The stress pattern of Dutch monomorphemes

Van Oostendorp, M.; Voeten, Cesko

2015

The stress pattern of Dutch monomorphemes: a pilot empirical study

CESKO VOETEN

Radboud University Nijmegen & Meertens Institute (Amsterdam)

MARC VAN OOSTENDORP

Meertens Institute (Amsterdam) & Leiden University

Because the stress system of Dutch is mixed between free and bound stress, Dutch metrical stress theories have traditionally been plagued by the issue that any given theory has different opinions on which patterns are rule-governed and which patterns are lexically-stored exceptions. We present the results of a-large-scale study (1,774 participants × 307 words; 76,782 cells filled) which used exclusively nonsense words, thereby ensuring that it probes the actual phonological grammar, rather than any lexically-stored stress patterns. The results show that, when ignoring author-specific 'single-problem' postulations, the stress system of Dutch matches well with the concensus in the literature: iterative right-to-left moraic trochees with weight-by-position, whereby tense vowels are metrically short, and where lax ultimas in particular may be extrametrical (cf. Gussenhoven 2014). We also observe an interesting novel finding, viz. that sonorant codas appear to be considered heavier than obstruent codas. We thus show that this novel methodology of using nonsense words in a large-scale study can not only put the literature to the test, much like a meta-analysis does, but can also unearth genuinely novel effects hitherto not considered.

1. Introduction

Dutch has a mixed stress system, meaning that stress is partly rule-governed and partly free. An unfortunate consequence of this fact is that the literature that has aimed to describe the rule-governed part of the stress phonology of Dutch is based on a precarious empirical basis, since any stress pattern that does not conform to the rules of any given theory can simply be analyzed as anexception. Put differently, what one theory might analyze as a genuine stress pattern of Dutch another theory might analyze as an accidental historical contingency, part of the set of lexically-stored fixed stress patterns. Because of the many uncertainties in the literature regarding what is (or should be) the rule and what is (or should be) an exception, empirically studying the stress system of Dutch is not a simple task: to any analysis based on real words, objections can be raised that, from any given theoretical point of view, it analyzes exceptions as being rules and/or rules as being exceptions.

Such discutability can be prevented, however, when an empirical study is based on *nonsense* words – by definition, these words are not present in the lexicon of any given participant, and will therefore test exclusively the rule-governed intuitions that participant will have. We present a study of nonsense stress-assignment data that was undertaken as part of the Dutch *Language Portal* project (see *www.taalportaal.org*, in English) at the Meertens Institute, Amsterdam. The data we present consist of stress assignments over 307 bi-, tri-, or quadrisyllabic nonsense words by 1,774 participants – though it should be noted that many participants did not fill in the entire survey, yielding 'only' 76,782 filled cells. By virtue of its exclusive use of nonsense words, this extremely-large-scale study probes only the participants' productive phonological grammars, thereby giving an objective measure of the *phonology* of Dutch word stress,rather than modelling potential historical accidents.

The structure of this paper is as follows. In Section 2, we present the various theories to be tested. In principle, Section 2 is aimed at coming up with a global overview of the field that is Dutch metrical stress theory; in order to do justice to author-specific peculiarities which, we believe, deviate from the currently-established assumptions in the field, we delegate specific solutions to specific problems to Section 2.4, where we list them by author. Section 3 is where our analysis of the empirical facts is explained, presented, interpreted, and discussed with respect to the concensus analysis with which we end Section 2. Section 4 then interprets the findings in metrical-stress-theoretical terms and the implications for specifically the phonology of Dutch. Section 5, finally, concludes the paper with a discussion of the general theoretical and methodological implications of the study.

2. Established theories on the stress system of Dutch

Classic theories of Dutch metrical stress usually start out with iterative right-to-left moraic trochees (Lahiri & Koreman 1988). Furthermore, there is concensus in the literature that main stress is limited to a three-syllable window at the right edge of the phonological word (e.g. Kager 1989, but see also Van Oostendorp 2012). Thirdly, some information must be stored lexically in at least some cases:

- at least one member of minimal pairs such as *kánon*/*kanón*[1];
- stress patterns violating the three-syllable window: *Wágeningen, Schéveningen* (note that these have also been analyzed as opaque prosodic compounds *waag/scheef* + *eningen*, in which case these forms are actually regular via compound stress; van Oostendorp 2012)*,* *álgoritme*;
- final stress on light (monomoraic) syllables. We leave undecided whether what is stored is the stress pattern of such words as a whole (most authors, e.g. Trommelen & Zonneveld

1989) or a catalectic mora (Nouveau 1994);

- some authors (e.g. van Oostendorp 2012) claim that *all* stress should be stored lexically; this will be discussed below.

Fourthly, it is assumed more often (van der Hulst 1984; Trommelen & Zonneveld 1989; Kager 1989) than not (Booij 1995; Langeweg 1988) that Dutch is quantity-sensitive. Furthermore, most authors implicitly assume weight-by-position (henceforth WBP) for coda consonants here (e.g. Kager 1989). Fifthly, Dutch stress analyses generally assume high-ranked *CLASH (Gussenhoven 2009) and End Rule Final (e.g. Kager 1989).

It follows from the above that the default position for the stress accent is generally the penultimate syllabe, unless the ultima is heavy, in which case it is the ultima that receives stress. In order to be able to explain antepenultimate stress, or penultimate stress with a heavy ultima, a lexically specific instance of NONFINALITY can be used under an Optimality-Theoretic analysis (Nouveau 1994). Furthermore, preantepenultimate stress and antepenultimate stress with a heavy ultima or penult can be explained by augmenting the 'regular' NONFINALITY, which is syllable-based, with a NONFINALITY2 constraint, which renders an entire foot invisible to the stress algorithm (Nouveau 1994). Of course, this knowledge must be stored in some fashion – depending on one's theory, it may be stored (a) as a generalization over a certain type of stratum of words, (b) as a specification added to the lexical entry of individual words, (c) as lexically stored stress (in which case we do not need NONFINALITY(2)), or (d) in some other way.

