgian and Central Franconian area (with the exception at least of the peripheral area around Horst, to the north of Venlo), the audibility of the low pitch (in H*L alignments) on the sonorant consonant was apparently enough to make the learners analyse [zóǹ] as the structure in (62).

[24] This word may not be actually attested as having an acute, because it is possible that all the areas in which short vowels before obstruent consonants have contrastive tone have the Ripuaric form **ɦɑ́n** instead (e.g. Jongen 1972).

(57) *Surface structures after schwa drop, sonorant version*



Present-day reflexes. In words with schwa drop after a sonorant geminate, an acute is found in nearly all of the Limburgian and Central Franconian area, with the exception at least of the peripheral area around Horst, to the north of Venlo, where no tone contrast exists on this type of syllables. Present-day forms typical of Geleen, Roermond and Sittard are listed in (58).

(58) *Schwa drop after geminate sonorants*

> **vɑ́l** f. ‘trap’ < **vállə** (NHG *Falle*), **vɑ́l** ‘fall-1SG’ < **vállə** (NHG *falle*), **kɑ́l** ‘speak-1SG’ < **kállə**, **ɦǽl** ‘hard-PL/FEMSG’ < *ɦǽllə (NHG *helle*), **wɔ́l** (Rm. **wól**) f. ‘wool’ < **wɔ́llə** (NHG *Wolle*, Low Saxon *Wulle*), **vɔ́l** f. ‘full-PL/FEMSG’ < **vɔ́llə** (NHG *volle*), **ʃtél** ‘silent’ < ***stéllə**, **vǽr** ‘far’ < ***vǽrrə** (NHG *ferne*), **kǽr** f. ‘car’ < ***kǽrrə** (NHG *Karre*, *Karren* m.), **tǽr** m. ‘tar’ < ***tǽrrə**, **zón** f. ‘sun’ < ***zónnə** (MHG *sunne*), **tón** f. ‘barrel’ < ***tónnə** (NHG *Tonne*), **mǽm** f. ‘female breast’ < ***mǽmmə** (MHG *memme*), Sitt. **ʃtœ́m** Rm. **ʃtøm** f. ‘voice’ < ***stɛ́mmə/*stémmə** (NHG *Stimme*)

The new phonotactics. This change led to a new kind of lexical tone contrasts, namely in final sequences of a short vowel and a sonorant consonant, i.e. between the forms listed in (31) and those listed in (58). One of the many minimal pairs is **vɑ́l** ‘trap; fall-1SG’ versus **vɑ̄l** ‘fall’.

7.7. Geminate lenition and schwa drop

Original sequences of a lateral or a nasal followed by a voiced homorganic plosive were at some point simplified to a geminate. All the later developments of this class of consonant clusters are identical to the developments in the original sonorant geminates. We can most easily account for this fact if we assume that this change occurred before the advent of the tone contrast. For instance, ***bɑŋgə** ‘scared’ > ***bɑŋŋə** > ***bɑ́ŋŋə** > ***bɑ́ŋ**. Another possibility is that the gemination occurred later; in that case, however, there must have been a tone change when gemination happened: ***bɑŋgə** > ***bɑŋ̄gə** > ***bɑ́ŋŋə** > ***bɑ́ŋ**. This question of this timing is not unlikely to be quite relevant to the whole question of Limburgian tonogenesis.

Present-day reflexes. The present Limburgian forms that derive from a schwa drop after a lenited geminate all have the acute tone.

(59) *Schwa drop in geminate lenition environments*

From *nd: **mą́ɲ** f. 'basket' < ***mą́ɲɲə** < ***mąɲɟə** (MHG *mande*), **ɦǽɲ** f.pl. 'hands' < ***ɦɛ́ɲɲə** < ***ɦɛɲɟə** (MHG *Hände*)[25], **tǽɲ** f.pl. 'teeth' < ***tɛ́ɲɲə** < ***tɛɲɟə** (MHG *Zähne*)[26], **wóɲ** f. 'wound' < ***wóɲɲə** < ***woɲɟə** (NHG *Wunde*), **ɦóɲ** or **ɦøɲ** m.pl. 'dogs' < ***ɦóɲɲə** < ***ɦoɲɟə** (NHG *Hunde*), **véɲ** 'find-1SG' < ***véɲɲə** < ***veɲɟə** (NHG *finde*), **véɲs** 'find-2SG' < ***véɲɲəs** < ***veɲɟəs** (NHG *findest*), **véɲc** 'find-3SG' < ***véɲɲət** < ***veɲɟət** (NHG *findet*)

From *ŋg: **báŋ** 'scared' < ***báŋŋə** < ***bɑŋgə** (NHG *bange*), **láŋ** 'long (adv.)' < ***láŋŋə** < ***lɑŋgə** (MHG *lange*), **ʃpréŋ** 'jump-1SG' < ***spréŋŋə** < ***spreŋgə** (MHG *springe*), **ʃpréŋs** 'jump-2SG' < ***spréŋŋəs** < ***spreŋgəs** (MHG *springest*), **ʃpréŋk** 'jump-3SG' < ***spréŋŋət** < ***spreŋgət** (MHG *springet*), **tóŋ** 'tongue' < ***tóŋŋə** < ***toŋgə** (NHG *Zunge*)

From *mb: **dóm** 'dumb-PL/FEMSG' < ***dómmə** < ***dombə** (MHG *tumbe*)

From *ld: **wéʌ** 'wild-PL/FEMSG' < ***wéʌʌə** < ***weʌɟə** (NHG *wilde*)

The new situation. This change led to an extension of the oppositions for forms ending in **m**: **dóm** (< ***dombə**) joins **ʃtóm** (< ***stommə**) in contrasting with **ʃtom̄**. More interestingly, a new contrast arose between forms like **ʃpréŋk** (< ***spreŋgət**) and **reŋ̄k** (< ***reŋk**), leading to some minimal pairs such as **zéŋk** 'sings' (< ***zeŋgət**) versus **zeŋ̄k** 'sinks' (< ***zeŋkət**). Also, a new correlation between segments and tones arose in the language: all words ending in a short vowel followed by **ɲ**, **ŋ**, or **ʌ** received (and still have) an acute tone; this is ultimately due to the fact that all such forms derive from words ending in a schwa (-**ɲɟə**, -**ŋgə**, -**ʌɟə**), which again is due to the fact that **ɲ**, **ŋ**, or **ʌ** originally had to be licensed by a following plosive (in the case of **ɲ** and **ʌ** because palatalization required a plosive, and in the case of **ŋ** because **ŋ** was not a separate phoneme of West-Germanic and in fact Proto-Indo-European) and the plosive had to be voiced (and therefore followed by a vowel, because of the first final devoicing) in order to be deletable.

