3rd European Conference on
Speech Communication and Technology
EUROSPEECH'93
Berlin, Germany, September 19-23, 1993

A PERCEPTUAL EXPLANATION OF THE WEIGHTLESSNESS OF THE SYLLABLE ONSET

Rob Goedemans & Vincent J. van Heuven

Dept. Linguistics/Phonetics Laboratory/Leiden University
and
Holland Institute of Generative Linguistics
Leiden/Amsterdam, The Netherlands

10.21437/Eurospeech.1993-339

ABSTRACT

Quantity-sensitive stress languages typically stress the heaviest syllable in a word. Segments in the syllabic nucleus (i.e. the vowels) and in the coda (i.e. the post-vocalic consonants) contribute to syllable weight; pre-vocalic consonants (syllable onset) add no weight. This study presents a phonetic explanation for the weightlessness of the syllable onset. Both a speech production and a perception experiment were carried out. On the assumption that syllable weight can be operationalised in terms of duration, the production data provide no support for the weightlessness of the onset: added onset consonants contributed as much, if not more, to syllable duration than consonants added in the coda. However, the perceptual data provide a clear correlate of the weight difference between onset versus nucleus and coda. The results of a duration adjustment task indicate that variations in the onset consonant are not heard; identical variations in the coda consonant are reproduced adequately, whilst variations in the vowel are even perceptually overestimated.

1. INTRODUCTION

In so-called quantity-sensitive stress languages (e.g. English, German, Dutch) the location of the main stress position (defined as the docking site for a pitch-accent, cf. [1,2]) in a word is determined by syllable weight. The heaviest syllable in the word attracts the stress position. The weight of a syllable is determined by counting the segments in the vocalic nucleus and in the coda (the post-vocalic consonants); consonants and short vowels count as one segment, long vowels and diphthongs as two segments. As a first approximation, syllable weight can be operationalised in terms of duration: syllables are longer as they contain more segments. Crucially, however, the segments in the syllable onset (the pre-vocalic consonants) do not add to the weight of a syllable [3,4].

Why should language systems ignore pre-vocalic consonants in the determination of syllable weight? The present paper reports two related experiments set up to explore possible causes for what we have come to call the weightlessness of the syllable onset.

In our first experiment we sought to explain the weightlessness of the syllable onset from a speech production perspective. The linguistic behaviour of the onset would be understandable if the duration of the onset were constant for any number of onset consonants, or at least were less affected by the number of pre-vocalic segments than the nucleus or coda would be for the number of vowels and postvocalic consonants. A more involved mechanism, that would yield the same result, would be that lengthening the onset, by changing the number or duration of pre-vocalic consonants, would be compensated for by shortening the remainder of the syllable, whilst no (or at least less) compensation would occur in the onset when the number or duration of (post-)vocalic segments is changed. Whichever the case may be, syllable duration should be longer as the number of segments in the vocalic nucleus or the coda is increased but remain (more) constant when the number of segments in the onset is changed.

In order to check this possibility we measured onset and syllable duration as a function of the number of onset segments (empty onset, 1, 2, or 3 pre-vocalic consonants), and compared these measurements with the duration effects of similar variations in the number of segments in the vocalic nucleus (one vs. two segments) and coda (empty coda, 1, 2, or 3 post-vocalic consonants).

Alternatively, the weightlessness of the syllable onset may have a perceptual origin. In our second experiment we tested the hypothesis that the perceived duration of the syllable depends more on the duration of the nucleus and the coda than on the duration of the onset. If this is the case, changing the duration of a syllable onset should influence the perceived duration of the syllable to a smaller degree than changing the duration of the nucleus and/or the coda.

2. PRODUCTION EXPERIMENT

Method

Five tokens of the following meaningful Dutch words were recorded by two native speakers of Dutch, in the fixed carrier sentence *Wil je* [target] *eens zeggen* /wIl j@ ... @ns zEG@/ 'Would you ... please say'.

subset a:	op	'on'
onset, V-nucleus	sop	'suds'
	stop	'stop'
	strop	'noose'
subset b:	ga	'go'
coda, VV-nucleus	gaaf	'neat, unscathed
	gaafs	'something neat'
	gaafst	'most neat'
subset c:	laf	'cowardly'
coda, V-nucleus	lafs	'something cowardly'
	lafst	'most cowardly'

This set constitutes a non-ideal but workable compromise between the phonotactic and lexical limitations of Dutch and the full expansion of the ideal symmetrical pair of schemata $(((C_1)C_2)C_3)V_1(V_1)C_4$ and its mirror image (2*8 combinations). Note that, in our selection, the coda consonant is kept constant when the onset is the target, i.e. varied in number (subset a), and that the onset is kept constant (within the subset) when the nucleus and coda are the targets (subset b for coda targets with long vowels; subset c for coda targets with short vowels). An empty onset is phonotactically legal in Dutch but its position will typically be realised by a glottal stop at the onset of an accented word [5]. The coda must not be empty after a short vowel (hence the absence of an empty coda in subset c). An empty coda (open syllable) is perfectly legal after a long vowel; vowels in open syllables are not followed by a glottal stop but will be longer than in closed syllables [6].
The recordings were made on a REVOX B-77 tape-recorder in a sound attenuating recording booth with a Sennheiser MKH 416 condenser microphone.
Segment boundaries were determined by examining oscillograms, using a high resolution waveform editor (10 kHz, 12 bit, 4.5 kHz LP), and stored on computer disk.