Furthermore, it is uncontested that stress cannot be borne by schwa. For this reason, underlying schwa tends to attract stress to the syllable preceding it (van der Hulst 1984). This has been refined by Kager (1989), who assumes that schwa has no weight of its own and is hence not capable of projecting a full syllable, which causes any onset consonants of a so-called SCHWALLABLE to be metrically adjoined to the coda of the preceding syllable (as a so-called SCHWAPPENDIX), hence adding to that syllable's weight, making it more attractive towards stress.

This refinement (known as Kager's SCHWALLABLE RESTRICTION) makes it possible to explain the lack of stress of the type **sauríër*: this word has no (underlying) consonants between *i* and schwa, in which case the *i* syllable does *not* receive any additional weight from the schwallable (since the latter has no onset to adjoin to the former's coda). In this case there is insufficient material present with which to form a foot (viz. only a single mora) and hence this syllable must be skipped – due to SYLLABLE INTEGRITY (Kager 1999) – in order to obtain a single foot (*sau*) for this entire word. Van der Hulst (1984), on the other hand, would 'blindly' and incorrectly assign stress to the syllable preceding the schwa.

It is also uncontested that Dutch allows for various possibilities for extrametricality, though only where necessary. Rougly speaking, the main facts are as follows: the Dutch syllable consists of maximally two 'endogenous' consonants (one if the vowel is tense or a diphthong), on top of which an unlimited number of coronal obstruents (i.e. [s,t]) is allowed; furthermore, at the end of the phonological word a single optional extrasyllabic consonant is allowed for (van Oostendorp 2000). Extrasyllabicity is by definition a property of the individual word and hence does not require any special storage facilities; as concerns extrametricality, authors are divided. We will assume that, considering the restrictions mentioned imposed upon extrametricality, in patricular the observation that it may only be applied when this is required for the syllable structure to be well-formed, the type of extrametricality mentioned thus far is rule-governed and hence not in need of special storage facilities. Note that there exists a second type of extrametricality which is more 'arbitrary' (cf. Trommelen & Zonneveld 1989), in the sense that certain words apparently arbitrarily (need to) exclude a final syllable from prosodification in order to arrive at the correct stress pattern for these words; this type of extrametricality does require being stored, of course.

Note that we limit the discussion to monomorphemic words. The reason is that only these words make it possible to truly reveal the stress pattern of Dutch to the extent to which it is governed by rules. This is because (a) longer affixes (i.e. consisting of at least a single whole syllable) usually either carry their own prespecified lexical stress pattern or contain a schwa, which disrupts the regular stress pattern, and (b) compound stress works completely differently altogether, creating an extra stress line on top of the compound's constituents, thereby overriding the End Rule for main stress, which for compounds is always left-dominant (Trommelen & Zonneveld 1989; Kager 1989). Thus, we are limited to (a) monomorphemes and (b) simple inflected forms such as the -s plural, the latter of which could be considered dangerous territory because the inflectional consonant might count for syllable weight or might be extraprosodic. To sum up, it is only the monomorphemic words of which we can be wholly certain that they exemplify the 'true', i.e. unconfounded, stress pattern of Dutch.

2.1 Vowel quality and quantity

The Dutch vowel quantity system appears to be rather bizarre, in that tense long vowels appear to be considered light rather than heavy. This 'problem' has been tackled from various angles. Trommelen & Zonneveld (1989) deny that there is a problem in the first place: to them, length is underlying, and their stress analysis is modeled around this assumption. Kager (1989) assumes that the reason why tense long vowels appear to be treated as light is that Dutch vowel quantity is computed at the X-slot level, where long vowels receive a single V-slot which just happens to project two moras. This, however, is inconsistent with a multitude of segmental research, which consistently shows that these vowels, with the exception of (a), have been diphthongs (van der Harst 2011), even at the underlying level (Voeten 2013), since 1960 (Voortman 1994; Van de Velde 1996), and hence are melodically complex after all (i.e. projecting *two* V-slots). A different solution is proposed by van Oostendorp (1995, 2000), who assumes that tense vowels are not at all long underlyingly, but are in fact short. Independent evidence for this position is found by Gussenhoven (2009), who shows that these long vowels are actually not long at all, except solely when stressed. This is not at all unique for Dutch, but rather is the embodiment of the cross-linguistically widely attested SWP (STRESS-TO-WEIGHT PRINCIPLE) constraint.

This analysis manages to explain in one fell swoop both why the Dutch quantity system appears to be so highly odd and why the long vowels are actually not always long in reality; in addition, it also shows the role of the general constraint that is SWP. Still problematic, however, is the fact that the tense MID vowels are not just long, but are actually diphthongs, even at the underlying level (Voeten 2013).

2.2 The three-syllable window

A controversial component of the metrical stress analysis as sketched up to now is the three-syllable window. Briefly summarized, main stress is clearly limited to the final three syllables of the phonological word (with a few highly select exceptions, see van Oostendorp 2012 for an inventory), and is allowed considerable leeway within this three-syllable window through means of local extrametricality (accepted by many – Kager 1989; Trommelen & Zonneveld 1989; but not Langeweg 1988 – for the final syllable; less accepted for the final two syllables or the final foot – only Nouveau 1994 – in which cases it is usually assumed – Trommelen & Zonneveld 1989 – that the whole stress pattern is lexically stored altogether). A major problem with this analysis is that it is rather uninformative, given the grave scarcity of monomorphemic words consisting of more than three syllables – which, if found, can furthermore also be analyzed as opaque prosodic compounds of two cranberry morphemes (van Oostendorp 2012).