7.8. Schwa drop after a vowel

Early Middle Limburgian had no cases of short vowels followed by a schwa, so Late MLb had no cases of lengthened vowels followed by a schwa. The only circumflex vowels before schwa could have been the high vowels **iī**, **yȳ**, and **uū**. Thus, there must have been forms like ***biiə** 'bee' and ***buuə** 'build-1SG'. However, §4.6 has proposed that in Early MLb the high vowels were monomoraic, so the structures might have been better represented as ***bijə** and ***buwə**, but instead I propose that the actual structures were ***bijjə** and ***buwwə**, with geminated glides. The reason for this proposal is that in the general Late MLb bimoraic reanalysis of stressed syllables such structures would have become ***bíjjə** and ***búwwə**, with acutes that we still see today in the plurals ***bíjən** 'bees' and ***búwən** 'build-1PL'. The forms with a final schwa received a less surprisingly acute: **bíj** and **búw**.

[25] The present **æ** instead of the expected **ɛ** is an example of the new regularized umlaut.
[26] The present **æ** instead of the expected **ɛ** is an example of the new regularized umlaut.

Present-day reflexes. All words with a schwa drop after a vowel have an acute in present-day Limburgian. Some cases are listed in (60).

(60) *Schwa drop after a vowel: acute*

> **bíj** 'bee' < *__biiə__ (MDu *bije*), **vríj** 'free-PL/FEMSG' < *__vriiə__ (NHG *freie*), **dýj** 'push-1SG' < *__dyyə__, **nýj** 'new-PL/FEMSG' < *__nyyə__, **búw** 'build-1Sg' < *__buuə__ (NHG *baue*), **ʃúw** 'shy-PL/FEMSG' < *__sxuuə__ (NHG *scheue*)

The new situation. The change led to the first tone contrasts in final vowels, some minimal pairs being **bíj** 'bee' ~ **bij̄** 'by', **vríj** 'free-PL/FEMSG' ~ **vrij̄** 'free', **nýj** 'new-PL/FEMSG' ~ **nyj̄** 'new', and **ʃúw** 'shy-PL/FEMSG' ~ **ʃuw̄** 'shy'.

7.9. Schwa drop and ð loss

At some point in early New Limburgian, a **ð** was lost between a vowel and a schwa. The schwa was also dropped, and if the preceding syllable had a circumflex, this changed into an acute, which comes to no surprise.

Present-day reflexes. All present-day forms with word-internal schwa deletion have an acute.

(61) *Drop of d and schwa: acute*

> **ɣáarən** 'collect' < *__ɣaāðərən__ < *__ɣaðərən__, **éɛk** 'vinegar' < *__ɛēðək__ < *__ɛðək__, **léɛr** 'leather' < *__lɛēðər__ < *__lɛðər__, **wéɛr** 'weather' < *__wɛēðər__ < *__wɛðər__, **léɛx̲** 'empty' < *__lɛēðəx̲__ < *__lɛðəx̲__, **bɔ́ɔm** 'bottom' < *__bɔɔ̄ðəm__ < *__bɔðəm__, **bǿøl** 'hangman' < *__bøøðəl__ < *__bøðəl__, **zíj** 'silk' < *__ziiðə__ (NHG *Seide*), **zíj** 'side' < *__ziiðə__ (NHG *Seite*), **lýyk** 'Liege' < *__lyȳðək__, **ɔ́ɔr** f. 'vene' < *__ɑáɑðər__, **ɔ́ɔm** m. 'breath' < *__ɑáɑðəm__, **vóorən** 'fodder' < *__vóoðərən__

The new situation. Before this change, the lengthened vowels **aa** and **ɛɛ** and the high vowels **ii**, **yy** and **uu** could only have an acute in positions before a voiced consonant. While there are certainly statistical tendencies to that extent in present-day Limburgian, the list in (61) shows several new acutes for these vowels before a voiceless consonant (**éɛk**, **léɛx̲**, **lýyk**). The list also shows new acutes for these vowels in disyllabic forms (**ɣáarən**).

8. Neutralization in voiced environments

The change from circumflex to acute in voiced environments described in §7.3 was conditioned by a drop of schwa in the next syllable. In the larger part of the present tone area, however, this change was not conditioned by schwa drop: it occurred as well in words that stayed disyllabic, such as in **léɛvən** 'to live' and **káamər** 'room'. In this area, then, all disyllabic forms with voiced intervocalic consonants have an acute, and **bəlɔ́ɔvən** 'to promise', from *__bəlɔvən__, rhymes with **ʃtrɔ́ɔvən** 'to punish', from *__strɑɑvən__.

The area in which this neutralization took place is called the *Rule A* area and roughly comprises the Central Franconian area (Ripuaric, Moselle Franconian) as well as the Limburgian-Ripuaric transition area. The more conservative area, where

bəlɔɔ̄vən still does not rhyme with **ʃtrɔ́ɔvən**, is called the *Rule A2* area and roughly comprises the Low Franconian (Limburgian) area.

There are at least two ways in which the neutralization could have happened. The first is by starting from a paradigmatic analogy, then overgeneralizing phonologically: the form **lɛɛ̄vən** could have become **léɛvən** by analogy to the related form **léɛv** 'live-1SG', after which the number of disyllabic voiced circumflexes in the language was so low that the few non-alternating remaining ones, like **kaāmər**, were changed to acute as well.

The second possibility is that the cause of the neutralization is purely phonological. We know that there are neighbouring non-tonal dialects that invariably have an acute in all long vowels, thus **léɛvən** as well as **éɛtən**. It is not unlikely that some of these dialects have been tonal at some stage in their development. When compared to present-day Limburgian, these dialects must therefore have gone through a general innovation exemplified by the changes **lɛɛ̄vən** → **léɛvən** and **ɛɛ̄tən** → **éɛtən**. The question now is: if there were any intermediate dialects that changed circumflex to acute only in one of these two cases, what would such a dialect look like, i.e. would it change **lɛɛ̄vən** to **léɛvən** or would it prefer to change **ɛɛ̄tən** to **éɛtən**? My answer is that it would prefer to change **lɛɛ̄vən** to **léɛvən**, simply because the auditory forms [léɛ́ɛ̀ə̀ŋ] and [léɛ̀ɛ̀ə̀ŋ] are auditorily closer than [éɛ́tə̀ŋ] and [éɛ̀tə̀ŋ]: in the former pair, the HL alignment difference is one segment, in the latter pair it is between one and two segments.

9. The Rule A – Rule B distinction

Here I first repeat something that I wrote to Carlos Gussenhoven in March 1998, before Köhnlein (2013) showed that this story would not account for interrogative intonation.