Results and discussion

Mean onset and coda duration is plotted in figure 1A and 1B, respectively, as a function of nucleus duration (plotted along the horizontal dimension) and number of segments in the target consonant cluster (numerical codes in the plot). The number of empty onsets ("0") is smaller in figure 1B since there are no empty codas after short vowels. In figure 1B we also see that the measurements for short and long vowels group into two clusters (short vowels lie to the left of the boundary, long vowels to the right). Unlike in panel A the empty cluster has a duration of 0 ms. This is so because empty codas are not filled with glottal stops. Unlike empty onsets, empty codas are also physically empty.
Figure 1A shows that the duration of the onset increases monotonically with the number of segments in it, even when empty onsets are not considered, $F(2,27)= 61.1$ ($p<.001$). Similarly, in figure 1B the duration of the coda increases with the number of coda consonants, $F(2,27)=38.0$ ($p<.001$) and $F(2,27)=39.4$ ($p<.001$) for codas after long and short nuclei, respectively (excluding empty codas). Longer consonant cluster duration is generally compensated for by shortening the nucleus. Crucially, however, there is less compensation in the vowel for longer onset duration ($r=-.30$, $p=.10$) than for longer coda duration ($r=-.63$, $p<.01$ for short vowels, even when excluding empty codas; $r=-.48$, $p<.05$ for long vowels).

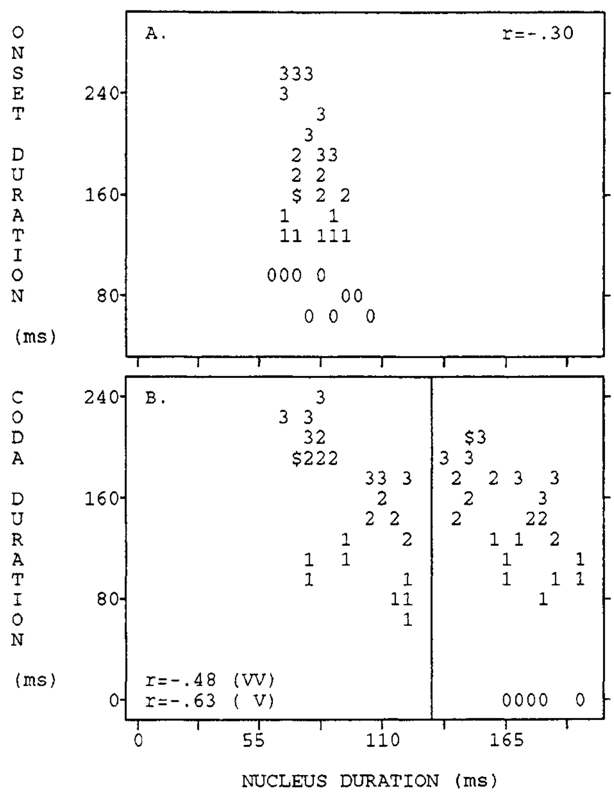

Figure 1: Onset duration (A) and coda duration (B) as a function of nucleus duration broken down by number of segments in target cluster (indicated by numerals in plot; $ represents multiple occurrence).

In sum, the effect of number of consonants on the duration of the syllable as a whole, as well as on the duration of the target cluster, is larger for onset complexity than for coda complexity. These results run counter to the suggestion made above that the duration of onset clusters is constant, and does not contribute to syllable duration. We may therefore safely rule out the possibility that the weightlessness of the onset has its cause in speech production.

3. PERCEPTION EXPERIMENT

Now that we have not been able to relate the weightless behaviour of the syllable onset to speech production mechanism, we shall approach the problem from the perceptual angle. In this experiment we shall test the hypothesis that the perception of duration variations in the onset is less accurate than the perception of such variations in the nucleus and/or coda, which do contribute to the weight of the syllable. To test our perceptual hypothesis, subjects adjusted the duration of a comparison stimulus (white noise burst) so as to be identical in duration to a reference syllable, in which either the onset consonant, the vowel or the coda consonant was changed in duration.