To van Oostendorp (2012), this uncertainty regarding the status of the three-syllable window (low predictive power + high degree of freedom within the window) is a reason to reject the three-syllable window as being phonological knowledge. That is, while the strong generalization that main stress is usually limited to the final three syllables is not rejected, van Oostendorp considers this a historical accident rather than an *active* phonological constraint within the grammar of the Dutch native speaker. With the rejection of the three-syllable window – based on the lack of productive evidence speaking in its favor – its three formal counterparts of (i) local extrametricality for syllables, (ii) local extrametricality for feet (both for stress located 'too far' from the right edge), and (iii) catalectic moras (for stress located 'too far' TO the right edge, in words like *kopíé*) also appear to lose their theoretical motivations. Consequently, according to van Oostendorp, stress appears to be no longer rule-governed – in this vision, what we have been modeling as linguists is not so much the *synchronic* productive grammar of the Dutch native speaker, but rather the combination of various types of internal and external influences which together constitute the *history* of Dutch.[2]

Similar arguments have been raised against the assumption that Dutch is quantity-sensitive. While it is uncontested that schwa is maximally light and diphthongs are maximally heavy (van der Hulst 1984; Kager 1989; van Oostendorp 2012), less ambiguous evidence has been unvealed for the intermediate positions, in particular the contrast between tense-and-light VV and lax-and-heavy VC. It should be noted that this point of discussion has been relatively underemphasized within the literature; the generally-accepted analyses generally take no issue with assuming the regular four-way quantity contrast of the *schwa–VV–VC–diphthong* type[3]; due to the apparent silent consensus on this topic, this is also what we will assume.

2.3 Conclusions: the consensus theory

A synthesis of the considerations thus far leads to an analysis of Dutch monomorphemic stress as using iterative right-to-left moraic trochees with weight-by-position, End Rule Final, optional extrametricality of at least a final syllable and to some authors also a final foot. Furthermore, it should be assumed that tense vowels are underlyingly short and are analyzed as if they were monophthongs; they are lengthened only when stressed by means of SWP. In addition, schwa appears to have no weight (i.e. mora) of its own, for which reason onset consonants of schwallables are adjoined to the preceding syllable in the form of a schwappendix. Everything that this analysis does not regulate is stored in some way.

2.4. Noteworthy deviations per author

2.4.1 Van der Hulst (1984)

Van der Hulst has a foot type of his own, which I will call the *VDH*-foot. In principle, this foot is a syllabic trochee, except when the weak syllable of this foot is heavier than the strong one; in that case, stress is shifted to the foot's second syllable. If we, in accordance with van der Hulst (1984), construct this foot iteratively from right to left, the resulting pattern matches a priori with a moraic trochee with local final extrametricality in those cases in which the penult is heavy or superheavy. This deserves explicit mention, because this gives rise to a third type of extrametricality: next to the full extrametricality of van Oostendorp's (2000) extrasyllabic consonants and the 'arbitrary' extrametricality created by local invocation of (any type of) NONFINALITY (e.g. Nouveau 1994), the VDH-foot creates another, wholly predictable, type of extrametricality, which might perhaps overlap fully with those cases in which local NONFINALITY is commonly employed.

2.4.2 Langeweg (1988)

A unique deviation of most of what has been discussed so far is offered by Langeweg (1988). She proposes a bidirectional stress pattern for Dutch, starting with a single VDH-foot at the right word edge, followed by a single unbounded[4] syllabic trochee from left to right, where, optionally, *CLASH is allowed to block this trochee's secondary stress if necessary. Though this appears highly unconventional, according to Langeweg it works surprisingly well, and the exceptions to this algorithm are in fact mostly the standard well-known exceptions which are problematic in any case.

2.4.3 Gussenhoven (2009)

Gussenhoven (2009) is special in that he assumes a specific variant of WBP which is reserved only for syllables *following* the one receiving primary stress. We are not convinced that this is necessary, though. First and foremost, this analysis runs into the problem that it seems dangerously circular. Note further, however, that the reason *why* Gussenhoven (2009) posits this limited WBP is only to get rid of an unwanted secondary stress in words like **àrmáda*; we believe, however, that this initial secondary stress can also be just as easily removed by *CLASH.

3. The empirical study: nonsense words

3.1 Introduction and method

As explained in the introduction, we conducted a study that alleviates the objections against most prior studies by making use of *nonsense* words, which to our knowledge has only been done before by Nouveau (1994). Her study, however, uses a mere 24 stimuli and focuses mostly on the (arguably immature) stress systems of 40 children (in his nonsense condition) and, as a control group, 20 adults. Our study, on the other hand, consists of 307 phonotactically different di-, tri-, or quadrisyllabic nonsense words run by 1,774 participants (with 76,782 cells actually filled). This is important, because, as we will detail below, the between-participant variation is rather impressive; by countering this obstacle with a fairly large number of participants, we believe our study gains significant predictive power by estimating significantly more stable parameters.

One important limitation of this study's set-up must be taken into account: the methodology employed was an on-line web survey among laymen who did not master the phonetic alphabet. Hence, stimuli were presented orthographically. The unfortunate consequence of this manner of presentation is that no schwas could be included in the stimuli. Of course, the stimuli have been crafted as such that no accidental schwas could be construed out of orthographical readings of <e>.

Because of practical as well as theoretical considerations, it was decided to exclude the factor 'participant', which had 1,774 levels, from the analysis. As for the practical part, the number of participants was so large that an analysis with this factor entered as random factor was simply impossible purely on computational grounds. In addition, between these participants the entire spectrum of variation in stress patterns was accounted for: when taking only the disyllables as a sample and counting per participant the proportion of stimuli to which final stress had been assigned, practically all proportions between 0 and 1 were represented, without any clear clusterings. Because of this total continuity of variation, combined with the aforementioned computational constraints, it was decided to allow the variation between participants to be random (i.e. error) variance.