Circumflex accents nowadays tend to be longer than acute accents, although (in the Rule A and Rule A2 regions) they derive from vowels that were shorter (high vowels tend to be shorter than non-high vowels, and the lengthened vowels were shorter than the originally long vowels before OSL). So the Rule A and A2 regions have historically undergone a *length reversal*. I hypothesize that this is because there is a universal tendency for circumflexes to be longer than acutes.

Where does this universal tendency come from? I propose that the direct cue for mora alignment is the question whether or not the syllable contains a pitch movement. Thus, if the syllable contains a pitch movement, the accent must be acute, and if the syllable stays on the same pitch, the accent must be circumflex. Now, movement (e.g. pitch movement) can be perceived even if it is fast, whereas detecting constancy (e.g. monotonicity) takes a while. Hence the Limburgian correlation between pitch movement and duration. Similar correlations are found in some Slavic languages (Van Wijk 1935, 1939).

I propose that in the Rule A and A2 areas the accent distinction was about mora alignment, so that the pitch movement was the main cue and the durations adapted themselves later according to the universal movement-duration correlation. In the Rule B area, however, something different happened: these dialects did not create a mora alignment distinction. Rather, the original duration difference persisted. When vowels lengthened in open syllables, the originally long vowels just became longer (push chain). The dialects thus created a three-way length contrast, and the pitch movement became a cue only later. This new secondary cue was added in line with the universal movement-duration correlation, hence the longest vowels (i.e. the original long vowels)

received a constant pitch (which sounds like a circumflex accent), and the middle-length vowels (i.e. the vowels lengthened in open syllable, plus the high vowels and the diphthongs) developed a moving pitch (which sounds like an acute accent).

In this way, the Rule B dialects ended up having an acute everywhere the Rule A dialects had a circumflex, and the reverse. The resulting cross-linguistic apparent reversal is remotely comparable to that between Old Greek and Lithuanian (Kiparsky 1973), and a bit more directly comparable to that between Lithuanian and Latvian.

Recently, Köhnlein (2011) has shed doubt on this account, on the basis of the observation that while the declarative pitch contours are diametrically opposite in Rule A/A2 and Rule B, the interrogative pitch contours are the same in both regions. My analysis above could therefore hold only if the declarative contour is original in the Rule B area and the interrogative contour is somehow secondary, for instance borrowed by language contact. Köhnlein (pp. 197–199), however, sees the interrogative contour as original and the declarative contour as secondary, arguing that all dialects went through a stage when there was no opposition between declarative and interrogative intonation, a situation that stills exists in Hasselt (West-Limburgian). A further development of this theory appears in Köhnlein (2013).

10. Morphological uses of the accent contrast

During the development of Limburgian, much of the inflectional paradigm of nouns and adjectives was lost. In nouns the paradigm became restricted to the singular-plural distinction; in this case, the current singular continues the original nominative singular and the current plural continues the original nominative plural (these nominatives were often, but not always, identical to the accusative forms). In adjectives the paradigm became restricted to the attributive form, the masculine singular, the feminine singular, the neuter singular, and the plural (which is identical to the feminine singular); in this case, the forms are again based on the original nominative, except for the masculine singular, which is based on the original accusative, e.g. present-day **naātən** 'wet-MASCSG' ~ and **naātə** 'wet-FEMSG'.

The reduction of inflexion led to the situation that many nouns now had only a single form. This includes neuter nouns with an original zero plural ending (***bɛīn** 'bone, leg-SG&PL, ***woōrt̲** 'word'-SG&PL, ***ɦɔɔ̄f̲** 'court-SG&PL', ***lɑnt̲** 'land-SG&PL') and mainly masculine and neuter nouns whose plurals came to be identical to the singular as a result of schwa drop (***ʃɔ́ɔp** 'sheep-SG&PL). In all cases, a new plural was constructed. In the case of back vowels followed by a voiceless consonant, the choice was the addition of umlaut in the plural: **ʃɔ́ɔp** 'sheep-SG'~ **ʃœœp** 'sheep-PL' (cf. NHG *Schafe* 'sheep-PL'). In neuter nouns, this change was often accompanied by the addition of -**ər**, as in **daākər** 'roofs'. This changes greatly added to the already existing use of umlaut in plurals not ending in -**ə(n)** and to the already existing use of -**ər** for neuter nouns. In the case of circumflex vowels followed by a voiced consonant, the tone changed to acute, as in the neuter **bɛīn** 'leg-SG' ~ **bɛ́in** 'leg-PL', completely analogously to the already existing masculine **ʃtɛīn** 'stone-SG' ~ **ʃtɛ́in** 'stone-PL'; if possible, this change is combined with umlaut and/or lenition: **ɦɔɔ̄f̲** m. 'court-SG' ~ **ɦœœv** 'court-PL'; **lɑŋ̄c** n. 'land-SG' ~ **læŋ̄** 'land-PL' (also **læŋ̄ər**); Rm. **kɑm̄p** m. 'comb-SG' ~ **kǽm** 'comb-PL'.

The chronology of these events probably varies widely. Given the fact that many of these changes also occurred in High German (*Beine* 'legs', *Dächer* 'roofs',

Worte/Wörter ‘words’, *Höfe* ‘courts’, *Länder* ‘lands’), it is possible that some of them precede even schwa drop. Thus, the sequence **bɛīn** ‘bones’ > **bɛīnə** > **bɛ́in** is not out of the question. Also, **wɔɔ̄rt̲** ‘words’ > **wœœ̄rdə** > **wǿœrd** is possible; the **d** would then have been based on a still existing dative singular or on compounds such as ***wɔɔ̄rdənbóok** ‘dictionary’, or on the diminutive (present-day **wǿœrtjə**, not **wœœ̄rtjə**). More examples will be seen in the next section.

11. Second final devoicing

After schwa drop, many words ended in a voiced obstruent. These include first person singular verb forms such as **blíiv** ‘stay’ and **zɛg** ‘say’, feminine singulars such as **drúuv** ‘grape’ and **ɦɛg** ‘hedge’, and masculine plurals such as **déev** ‘thieves’. In some regions, the voiced fricatives made it into the 20th century and some voiced plosives still exist (Dupont 1910–11, Leenen 1915, Goossens 1977). In most regions, however, all final obstruents were devoiced, turning e.g. **zɛg** into **zɛk̲** (where the underlining as usual denotes the fact that longer related forms still contain a **g**).