Method

The basic reference stimulus was the Dutch word *sas* /sAs/ 'good humour', synthesised from diphones (LPC10 diphones: 5 formants/bandwidths, 10 ms frame duration, speaker HZ [7]). By deleting selected frames, the durations of onset consonant, vowel, and coda consonant were set at 100 ms each. Six additional reference syllables were derived from the base stimulus by lengthening or shortening each of the three segments by 30 ms. The resulting set of seven reference stimuli is specified in table I.

Table I: Stimulus types with durations (in ms) of their constituents.

stim.	onset	nucleus	coda	total
1.	100	100	100	300
2.	**70**	100	100	270
3.	**130**	100	100	330
4.	100	**70**	100	270
5.	100	**130**	100	330
6.	100	100	**70**	270
7.	100	100	**130**	330

The rise time of the onset /s/ and the decay time of the coda were set at 10 ms, so as to avoid audible clicks.
Since we wanted subjects to estimate the duration of the entire reference syllable we did not want a comparison signal that encouraged them to attend to the vowel only. Because the /s/'s at the beginning and end of the syllable (which should definitely be taken into consideration when determining the duration) are noise-like, we opted for a noise burst, rather than some periodic sound, as the comparison stimulus.
Subjects were 24 native speakers of Dutch, with ages between 20 and 40, and without any self-reported hearing deficiencies. None received any payment. The experiment was held in the same sound proofed booth as the production experiment. Subjects adjusted the duration of a white noise burst that was presented 300 ms after the reference stimulus (offset to onset). The intensity of the white noise burst was set at a loudness level that was equal to that of the reference syllable. Reference-comparison pairs were repeated with 1000 ms silent intervals. During this interval the subject adjusted the duration of the noise burst, the effect of which could be heard during the next repetition of the pair. At the initiation of each trial the duration of the noise burst was close to 0 ms. Subjects were allowed to adjust the burst as often as they liked. Once they were satisfied that they had obtained the best possible match, the result was stored and the next trial initiated. Stimuli were presented to a subject in one of three different random orders. Each reference value was presented twice with the second half of the materials in reversed order.

Results and discussion

Not all subjects were able to perform their task with reasonable consistency. Subjects whose adjusted durations in the first and second presentation of the same reference stimulus correlated below r=.50 (excluding 70 and 130 ms onsets) were eliminated from the data set. Seventeen subjects passed the selection filter. The overall means of the noise durations set by these subjects were calculated for each of the categories in table I. The results are presented in table II.

Table II: Mean adjusted durations (set, in ms) and standard deviations (sd, in ms) of white noise burst as a function of the temporal structure of the reference syllable.

Variable	Onset		Nucleus		Coda		Ref.
Duration	Set	sd	Set	sd	Set	sd	dur.
-30 ms	235	61	195	55	202	59	270
base	237	50	237	50	237	50	300
+30 ms	251	47	313	70	260	60	330

We observe, first of all, that comparison signals are generally adjusted to a 20% shorter duration than the physical duration of the reference syllable. This effect, which is irrelevant to our purposes, is quite probably due to time-order bias, by which the duration of the second member of a pair of sounds is auditorily overestimated (hence underestimated in a reproduction task) [8,9]. The standard deviations, between 15 and 20% of the reference syllable duration, are sizeable but not unlike those found in earlier speech data [10, and references there]. When the reference stimulus is short (30 ms decrement) the mean reproduced duration of the comparison stimulus is 26 ms shorter on average than the mean duration reproduced for the base stimulus. Similarly, a 30 ms increment in the reference duration is reflected in a 38 ms longer reproduced duration in the comparison signal (again relative to the base adjustment). The effects have been broken down by position of the target segment in table III.
Crucially, table II shows that 30 ms **vowel** duration manipulations in the reference stimulus are reproduced larger than life in the adjusted duration of the white noise comparison stimulus, which deviate 59 ms on average from the base adjustment, i.e. the effects are perceptually overestimated by a factor 2. Duration variations in the **coda** are reflected more or less faithfully in the adjustments: 29 ms increment/decrement for a 30 ms change in the reference. Changing the duration of **onset** consonants, however, has hardly any effect at all: 8 ms of the 30 ms change is reflected in the reproduced duration, i.e. underestimated by a factor 4.

Table III: Deviations (in ms) from base for varied onsets, nuclei and codas.

Variable	Onset		Nucleus		Coda	
Ref.dur.	Dev	sd	Dev	sd	Dev	sd
-30	-2	23	-42	46	-35	38
base	0	0	0	0	0	0
+30	14	25	76	49	23	32

The effect of position of the target segment on mean over/underestimation of syllable duration are significant: t=-2.16 (p=.017, pairwise, one-tailed) for onset versus coda, and t=-2.29 (p=.012, pairwise, one-tailed) for coda versus nucleus.