In the following, the results will be discussed. The structuring is as follows. The global results are discussed first. Initially, these were split out per potential stress placement (ultima, penult, antepenult (not discussed)), but the effects turned out to be similar for both categories examined (ultima and penult) and are therefore discussed globally. Next, the results are discussed again, but this time split by word type (disyllabic, trisyllabic, quadrisyllabic (not discussed)). Thirdly, a complete regression model is presented for only the disyllables (the tri- and quadrisyllables could not be included in the same model due to methodological limitations). Finally, the implications for Dutch metrical stress theories are discussed.

Note that, in line with a Dutch tradition (e.g. van Oostendorp 2000)), tense vowels are referred to as *A-vowels*, and lax vowels as *B-vowels*. Lax vowels always occur in closed syllables only; A vowels occur in open syllables in this experiment, except sometimes when they are final.

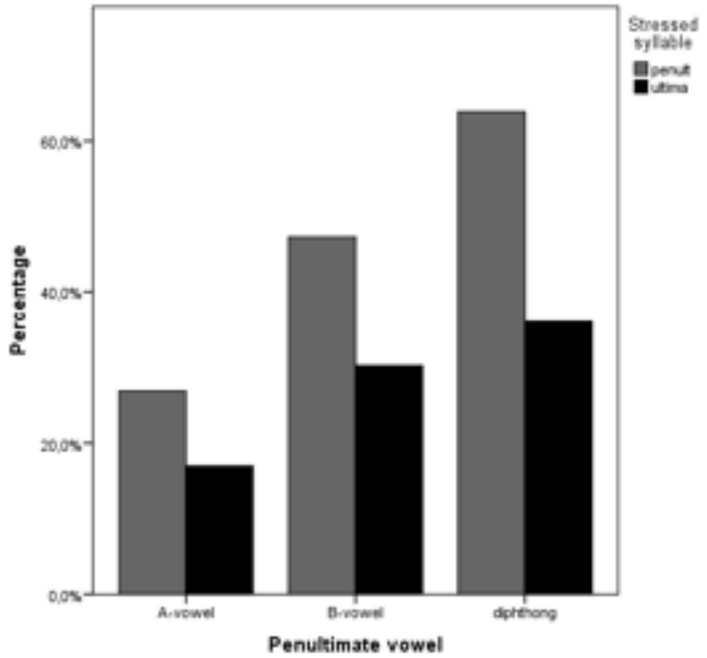
3.2 Global findings

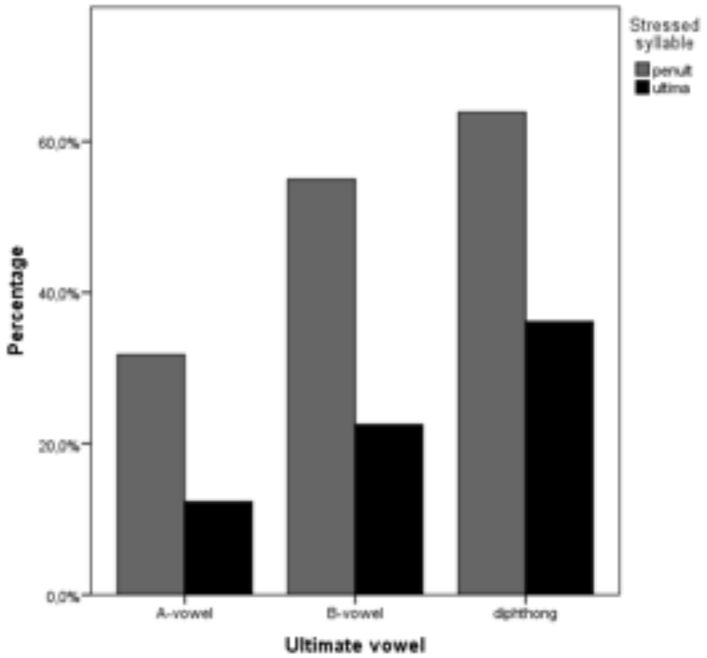
A first inspection of the raw numbers shows the following tendencies. First, when the ultima contains an A-vowel, there appears to be a strong tendency for stress to be penultimate – this follows naturally when one assumes that an ultimate A-vowel projects only a single mora and hence consists of insufficient material to be able to constitute a whole foot on its own. When we add a coda to such an A-ultima, this effect virtually remains: only *slightly* more stress is pulled towards the ultima in this case. This same effect holds for ultimate B-vowels: these show a *small* tendency to delegate stress towards the ultima, but for the most part (~66:33) they stick to the penultimate default. A strong effect of ultimate diphthongs is found: they most definitely attract stress.

The effects found are significant, both for the ultima per stress type (ultimate, $\chi^2(6) = 11430.852$; $p < .001$; penultimate, $\chi^2(6) = 17088.893$; $p < .001$; antepenultimate, $\chi^2(6) = 6556.462$; $p < .001$; preantepenultimate, $\chi^2(6) = 2369.855$; $p < .001$) and as a whole ($\chi^2(6) = 32977.998$; $p < .001$), and for the penult per stress type (ultimate, $\chi^2(4) = 18678.000$; $p < .001$; penultimate, $\chi^2(4) = 28354.000$; $p < .001$; antepenultimate, $\chi^2(4) = 13485.000$; $p < .001$; preantepenultimate, $\chi^2(2) = 5035.000$; $p < .001$) and as a whole ($\chi^2(4) = 65552.000$; $p < .001$). The data on which these findings are based are available on-line, at the website of the project: *www.taalportaal.org*.

3.3 Results for the disyllables

The graphs below show the distribution of stress per vowel type per syllable. A clear penultimate default can be observed, which must only be qualified by mentioning that an ultimate diphthong attracts stress towards the ultima.