The second final devoicing must also have led to some new cases of homophony between singulars and plurals. Thus, **déef̲** ‘thief-SG’ ~ **déev** ‘thief-PL’ merged into **déef̲** ‘thief-SG&PL’. Just as in the previous section this merger gave rise to analogies that created new plurals. In the case of **déef̲** ‘thieves’, umlaut was not an option (the vowel was already front), neither was tone change (the tone was already acute), so that the only remaining option was to add an ending. In the case of the masculine **déef̲**, this ending had to be -**ən** (the ending -**ər** is restricted to neuters), so that the present-day form is **deēvən**. The circumflex accent in this plural form, which does not correspond to the historically expected acute, is explained by the model on which this form must be based by speakers without access to the historical forms, namely the alternating-tone model of **dúuf̲** ‘dove’ ~ **duūvən** ‘doves’, which is the predominant pattern for nouns with -**ən** plurals. This -**ən** addition was originally restricted to cases where the less shocking modifications (umlaut, tone change) could not apply. In the twentieth century we find an extension of this change to all words with non-umlauting vowels (i.e. back vowels plus, as far as plurals are concerned, the vowel **aa**), independent of whether a tone change can apply. Thus next to the tone pair **daāx̲** ‘day’ ~ **dáax̲** ‘days’, which is still in use, one can nowadays hear the plural **daāɣən**. Another example is the extension of **jɔ́ɔr** ‘year-SG&PL’ to the new plural **jɔɔ̄rən** ‘years’, analogously to, say, the historical **kɔ́ɔl** ‘coal’ ~ **kɔɔ̄lən** ‘pieces of coal’.

12. Four new kinds of tones

First let me assert that present-day Limburgian is not just a lexical mora-accent language. Instead, it may be closer to a tone language, because (for instance) sentence-final acute and circumflex syllables strongly contrast with each other even if the sentence-final word is not in focus. Thus, a Roermond question intonation on the word **ɣəwéɛs** ‘been’ will sound like [ɣə́wɛ́̋ɛ̏s], with an octave drop in the final syllable, whereas a question intonation on the word **ɣədɔɔ̄n** ‘done’ will sound like [ɣə́dɔ̋ɔ̏ń], with both an octave drop and a fifth rise in the final syllable. Gussenhoven (2000ab) interprets [ɣə́wɛ́̋ɛ̏s] as having only a HL interrogatory boundary tone, and [ɣə́dɔ̋ɔ̏ń] as having the same HL boundary tone followed by a H tone that must have been supplied in the underlying form. In Gussenhoven’s view, then, the contrast between acute and circumflex for Roermond is that between the absence and presence of a lexical H tone.

Meanwhile, the Roermond sentence-internal contrast is still often one of mora alignment, as in (1) and (2).

But present-day sentence-internal contrasts are not *always* cases of mora alignment contrasts. After schwa drop, the language was still thoroughly mora-accenting, in the sense that the only acute-circumflex contrast was that between a monosyllabic bimoraic acute and a monosyllabic bimoraic circumflex. Beside these two types of accents, however, the later sound changes of degemination and schwa insertion caused the emergence of four more types: the monomoraic acute, the second monomoraic circumflex, the disyllabic acute, and the disyllabic circumflex. These four new accents are discussed in the following sections.

12.1 Degemination and the rise of the monomoraic accents

From §5.5 we know that the words **bɛddən** 'beds' and **kɑttən** 'cats' were pronounced [bɛ́ɛ̀də̀ə̀n]$_{Aud}$ and [kátːə̀n]$_{Aud}$ (we now discuss only the plurals of these words, because the singulars have lost their final schwas in §7.5). A subsequent probably speaker-induced sound change was a gradual shortening of the geminate consonants. A certain generation of learners must then have interpreted these consonants as singletons. The question then is: with what alignment of the HL contour to the remaining moras?

There is a by now familiar difference between the auditory forms [bɛ́ɛ̀də̀ə̀n]$_{Aud}$ and [kátːə̀n]$_{Aud}$: only in the former does a pitch movement occur within the first syllable (§5.5). If the geminate is reanalysed by a learner as a singleton, the first syllable must be reanalysed as monomoraic, and the original syllabic pitch movement can be reanalysed as a pitch movement within the new monomoraic syllable.[27] The resulting form can be abbreviated as **bɛ̂dən**. Likewise **brɔ̂gən** 'bridges', **ʃwɔ̂mən** 'to swim', and so on. Since the pitch movement is (very) early, a structure like **bɛ̂dən** can be called a *monomoraic acute*.

No acute reanalysis is possible in the case of a shortening geminate in [kátːə̀n]$_{Aud}$. The only possibility is a reanalysis with a H tone on the first (now monomoraic) syllable. Since this means that the pitch movement now occurs after the first syllable, the resulting form can be abbreviated as **kɑ̄tən** (likewise **ʃtɛ̄kən** 'pieces' and so on). This structure can be called a *monomoraic circumflex*.

The contrast just discussed has not received much attention, probably because the distinction is predictable on the basis of the voicing of the consonant. In fact, the post-lexical process of word-final prevocalic voicing causes acute-circumflex alternations within the same word form. Thus (Ben Hermans, p.c.), **brœk** 'bridge' and **ʃtœk** 'piece' are both pronounced with a circumflex in **ən brœk̄ tə zeēn** 'to see a bridge' and **ə ʃtœk̄ tə kriīɣə** 'to get a piece' (with question intonation: rising on **t**), and both with an acute in **woo də brɔ̂eg es** 'where the bridge is' and **woo ət ʃtɔ̂eg es** 'where the piece is' (with question intonation: rising in **œ**).

Nevertheless, the phonetic distinction between **bɛ̂dən** and **kɑ̄tən** is large, especially in sentence-final focal position with question intonation: **bɛ̂dən** has its pitch peak somewhere in the first vowel or the following consonant ([bɛ̏ɛ̋́də]$_{Aud}$) whereas **kɑ̄tən** has its pitch peak in the second vowel ([kɑ̀tɑ̋́ˋ]$_{Aud}$) (Ben Hermans p.c.). Moreover, scattered monomoraic circumflexes occur in voiced environments, for

[27] I say "can be reanalysed", not "must be reanalysed", because if one looks closely at the auditory form, an alternative reanalysis with H-toned first syllable seems to be possible as well.

instance in the Geleen form **ɦøːmən** ‘shirt’. In this form, the cause of the exceptional circumflex can lie in a late degemination caused by an original heterorganic nasal-plosive sequence: ***ɦɛmədə** > ***ɦɛmdə** > ***ɦɛm̄də** > ***ɦɛm̄bə** > ***ɦɛm̄mə**, with a drop of the first schwa bleedingly before Open Syllable Lengtening (§5.1), and the ***d** > ***b** change counterfeedingly after geminate lenition (§7.7). This word has an exceptional circumflex in Maastricht as well (C. Gussenhoven, p.c.), and the long vowel in **ɦøømə** in Weert Limburgian (which is nowadays a toneless dialect according to Heijmans and Gussenhoven 1998) must also reflect an earlier circumflex.