These are clearly the results we were looking for: the perceptual effect of a change in duration of the syllable onset are negligible. The incapability of subjects to perceive variations in the duration of the onset, while they can hear such variations in nucleus and coda, explains why the onset does not contribute to the perceived weight (i.e. duration) of the syllable.

Interestingly, duration variation in the nucleus exerts a stronger influence on perceived syllable duration than the same variation in the coda segment. This asymmetry raises the possibility that segment intensity (or sonority) plays a role in the determination of syllable weight as well. The asymmetry also ties in with the observation that there are numerous languages in which not all two-segment rhymes (i.e. nucleus plus coda combinations) count as heavy. These languages typicaly mark syllables with long vowels (-VV) as heavy and closed syllables (-VC) as light.

4. CONCLUSION

We have found that there is nothing in speech production (experiment I) that justifies the insignificance of syllable onsets for stress assignment rules. A perception experiment (experiment II), however, showed that the weightlessness of the syllable onset does have a phonetic basis after all. Subjects proved virtually unable to hear duration variation in the syllable onset, whereas they did respond to similar variation in the syllabic nucleus, where the stimulus variation was in fact exaggerated by a factor 2, and in the coda, where the stimulus variation was reproduced faithfully.

We do not yet understand **why** manipulations of the onset are not responded to. This is the subject of a series of psychophysical experiments that will be carried out in the next three years.

The results of our perceptual experiment independently motivate the phonological view that syllable onsets are weightless, and are consequently ignored by stress assignment rules. This means that languages whose stress rules refer to the syllable onset, are highly marked, but it does not necessarily rule out the existence of such languages completely. In fact, Davis [10] reports that five such languages exist, but they are - predictably - small in number. With the exception of one (Piraha, an Amazonian language) all are more or less genetically related Australian languages. More recent reanalyses adequately account for the Piraha stress system [3,12] that no longer recognise the weight of syllable onsets. Similarly, Goedemans [13] reanalysed some of the Australian languages within the common framework of metrical phonology, and argues that the phenomena can be captured equally well or even better by stress rules that refer only to nucleus and coda.

ACKNOWLEDGMENT

This research was carried out under project # 300-173-026, which is supported by the Linguistic Research Foundation, which is funded by the Netherlands Organisation for Scientific Research, NWO).

REFERENCES

[1] Bolinger, D.W. (1958). A theory of pitch accent in English, *Word*, 14, 109-149.

[2] Heuven, V.J. van. (1993). What is the Smallest Prosodic Domain? In P.A. Keating (ed.) *Papers in Laboratory Phonology III: Phonological structure and phonetic form.* Cambridge University Press, London. (in press)

[3] Halle, M., Vergnaud, J.-R. (1987). *An essay on stress.* MIT Press, Cambridge, MA.

[4] Hayes, B.P. (1981). *A metrical theory of stress rules.* Doctoral dissertation, MIT, Cambridge, MA (distributed by the Indiana University Linguistics Club, Bloomington, IN).

[5] Jongenburger, W., Heuven, V.J. van (1991). The distribution of word-initial glottal stop in Dutch. In F. Drijkoningen, A. van Kemenade (eds.) *Linguistics in the Netherlands 1991*, John Benjamins, Amsterdam, 101-110.

[6] Quené, H. (1989). The influence of acoustic-phonetic word boundary markers on perceved word segmentation in Dutch, doct. dissertation, Utrecht University.

[7] Bezooijen, R. van, Pols. L.C.W. (1993). Evaluation of text-to-speech conversion for Dutch. In V.J. van Heuven, L.C.W. Pols (eds.) *Analysis and synthesis of speech: Strategic research towards high-quality text-to-speech generation.* Mouton de Gruyter, Berlin, 339-360.

[8] Woodrow (1951). Time perception. In S. Stevens (ed.) *Handbook of experimental psychology.* J. Wiley & Son, New York.

[9] Broecke, M.P.R. van den, Heuven, V.J. van (1982). Perceived duration of auditory stimuli as a function of sequential position and decay time. *Progress Report of the Institute of Phonetics Utrecht*, 7:2, 30-40.

[10] Heuven, V.J. van, Broecke, M.P.R. van den (1983). Perceived Duration in 3-syllable Structures. *Progress Report of the Institute of Phonetics Utrecht*, 8:1, 3-14.

[11] Davis, S. (1985). Topics in syllable geometry. doct. dissertation, University of Arizona.

[12] Davis, S. (1988). Syllable onsets as a factor in stress rules. *Phonology*, 5, 1-19.

[13] Goedemans, R. (1993). Rhymes and riddles, investigations in the phonetics and phonology of the weightlessness of syllable onsets, Master's thesis, Dept. Linguistics, Leiden University.