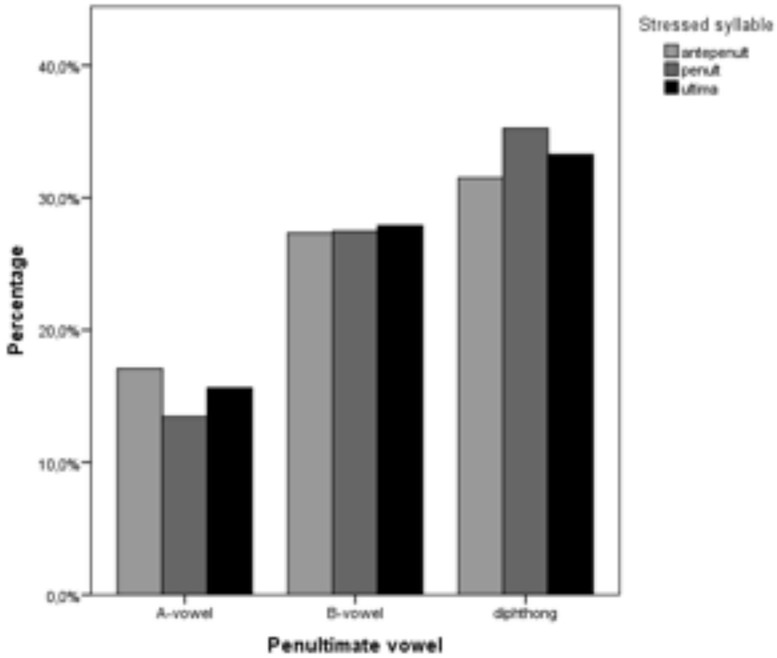


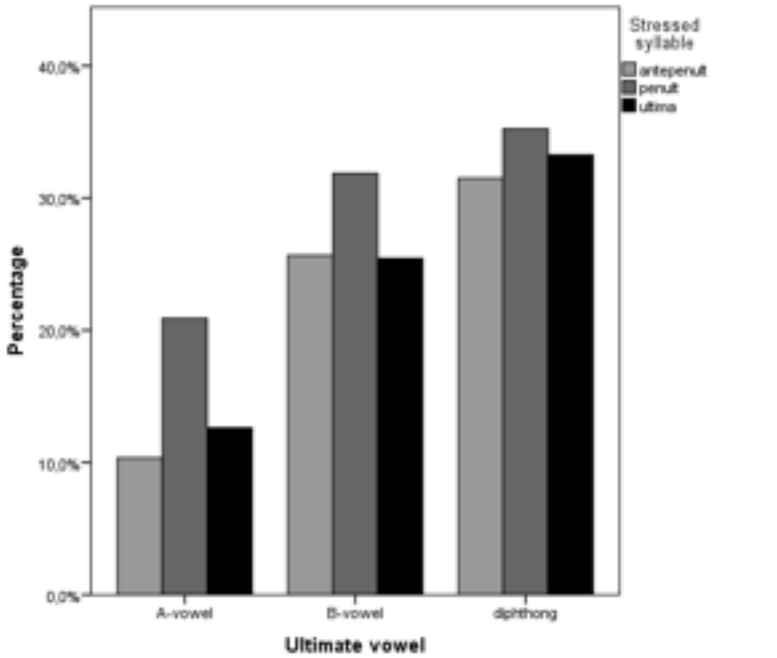
A more precise inspection of the numbers gives rise to the following findings. When the ultima contains an A-vowel, stress is again nearly never ultimal when no coda is present. If there is a coda, the stress distribution is about 66:33: a sonorant coda to the ultima attracts stress (~66%), an obstruental coda much less so (~33%). When the ultima contains a B-vowel, stress appears to remain penultimate no matter the type of coda consonant, unless the coda is complex: in that case, the stress distribution is about 45:55: ~45% of the stresses move towards the ultima, the other ~55% remains penultimate in spite of the complex coda. Ultimate diphthongs attract stress with about the same approximate 45:55 distribution, despite any coda consonants (though it does appear that obstruents in the ultima show the same non-stress-attracting effect as with the A-vowels).

The effects found are significant, both for A-vowels ($\chi^2(2) = 1157.802$; $p < .001$), B-vowels ($\chi^2(2) = 260.018$; $p < .001$), and diphthongs ($\chi^2(2) = 46.792$; $p < .001$) and as a whole ($\chi^2(3) = 400.476$; $p < .001$).

3.3 Results for the trisyllables

For the trisyllables, the effect of diphthongs is the same as for the disyllables: a diphthong in the penult or ultima attracts stress towards that same syllable. The same goes for a lax vowel, but to a much less degree. Ultimate A-vowels again appear to delegate stress towards the penult. The penult itself does not appear to exert any effect, save for the observation that a diphthong in the penult appears to detract stress from the ultima (which is obvious, given the above) and the lesser effect of stress attraction by lax vowels. The structure of the antepenult appears to be irrelevant altogether; in the interest of space, graphs segmented by this syllable's vowel type are hence not included.





From the numbers, ultimate A-vowels appear to delegate stress towards the penult if they are open (as discussed earlier). When a coda is added, this effect disappears and the ultima attracts stress with a ~55:45 distribution. Note that in this case there is no difference between obstruents and sonorants. Ultimate B-vowels barely show an effect, apart from the fact that they delegate stress away from the ultima to a prior syllable with a ~66:33 distribution. Again, no difference can be observed between obstruental and sonorant codas. Ultimate diphthongs still attract stress, though a division appears to be created: ~45% of the stress is attracted towards the ultimate diphthong itself, ~20% towards the penult, and ~35% to the antepenult. In both cases, a potential explanation could be that the ultima might constitute a single foot on its own, and that, subsequently, a second right-to-left trochee is created, which then seems to receive main stress.

The effects found are significant, both for A-vowels ($\chi^2(2) = 2379.190$; $p < .001$), B-vowels ($\chi^2(2) = 387.939$; $p < .001$), and diphthongs ($\chi^2(2) = 60.968$; p < .001) and as a whole ($\chi^2(3) = 2301.505$; $p < .001$). Again, the data are available on-line.