12.2 Schwa insertion and the rise of the disyllabic accents

When a liquid was followed by a non-coronal obstruent, a schwa was inserted. The circumflex words **bæːr̲x** ‘mountain’ and **wɔːl̲f** ‘wolf’ therefore became **bæːrəx̲** ‘mountain’ and **wɔːləf̲**, still with the tone movement during or after the final consonant. The resulting accent can therefore be called a *disyllabic circumflex*. Likewise, **æːrəm** m. ‘arm’ and **ʃæːrəf̲** n. ‘splinter’ (MDu *scerf*, not NHG *Scherbe* f.), which could be added to (31), and with voiceless final schwa drop **bɑːləkˌ** m. ‘beam’ (NHG *Balken*) and **keːrəkˌ** f. ‘church’ (NHG *Kirche*), which could be added to (45).

In large parts of Limburg a schwa is also often appended phrase-finally after a word that ends in a high vowel, as in **nuːwə** ‘now’, which is pronounced with a pitch pattern identical to that in **huːs** ‘house’. In Sittard I have heard the same disyllabic circumflex on words where a sonorant consonant is flanked by two short high vowels: **jyːni** ‘June’, **jyːli** ‘July’, **fəmiːli** ‘family’, **ʃiːli** ‘chilli, Chile’, **kiːwi** ‘kiwi’, **bəkiːni** ‘bikini’; whether this phenomenon is widespread or idiosyncratic still has to be investigated.

A similar change occured in acute words: **bérɣ** ‘mountains’ (< ***bæːrɣə**; NHG *Berge*) became **bérəɣ**. The simplest assumption about the pitch movement is that it stayed in the middle of the word, which in **bérəɣ** must be close to the syllable boundary. This structure can be called a *disyllabic acute*. Likewise **érəv** n. ‘premises’ (< ***ɛːrvə**; OLF *ervi* ‘inheritance’, NHG *Erbe*), **érəm** pl. ‘arms’ < ***ɛːrmə**. These examples could be added to (51).

There is a clear phonetic distinction between the monomoraic circumflex as in **kɑːtən** (pitch movement during **t** or **ə**) and the disyllabic circumflex as in **bæːrəx̲** (pitch movement during or after **x**). Whether there is also a phonetic distinction between the monomoraic acute as in **bɛ̂dən** (pitch movement during **ɛ**) and the disyllabic acute as in **bérəɣ** (pitch movement between **æ** and **r**?) remains to be seen.

13. Comparing theories on Franconian tonogenesis

The crucial evidence that establishes the order of events in the Rule A2 regions is the retention of the original accent in disyllabic words, namely the retention of the acute in originally long non-high vowels, both in the voiceless cases of (11) and in the voiced cases of (12), and the retention of the circumflex in lengthened vowels, both in the voiceless cases of (20) and in the voiced cases of (21). This evidence clearly places the beginning of the lexical accent contrast in the result of the otherwise neutralizing process of open syllable lengthening. No other proposals than the present one realize that this evidence is crucial for establishing the order of events; all other proposals than the present therefore misplace the first occurrence of the contrast, typically locating it later, such as in the result of schwa drop or in the result of analogical lengthening. As a

consequence, these other proposals can handle only a small part of the data. I will now review several of these proposals.

13.1 De Vaan (1999)

Of all the other proposals, De Vaan (1999) comes closest to the correct chronology, albeit sometimes for the wrong reasons. I will therefore discuss his proposal in some detail.

De Vaan (p. 35) starts his chronology by stating, as I do here in §4.2, that in pre-OSL times long non-high vowels had a moving (De Vaan says "falling") pitch, and that other vowels had a level (De Vaan says "rising") pitch. Presumably (though De Vaan gives no examples here) this means that e.g. the word 'seek' from (11) was pronounced [zǒǒkə̀ǹ]. Next, De Vaan states that as a result of Open Syllable Lengthening, originally short vowels became level (De Vaan says "rising") before single medial voiceless consonants. Presumably, the means that e.g. the word 'kitchen' from (20) became pronounced as [kǒǒkə̀ǹə̀]. Now, I argued in §5.1 that the result is a lexical accent contrast, because words like [zǒǒkə̀ǹ] and [kǒǒkə̀ǹə̀] differ only in their pitch pattern on the first syllable, which is unpredictable from the surrounding segments. De Vaan, however, does not realize that OSL caused the lexical contrast. Instead, he states next that "the tonal opposition of falling [i.e. moving] pitch versus rising [i.e. level] pitch reached phonemic status through [the process of] apocope [which happened after OSL]".

De Vaan's oversight may be due to his failure to look at the system of oppositions at each stage of the language. This happens more often in his paper. In a footnote (p. 29), De Vaan dismisses a proposal by Goossens according to which Limburgian originally had Rule A and then turned many acutes into circumflexes "through analogy", thus leading to Rule A2. Presumably, this would mean that an acute on lengthened vowels in voiced positions, as in **léɛzən** 'read', turned into a circumflex, as in **lɛɛ̄zən** from (21). De Vaan correctly notes that there is not much to base an analogy on, but his dismissal is further only based on a principle of "economy". The logical-chronological reason, however, why the words in (21) could not have come from acutes is that in that case the words in (12), such as **kéezən** 'choose', would also have turned their acute into circumflexes, as in **keēzən**, a neutralizing change that according to the list (12) did not happen: circumflexes form a small minority in (12), while acutes form a small minority in (21). Again, De Vaan misses the crucial systemic argument that would derive from looking at the synchronic contrasts and oppositions at every stage of the language. On p. 34, De Vaan correctly states that the Rule A2 dialects "have preserved a more archaic situation" than the Rule A dialects, but he does so for the wrong reason, namely the geographical peripherality of the Rule A2 dialects. This is an invalid argument. In De Vaan's view, the Rule A2 dialects form a contiguous area in the North and West; since there are more contiguous peripheral areas, not all of them can be archaic, which strongly indicates that peripherality points at archaism only if these peripheral areas are scattered and unconnected, which they are not. The logical reason for the greater age of Rule A2 is their non-neutralization, i.e. the fact that they retain the contrast between the forms in (12) and in (21), a distinction lost in the Rule A areas. Logically, neutralizations are innovations, and retentions of etymological contrasts are archaisms. The only example that De Vaan gives to illustrate the archaism of Rule A2 is **váart** ~ **vaārən**, and this does illustrate the problematicity of an analogical cause of

Rule A2 (why would **váarən** change into **vaārən** if the other forms have an acute as well?), but it does not refer to the more crucial fact, which is the retention of, say, the **vaārən** ~ **kéezən** contrast.

A third case of unclear argumentation for an original Rule A2 is the statement (p. 35) that "the sequences [...], VD [i.e., originally short vowels before voiced consonant] developed a falling pitch" during OSL. If this had been true, the lengthened V:D would have fallen together with the originally long V:D, as in Rule A. Since in Rule A2 the former nowadays have a circumflex and the latter an acute (as De Vaan acknowledges), this cannot be correct.

I conclude that De Vaan was correct in proposing phonetic pre-OSL contrasts, correct for the wrong reason in proposing the primacy of Rule A2, and incorrect in locating the inception of the lexical accent contrast in the results of schwa drop.

13.2 Gussenhoven (2000c)

Of all the proposals, the one by Gussenhoven (2000c) has been most widely distributed over the world. At the same time, it is the least credible of all the proposals, because it accounts for only a little corner of the data and has logical-chronological problems with the remainder of the data.

Gussenhoven proposes that the lexical accent contrast originated through a mechanism of "fake analogical lengthening" (Fake AL; p. 231). At a certain point in time, Franconian would have had a length contrast between **dax** 'day' and **daax** 'days', the length in **daax** resulting from open syllable lengthening followed by schwa deletion (**dayə** > **daayə** > **daax**). At the same time, an influential neighbouring dialect would have had the older **dax** ~ **daayə** alternation but decided that vowel length should be homogeneous throughout the paradigm, leading to **daax** ~ **daayə**. This is the usual process of *analogical lengthening* (§6). At this point, dialect 1 would have had **dax** ~ **daax**, dialect 2 **daax** ~ **daayə**. According to Gussenhoven, the speakers of dialect 1 would decide to mimic speakers of dialect 2 in having a long vowel in the singular. However, doing so would have turned the paradigm into **daax** ~ **daax**, thus removing the distinction between singular and plural. Gussenhoven proposes that the speakers then decided to make a difference between the two forms, somehow adding a high tone to the singular: **daa˝x** ~ **daax**. This is what Gussenhoven calls "fake analogical lengthening". Despite the evidence presented in the current paper, Gussenhoven has repeated this hypothesis numerous times, most recently in Gussenhoven (2012).

The largest problem with Gussenhoven's proposal is that it cannot explain how the tone contrast spread from **daa˝x** ~ **daax** to the rest of the lexicon. Presumably, fake analogical lengthening would have to have postdated schwa drop (otherwise, no threat of homonymy), which must have postdated open syllable lengthening (otherwise, no vowel lengthening in the plural **daayə**), and this is indeed the order in which Gussenhoven arranges these three changes on p. 233. Also, fake analogical lengthening by itself must have postdated open syllable lengthening, otherwise there would have been no long vowel that could be the source of the analogy. Now, with the order OSL → schwa drop → Fake AL, it is virtually impossible that the change could have spread through the lexicon along etymological lines. If tonogenesis (i.e. Fake AL) follows schwa drop (which comes with final devoicing, as G. states), an original **huu̲s** 'house' and **druuvə** 'grape' would have neutralized their structures as **huu̲s** and **druu̲f**. How could subsequently **huu̲s** (and all other such forms without schwa drop) get a H tone,

whereas **druuf̲** (and all other such forms with schwa drop) did not? In other words, how could the neutralization be undone along etymological lines?[28] Even more problematically, if tonogenesis (Fake AL) followed OSL, how did **køøkənə** 'kitchen' (and all other such forms with originally short vowels) get a H tone, whereas **zøøkən** 'seek' (and all other such forms with originally long non-high vowels) did not? In other words, if Gussenhoven's theory is correct, the spread of tones through the lexicon must have undone the neutralization caused by schwa drop, such as that between **huus̲** and **druuf̲**, as well as the neutralization caused by open syllable lengthening, such as that between **køøkənə** and **zøøkən**. In reality, the original schwa-based contrast between **huus̲** 'house' and **druuvə** 'grape' (and many other such forms) has been preserved as the present-day tonal contrast **huūs̲** ~ **drúuf̲**, and the original short-long contrast between **køkənə** 'kitchen' and **zøøkən** 'seek' (and many other such forms) has been preserved as the present-day tonal contrast **køōkən** ~ **zǿøkən**. Neutralization can have no place in these changes, so Gussenhoven's proposal must be rejected on the basis of an incompatibility with one the main heuristic axioms of historical phonology.

One can wonder what the causes are of Gussenhoven's neglect of the **huūs̲** ~ **drúuf̲** and **køōkən** ~ **zǿøkən** cases, i.e. of the majority of the data. One cause seems to be an incorrect view of the data themselves. Within six lines on p. 237, Gussenhoven incorrectly lists Tongeren **zwɔ́ɔn** 'swan' (< ***zwanə**) and **brýək** 'fracture' (< ***brøkə**) as having originally long vowels, and **ɦúək** 'hook' (< **ɦɑɑkə**) as having an originally short vowel; under such circumstances it is indeed difficult to arrive at correct generalizations, especially if multiple competing factors play a role, as they do here. Another cause seems to be the failure to acknowledge the known interactions of these competing factors. In a footnote on p. 254, G. claims that "OSL is an unlikely locus for the tonogenesis, since it systematically fails to appear in forms in Table 2 that underwent it." Gussenhoven's Table 2 is a list of singular–plural pairs where the singular has AL; the plurals, which underwent OSL, indeed have an acute (**wɛɛ̄x** 'way' ~ **wéɛɣ** 'ways'; **ʃmeēt̲** 'smith' ~ **ʃméej** 'smiths'), but this is just the group that received an acute secondarily during schwa drop (***wɛɛ̄ɣə** → **wéɛɣ**; ***smeēðə** → **ʃméej**), i.e. half of the forms in (51). The much greater number of forms in (20) and (21), which are still disyllabic, still have the circumflex.

I conclude that G.'s proposal accounts only for the accent in AL cases (here §6) and fails on all other monosyllabic and disyllabic cases, i.e. on over 90 percent of the data. I agree that it is tempting to start an account of the accent contrast from singular–plural pairs; in fact, the examples in (1) and (2) provide just such an example. However, the hypothesis developed here, which is that the accent contrast originated in OSL, explains the **daāx̲** ~ **dáaɣ** alternation by turning an original **dax̲** ~ **daɣə** first into **dax̲** ~ **daāɣə** by OSL (with a contrastive circumflex accent), then into **daāx̲** ~ **daāɣə** by AL (not "Fake", because the accent is the same in singular and plural), and finally into **daāx̲** ~ **dáaɣ** by schwa drop (it is a true minimal pair, because the final consonant is currently voiceless in both words). In other words, these singular–plural pairs are not the source of the contrast, but instead form the most complicated cases, resulting from an interaction of at least three changes.