3.4 Regression analysis of the disyllables

As an addition to the overall chi-square tests that we ran, the disyllables were analyzed statistically by means of a logistic regression. The tri- and quadrisyllables could not be included into the same model because of methodological limitations (including *n+1*-syllable words will, a priori, create complete separation for the *n*-syllable words, invalidating the analysis). The dependent variable was, of course, the syllable which was stressed by the participant, dummy-coded so that the internal prediction to be made by the model was whether stress would fall onto the *ultima* or not (remember that we only included the disyllables, so if stress was not ultimate, it automatically was penultimate). Predictors were entered using the backward-stepwise method. The analysis is reported in the table below, by reporting all predictors that were included into the eventual model, sorted by effect size:

Factor	B	Exp(B)
Intercept	-1,507	0,222

Ultima has complex coda	3,185	24,166
Ultima has simplex sonorant coda	2,084	8,035
Ultima has B-vowel	-1,594	0,203
Ultima has diphthong	1,353	3,869
Ultima has simplex obstruent coda	1,328	3,773
Penult has diphthong	-,539	0,583
Penult has B-vowel	-,333	0,717

R^2 = .189 (Cox & Snell), .259 (Nagelkerke). Model $\chi^2(5) = 1617.550$, $p < .001$. All predictors $p < .001$.

Briefly summarized, firstly it became apparent that the penultimate default was extremely strong: in absence of all other predictors, the chances of penultimate stress were about 4.5 (i.e. 0.222^{-1}) times as large as were the chances of ultimate stress. As concerns the included predictors, the clearly strongest effect was that of having a complex ultima: this increased the chances of final stress by a factor of ~24 (Exp(B) = 24.166). Sonorants (Exp(B) = 8.035) and obstruents (Exp(B) = 3.773) showed a similar stress-attracting effect (challenging the above findings), though these effects were *much* smaller.

Ultimate B-vowels did not attract stress: the chances of final stress were almost five times as small (i.e. 0.203^{-1}) when the ultima contained a B-vowel. It should be noted here that, in order to reduce the model's complexity, no interactions were included in the model; it goes without saying, however, that having a B-vowel is concomitant with having a coda, which means that it cannot be concluded that B-vowels *reject* stress, since the effect of having a coda must still be added here. What *can* be concluded is that having a coda exerts a larger effect for A-vowels and diphthongs than for B-vowels, which does not seem unreasonable, given the fact that for the former two groups codas are optional extras, which is of course not the case for B-vowels.

The effect of a diphthongal ultima was about as large as the effect of having an obstruental coda; both attracted stress (Exp(B) = 3.869 resp. 3.773). Finally, penultimate diphthongs attracted stress away from the ultima towards the penult (Exp(B) = 0.583), as did penultimate B-vowels (Exp(B) = 0.717).[5]

4. Discussion

On the basis of the results, the following can tentatively be concluded.

First, the global tendencies show in any case that an ultimate A-vowel practically never attracted stress, i.e. did not constitute a foot on its own. This is in agreement with all literature that assumes that A-vowels are underlyingly monoraic, as well as with Voeten (2015), who assumes that A-vowels are obligatorily parsed *syllabically* into a trochee. Theories assuming underlying length for A-vowels, however, *in combination with quantity-sensitivity* (such as Trommelen & Zonneveld 1989), are faced with either a problem, or a highly regular scenario in which to deploy local NONFINALITY.

Adding a coda to an A-vowel did *not* cause stress attraction towards that A-vowel. This agrees with Voeten's theory that A-vowels are parsed syllabically, as well as for theories which would have saved themselves in the previous paragraph by using NONFINALITY. The more 'classic' theories (van der Hulst 1984, Kager 1989) assuming quantity-sensitivity *in combination* with weight-by-position, however, have a problem here, since in this view a (monomoraic) A-vowel combined with a coda should provide enough material with which to construct a foot, which should hence generate stress. A potential solution might be to assume that Dutch has no weight-by-position whatsoever: in that case a syllable containing an A-vowel plus a coda still has only a single mora, and hence no stress. We will discuss the evidence for this possibility further down.

A very strong effect that was found, is stress attraction by diphthongs, which is well-known from the literature. Under the standard metrical analysis of Dutch, diphthongs are the only vowels that are underlyingly bimoraic – the fact that they attract stress thus seems very strong evidence that Dutch is quantity-sensitive. This is not necessarily the case, however: a unique heaviness for diphthongs is also possible in a quantity-insensitive language, by ranking WSP >> FTBIN(σ) >> FTBIN(μ) (FTBIN(σ) >> FTBIN(μ) causes a syllabic trochee to be preferred over a moraic one, but higher-ranked WSP causes underlyingly long vowels to constitute monosyllabic feet despite this preference for a syllabic trochee).

Having made this observation, we are now faced with a choice: does Dutch have weight-by-position, thus saving the bimoraic minimum but allowing ourselves to be stumped by the apparent lightness of B-syllables plus coda, or doesn't it?

At this point, let us take a look at the results for the disyllables. These showed that A-vowels *did* show a stress-attracting effect of codas, but only when these were *sonorant*. The same held for B-vowels, but here it were not specifically the sonorant, but rather the *complex* codas showing an effect. From the regression coefficient it turns out that especially this latter factor plays a truly enormous role. Furthermore, the logistic regression *did* actually find stress-attracting effects of codas, especially when assuming the ambisyllabic analysis suggested in footnote 5. The only oddity that needs to be remarked with this analysis is that it did not appear to hold for *ultimate* codas of B-vowel syllables – after all, a B-vowel in the ultima, which of course always has a coda, *hampered* the stress-attracting effect of the coda.

All in all, the above can be summarized as follows. Dutch appears to have different levels of weight or quantity-sensitivity, which are captured by the following scale:

A-vowels < diphthongs < sonorant coda < complex coda

It should be noted that B-vowels are not included in the scale, since they are partially covered by the coda conditions, though it should be said that, without adding the coda effect, a B-vowel usually has a lightening effect, especially within the ultima.

The above is, despite everything that we have discussed before, consistent with weight-by-position: more coda consonants add more moras. The effect of B-vowels is at all times wholly compensated by the strong effect of weight-by-position, so we cannot speak of a stress-deterring effect for B-vowels; it *can* be observed, however, that a coda for a non-B-vowel appears to be considered heavier than a coda for a B-vowel. This might be related to the fact that in the case of B-vowels, codas are mandatory, and hence less noticeable; another possible explanation could be that B-vowels are 'secretly' the lightest vowels in all of Dutch, but that this effect has always been masked by the presence of an obligatory coda. Note that this final possibility would mean that Dutch is becoming more surface-true, since A-vowels are long at the surface level (and are even full-fledged diphthongs; Voeten 2013), as are diphthongs, but not B-vowels.

The only further remaining point needing discussion is the stress-detracting effect of B-vowels in the ultima, especially when considered against the same stress-*attracting* effect of *pen*ultimate B-vowels. Only one – obvious – resolution can be offered here: NONFINALITY. Apparently, ultimate B-vowels attract extrametricality, a point already raised implicitly by Gussenhoven (2014), who notes that 'main stress is on the penult, *unless the final syllable is […] heavy and the penult is light*, in which case it is on the antepenult if there is one' (Gussenhoven 2014: 2, emphasis ours).

In sum, the findings are as follows. In Hayesian terms, the correct analysis of the stress system of Dutch consists of iterative right-to-left moraic trochees with weight-by-position. Dutch is unequivocably quantity-sensitive, but there are two oddities. First, A-vowels, which are long at the surface level, appear to behave metrically as if they were short. This can be interpreted as evidence for the classic position that they actually are light (van Oostendorp 1995, 2000; Gussenhoven 2009; Kager 1989 (at the metrical level only)), or for the alternative interpretation that they are obligatorily parsed quantity-*in*sensitively (Voeten 2015). Underlying metrical bimoraicity (Trommelen & Zonneveld 1989), however, is *not* an option for A-vowels. Besides the observation that A-vowels appear to be positioned further down in the weight hierarchy for Dutch, a second oddity is that there appears to be a bias towards sonorants: while syllables containing complex codas are of course the heaviest of all, syllables with simplex sonorant codas appear to be heavier than syllables with simplex obstruent codas by a factor of ~2 (8.035/3.773): apparently, the stress system of Dutch has a preference for [+son] sounds, i.e. vowels and sonorants.

The apparent difference in weight between obstruents and sonorants is as far as we know a new finding with respect to Dutch metrical stress. This finding is exciting for two reasons. Firstly, it further validates our methodology: while up to now, our study has mostly been meta-analytic in nature, comparing key prior theories and testing them on the basis of our data, the sonorant bias we found demonstrates that large-scale nonsense-word studies such as ours can also unearth novel findings. Secondly, this finding in particular is interesting, because a similar obstruent-sonorant divide can be observed in the segmental phonology; two well-known examples are final devoicing – which only applies to obstruents, not sonorants – and diminutive formation – stems ending in a sonorant consonant get a schwa inserted between the stem and the diminutive suffix (e.g. *man-ə-tje*), whereas stems ending in obstruent consonants do not (e.g. *dakje*, as opposed to **dak-ə-tje*). Evidently, the key distinction between obstruents on the one hand and sonorants on the other concerns not only the segmental, but also the prosodic, phonology.

No evidence has been found for the special VDH-foot (van der Hulst 1984), for bidirectional stress assignment on the basis of such a foot (Langeweg 1988), or for a special type of weight-by-position ranging from the main stress onwards (Gussenhoven 2009). However, evidence has been found that Dutch seems to prefer to have ultimate B-vowels be extrametrical, which allows for an explanation of antepenultimate stress (cf. Gussenhoven 2014).

5. CONCLUSION

This paper started out with the observation that the literature is problematic from an empirical point of view, in that any theory using real words has always been free to make its own assumptions regarding what words conform to the rules it postulates and what words are exceptions; a problem which is inherent to analyses of mixed-stress systems. Though our study did not suffer from this problem, having made use of nonsense words rather than actual words, we are pleased to report that our study – even though the modeling part is based on only a subset of the data (only the disyllables) – has largely corroborated the concensus that is present in the literature: right-to-left moraic trochees with weight-by-position, whereby A-vowels are considered light.

We can discern three potential reasons why we have replicated the consensus in the literature, but not any of the author-specific 'one-off' solutions such as van der Hulst's (1984) VDH-foot or Gussenhoven's (2009) special type of weight-by-position from the main stress onward. The first potential reason is that the consensus is there for a reason: in this view, the rule-governed patterns in Dutch are actually quite clear, and the disagreement in the literature on specificities is due to any individual author's desire to incorporate as much of what actually are *exceptions* into his system of rules. If this is the case, then Dutch stress is simply far more free than has been hypothesized before, and the points of disagreement, for which we did not find evidence, only reflect a failure of the respective authors to recognize just how much stress patterns are in fact lexically stored (cf. van Oostendorp 2012). A second potential explanation is that, given the fact that our study dealt with *nonsense* words, the patterns we have found are in fact the most unmarked patterns in the entire language; these patterns then generalize so well to the existing studies simply because these studies happened to include mostly unmarked words (i.e. the studies with which we agree simply happened to base themselves upon highly representative data). A third potential reason for the relative uncontroversy of our findings could be tunnel vision from our part: in this (depressing) view, we have found what we did simply because we, by error of our part, accidentally predestined either the data or the analysis to reveal only what we wanted it to reveal; in other words, had the findings been different, we would probably have rejected the experiment as having been in error in some way or another. Fortunately, however, this option appears unlikely, given the fact that, in addition to the observations known from the literature, we also unearthed a novel effect of our own, viz. the apparent weight distinction between obstruents and sonorants. Had we unconsciously been creating or analyzing the data through a tunnel view, one would not expect us to have found such a non-standardly-assumed effect.