[28] For the **huus̲** ~ **druuf̲** case, G. tries to save the chronology (in an apparent defence against objections by the present author to an earlier version of G.'s paper) by proposing that there was a stage when schwa drop and Fake AL were optional rules that were in effect at the same time. This proposal cannot work for the **køøkənə** ~ **zøøkən** case, though, where the ordering problem involves OSL as well.

13.3 Schmidt (2002)

Jürgen Erich Schmidt criticizes both De Vaan and Gussenhoven for problems such as the phonetic naturalness of their proposed changes, but he does not criticize them for problems with their chronological logic. In fact, Schmidt claims that the lexically contrastive accent originated 'relatively late' (p. 207: "relativ spät") with schwa drop (p. 207: "nach dem Apokopierungsprozess" 'after apocope'; p. 220 "Apokopierung der Endsilbe" 'apocope of the final syllable'), summarized as "Auslöser der Tonakzentgenese war ein durchgreifender makrostruktureller Wortveränderungsprozess: der Ausfall einer Endsilbe (Apokope)" 'the trigger of tonogenesis was a radical macro-structural word change process: the drop of a final syllable (apocope)'. Given that Schmidt is very well acquainted with the historical conditioning of the contrast, one has to wonder how he could have missed the contrast between the old long non-high vowels in (11) and (12), which have acutes, and the lengthened vowels in (20) and (21), which have circumflexes.

The answer is partly in Schmidt's neglect of Rule A2. On p. 224, Schmidt claims that 'Rule A2 has been insecurely documented', because for the Rule A2 areas there sometimes exist dialect descriptions that describe them in terms of Rule A. Perhaps Schmidt thought here of work such as that by Tans (1938), who describes the situation in the Maastricht area (Rule A2) with Welter's (1910) Rule A conditions; however, Tans' *data* clearly show Rule A2 conditions, which Tans labels explicitly as exceptions to Welter's conditions (p. 21, 160, 216). Such a situation, with an apparently uninformed researcher listing exceptions to Rule A that he cannot put in a wider perspective but are regular from a larger point of view (i.e. from the independently reported Rule A2 perspective), should be regarded as *confirming* the Rule A2 situation rather than questioning it. Because of his inappropriate neglect of Rule A2, then, Schmidt does not count the contrast between the originally long vowels in (12), which are acutes nowadays, and the lengthened vowels in (21), which are circumflexes nowadays, as relevant or even existing, because all of these vowels (before voiced consonants) have circumflexes in the Rule A areas today. Likewise, Schmidt may have missed the contrast between (11) and (20) as relevant, because in the Southern part of the area the vowels in (20), standing before voiceless consonants, may never have lengthened.

In an area where Rule A applies *and* lengthening did not occur before voiceless consonants, tonogenesis may indeed have originated in schwa drop. If it indeed happened so late in these areas, it must still have happened as well in the Rule A2 areas centuries earlier, when open syllables lengthened there. That is, any late tonogenesis in the Southern areas must have followed a much earlier tonogenesis in the North.

13.4 Alternative possible correct proposals

In my discussions of others' proposals above, I not only pointed out their failures but also presented the hypotheses developed in the present paper as true. They do not have to be; there may be other thinkable logically-chronologically correct hypotheses.

The idea that tonogenesis originated with schwa drop can be saved by assuming that the originally long non-high vowels ***ɑɑ**, ***ææ**/***ɛɛ**, ***œœ**, ***ɔɔ**, ***ee**, ***øø** and ***oo** turned into ***ɔə**, ***e̯ə**, ***ø̯ə**, ***o̯ə**, ***e̯ə**, ***ø̯ə** and ***o̯ə** before Open Syllable Lengthening, i.e. all of them "broke" into a diphthong ending in schwa. When OSL then lengthened the originally short vowels, the lengthened vowels would have been kept distinct as a

group from the originally long non-high vowels. Subsequently, schwa drop, which came with tonogenesis, would have turned the seven vowels into ***ɔ́ɔ**, ***é̯ɛ**, ***ǿ̯œ**, ***ó̯ɔ**, ***é̯e̯**, ***ǿ̯ø̯** and ***ó̯o̯**, which are not far removed from the present-day Limburgian forms. The evidence for pre-OSL diphthongs is scarce: it may only be in the OLF writing of the higher-mid series as *ie* and *uo* (the present-day diphthongization of the lower-mid series cannot count as evidence, because it would have to follow schwa drop, i.e. monophthongization), but those spellings may have a variety of different causes (e.g. coming from a different region, or representing only an earlier stage than MLb, or representing raised mid monophthongs). Weighing the evidence, I cannot regard this scenario as likely.

14. Conclusion

We have seen that a sequence of synchronic phonological systems, gradual shifts by speakers and discrete reanalyses by listeners accounts for much of the data. The true phonetic and phonological details may differ from the ones presented, but I firmly believe that the order of events is as depicted; especially, the claims are that the accent contrast originated with Rule-A2-like open syllables by Open Syllable Lengthening, that schwa drop did nothing more than produce many new acutes, and that Rule A came about later as a neutralization. Comparisons with similar sound changes in Lithuanian (lengthening yield circumflexes and vowel drop turns circumflexes into acutes) showed that my account is not peculiar to the Limburgian case. Realizing that about 95 percent of the Limburgian accents have been accounted for, and that no systematic exceptions remain, I conclude that the chronology of the Franconian tonogenesis, at least for Rule A2 and Rule A, has been settled. The main question that remains is how the mora accent language turned into the present-day language with its synchronic tonal and metrical analyses.

References

Bach, Adolf. 1921. Die schärfung in der moselfränkischen mundart von Arzbach. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 45(2), 266–290.

Cosijn, Peter Jacob. 1872–1873. De Oudnederlandse psalmen. *De Taal- en Letterbode* 3 (pp. 25–48, 110–124, and 257–270) and 4 (pp. 149–176).

Cowan, H.K.J. 1959. De localisering van het Oudnederfrankisch der psalmenfragmenten. *Leuvense Bijdragen* 48, 1–47.

Cowan, H.K.J. 1969. Nogmaals de localisering van de Oudnederfrankische psalmenfragmenten. *Leuvense Bijdragen* 58, 114–132.

Dols, Willy. 1944. *Sittardse diftongering*. [published posthumously by Alberts' Drukkerijen, Sittard 1953]

Dupont, J. 1910–11. Het dialect van Bree. *Leuvensche Bijdragen* 9, 193–212.