Of the above three possible explanations for the lack of controversiality of our findings, we believe in the first: that we have managed to replicate the consensus because it is there for a reason. This is easiest to see when one keeps in mind our Hayesian framework of analyzing stress in terms of trochees vs. iambs, quantity-sensitivity vs. quantity-insensitivity, etc.: the options offered by the currently-accepted metrical stress framework are highly restrictive, which means that the selection of a wrong parameter for any of these options would be highly noticeable. This means that the framework implicitly guides phonologists into making the most optimal choices for their language of study. In addition, our study was somewhat 'theory-meta-analytic' in nature, in that our purpose has ultimately been to test and choose between the various theories that have been postulated; this, of course, automatically guided our analysis into being expressed in terms of these exact theories, which, per the preceding, turns out to be quite robust.

This paper thus ends on a methodological note: it turns out that a large-scale, nonsense-word-driven study is actually a highly reliable instrument for choosing between various different theories, especially when these theories all couch themselves in a robust framework delivering clear predictions. In addition to being able to test and decide between various theories, furthermore, the method is also able to unearth entirely novel effects which any of the individual competing theories have not managed to consider, *in casu* that sonorants appear to count heavier for weight than obstruents do. Thus, large-scale nonsense-word-driven surveys show themselves to be highly promising for future phonological research.

REFERENCES

Booij, Geert. 1995. *The phonology of Dutch*. Oxford: Oxford University Press.

Gussenhoven, Carlos. 2009. Vowel duration, syllable quantity, and stress in Dutch. In Hansson, Kristin & Sharon Inkelas (eds.). *The Nature of the Word. Essays in Honor of Paul Kiparsky,* 181–198. Cambridge/London: MIT Press.

Gussenhoven, Carlos. 2014. Possible and impossible exceptions in Dutch word stress. In Hulst, Harry van der. 2014. *Word Stress: Theoretical and Typological issues*, 276–296. Oxford: Oxford University Press.

Harst, Sander van der. 2011. *The vowel space paradox: A sociophonetic study on Dutch*. Ph.D. dissertation, Radboud University Nijmegen.

Hulst, Harry van der. 1984. *Syllable structure and stress in Dutch*. Ph.D. dissertation, Leiden University.

Kager, René. 1989. *A metrical theory of stress and destressing in English and Dutch*. Ph.D. dissertation, Utrecht University.

Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

Lahiri, Aditi & Jacques Koreman. 1988. Syllable weight and quantity in Dutch. *West Coast Conference on Formal Linguistics* 7, 286–299.

Langeweg, Simone. 1988. *The stress system of Dutch*. Ph.D. dissertation, Leiden University.

Nouveau, Dominique (1994). *Language acquisition, metrical theory, and optimality: A study of Dutch word stress*. Ph.D. dissertation, Utrecht University.

Oostendorp, Marc van (1995). *Vowel quality and phonological projection*. Ph.D. dissertation, Tilburg University.

Oostendorp, Marc van (2000). *Phonological projection: A theory of feature content and prosodic structure*. Berlin: Mouton de Gruyter.

Oostendorp, Marc van (2012). Quantity and the three-syllable window in Dutch word stress. *Language and Linguistics Compass* 6(6), 343–358.

Trommelen, Mieke & Zonneveld, Wim (1989). *Klemtoon en metrische fonologie*. Muiderberg: Coutinho.

Van de Velde, Hans (1996). *Variatie en verandering in het gesproken Standaard-Nederlands*. PhD dissertation, Radboud University Nijmegen.

Voeten, Cesko (2013). *On the diphthongization of the Dutch long mid vowels: Implications for the phonology of Dutch*. Ms., Radboud University Nijmegen.

Voeten, Cesko (2015). *The interaction between the segmental and the prosodic phonology in the midst of an on-going sound change: Resolving a contradiction in the synchronic phonology of Dutch*. Ms., Radboud University Nijmegen.

Voortman, Berber (1994). *Regionale variatie in het taalgebruik van notabelen. Een sociolinguïstisch onderzoek in Middelburg, Roermond en Zutphen*. Ph.D. dissertation, University of Amsterdam.

Author's address: Meertens Instituut, Joan Muyskenweg 25,1096 CJ Amsterdam, Nederland *marc.van.oostendorp@meertens.knaw.nl*

FOOTNOTES

[1] Note, however, that these classic words both have a different vowel in their first syllables.

[2] Given the enormous complexity of present-day stress analyses of Dutch, we would initially be inclined to agree with this position. However, it should be noted that it is a priori impossible for a language like Dutch to have *no* regulated stress whatsoever, since native speakers do have (strong) intuitions about novel loans or nonsense words. Therefore, at least *some* part of the stress grammar of Dutch must be a genuine phonological grammar, rather than a mere list of lexically stored fixed patterns.

[3] It is not at all clear what the position of superheavy syllables is here: both VV and VC syllables are allowed to add one further non-extrametrical coda consonant, and it is not clear where in the hierarchy such constructions should be placed.

[4] We fail to see why this trochee needs to be specifically UNBOUNDED, since we cannot discern any surface-level-visible difference between a single binary left-to-right trochee and a single unbounded left-to-right trochee.

[5] Note that, in order to remain theory-neutral, we did not take ambisyllabicity into account in the data processing. This means that words of the /pɒ.tɒt/ type have been analyzed as having a *monomoraic* penult, which means that the coda effect that would be ascribed to the penult under an ambisyllabic analysis has not been taken into account. It is possible, then, that the stress-attracting effect of penultimate B-vowels need actually be attributed to an ambisyllabic penultimate coda consonant.