Engelmann, René. 1910. Ein mittelfränkisches accentgesetz. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 36(2), 382–394.

Goossens, Jan. 1965. Die Gliederung des Südniederfränkischen. *Rheinische Vierteljahrsblätter* 30, 79–94.

Goossens, Jan. 1977. De tweede Nederlandse Auslautverscherping. *Tijdschrift voor Nederlandse taal- en letterkunde* XCIII 1–2, pp. 3–23.

Grootaers, Ludovic. 1909. Het dialect van Tongeren: eene phonetisch-historische studie. *Leuvensche Bijdragen* 8.

Gussenhoven, Carlos. 2000a. The boundary tones are coming: On the non-peripheral nature of boundary tones. In Michael Broe & Janet Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Acquisition and the lexicon*, 132–151. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2000b. The lexical tone contrast of Roermond Dutch in Optimality Theory. In M. Horne (ed.) *Prosody: Theory and experiment. Studies presented to Gösta Bruce*. Dordrecht/Boston/London: Kluwer.

Gussenhoven, Carlos. 2000c. On the origin and development of the Central Franconian tone contrast. In A. Lahiri (ed.) *Analogy, levelling, markedness: Principles of change in phonology and morphology*, 215–260. Berlin: Mouton de Gruyter.
Gussenhoven, Carlos. 2012. From Cologne to Arzbach: an account of the Franconian 'tone reversal'. To appear in *Proceedings of Nordic Prosody XI*.
Heijmans, Linda & Carlos Gussenhoven. 1998. The Dutch dialect of Weert. *Journal of the International Phonetic Association* 28, 107–112.
Helten, Willem Lodewijk van. 1896. Een en ander over en naar aanleiding van de Oudnederlandsche psalmvertaling. *Tijdschrift voor Nederlandsche Taal- en Letterkunde* 15, 146–171, 269.
Helten, Willem Lodewijk van. 1902. *Die altostniederfränkischen Psalmenfragmente, die Lipsius'schen Glossen und die altsüdmittelfränkischen Psalmenfragmente, mit Einleitung, Noten, Indices und Grammatiken*. Groningen: Wolters.
Hermans, Ben. 1984. Het Limburgs en het Litouws als metrisch gebonden toontalen. *Spektator* 14, 48–70.
Hermans, Ben. 2006. De tonologie van onbeklemtoonde lettergrepen in een Limburgs dialect. *Taal en Tongval* 19, 55–70.
Hermans, Ben. 2012. *The phonological representation of the Limburgian tonal accents*. In B. Botma & R. Noske (eds.) *Phonological explorations: empirical, theoretical and diachronic issues*, 227–244. Berlin: Mouton de Gruyter.
Jongen, René. 1972. *Phonologie der Moresneter Mundart: eine Beschreibung der segmentalen und prosodischen Wortformdiakrise*. Assen: Van Gorcum.
Kats, J.C.P. 1985. *Remunjs waordebook*. Roermond: H. van der Marck.
Kehrein, Wolfgang. This volume. There's no tone in Cologne: against tone–segment interactions in Franconian.
Kiparsky, Paul. 1973. The inflectional accent in Indo-European. *Language* 49, 794–849.
Köhnlein, Björn. 2011. *Rule reversal revisited: synchrony and diachrony in the Franconian dialect of Arzbach*. Doctoral dissertation, Leiden University.
Köhnlein, Björn. 2013. Optimizing the relation between tone and prominence: evidence from Franconian, Scandinavian, and Serbo-Croatian tone accent systems. *Lingua* 131, 1–28.
Köhnlein, Björn. This volume. Synchronic alternations between monophthongs and diphthongs in Franconian tone accent dialects: a metrical approach.
Kyes, Robert L. 1969. *The Old Low Franconian psalms and glosses*. Ann Arbor MI: University of Michigan Press.
Leenen, J. 1915. Dialecten in Belgisch Limburg. *Limburgsche Bijdragen* 13, 147–164.
Liberman, Anatoli. 1999. *Schärfung/stoottoon* and *trägheitsakzent/sleeptoon* in the Rhein-Limbourg area and their Scandinavian analogues. In E.C. Polomé & C.F. Justus (eds.) *Language change and typological variation. In honor of Winfred P. Lehmann on the occasion of his 83rd birthday*, Vol. I: *Language change and phonology*, 275–298. Washington: Institute for the Study of Man.
Maurmann, Emil. 1898. *Grammatik der Mundart von Mülheim a. d. Ruhr* (*Sammlung kurzer Grammatiken deutscher Mundarten* 4). Leipzig. [partial print, Marburg 1889; reprint Wiesbaden 1969]
Meer, S. van der. 1949. *Venloer Stadt-Texte 1320–1543: eine lautliche und orthographische Untersuchung*. Nijmegen: Dekker & Van de Vegt.
Nörrenberg, K. 1884. Studien zu den niederrheinischen mundarten. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 9, 371–421.
Oostendorp, Marc van. This volume. Tone, final devoicing and assimilation in Moresnet.
Reitz, Guido. 1985. Die Rheinische Akzentuierung im Morbacher Platt. Manuscript.
Sanders, Willy. 1968–1969. Zu den altniederfränkischen Psalmen. *Zeitschrift für deutsches Altertum und deutsche Literatur* 97, 81–107.
Schmidt, Jürgen Erich. 1986. *Die mittelfränkische Tonakzente (Rheinische Akzentuierung)*. Stuttgart: Franz Steiner.
Schmidt, Jürgen Erich. 2002. Die Sprachhistorische Genese der mittelfränkischen Tonakzente. In Peter Auer, Peter Gilles & H. Spiekermann (eds.) *Silbenschnitt und Tonakzente*, 201–233. Tübingen: Niemeyer.
Tans, J.G.H. 1938. *Isoglossen rond Maastricht*. Doctoral dissertation, Katholieke Universiteit Nijmegen.
Vaan, Michiel de. 1999. Towards an explanation of the Franconian tone accents. *Amsterdamer Beiträge zur Älteren Germanistik* 51, 23–44.
Welter, Wilhelm. 1933. Die niederfränkischen Mundarten im Nordosten der Provinz Lüttich. The Hague: Martinus Nijhoff.
Wijk, Nicolaas van. 1935. De klinkerrekking en de stoottoon vóór stemhebbende medeklinkers in het Limburgs en in andere dialekten en talen. *Nieuwe Taalgids* 29, 405–411.
Wijk, Nicolaas van. 1939. De Rijns-Limburgse polytonie, vergeleken met de Kasjoebse. *Onze Taaltuin* 8, 146–152.