STRESS AND LEXICAL ACTIVATION IN DUTCH

Anne Cutler and Mariëtte Koster

Max Planck Institute for Psycholinguistics
Postbus 310, 6500 AH Nijmegen, The Netherlands
Tel +31 24 352 1377, email anne.cutler@mpi.nl

ABSTRACT

Dutch listeners were slower to make judgements about the semantic relatedness between a spoken target word (e.g. *atLEET*, 'athlete') and a previously presented visual prime word (e.g. SPORT 'sport') when the spoken word was mis-stressed. The adverse effect of mis-stressing confirms the role of stress information in lexical recognition in Dutch. However, although the erroneous stress pattern was always initially compatible with a competing word (e.g. *ATlas*, 'atlas'), mis-stressed words did not produced high false alarm rates in unrelated pairs (e.g. SPORT - *atLAS*). This suggests that stress information did not completely rule out segmentally matching but suprasegmentally mismatching words, a finding consistent with spoken-word recognition models involving multiple activation and inter-word competition.

1. INTRODUCTION

Languages contain many thousands of words, but these words are constructed from a notably limited array of phonetic resources. Words can be distinguished by segmental differences: bellow vs. mellow, or rusty vs. trusty; but in many languages suprasegmental contrasts - variations in pitch, amplitude and duration of syllables - are also used to distinguish one word from another. A common means of distinguishing words suprasegmentally is via tone; Chinese languages make use of this option.

Another way of distinguishing words suprasegmentally is by stress. In lexical stress languages, one syllable of any polysyllabic word has higher stress than the other(s). Some such languages (e.g. Spanish) use stress distinctions which are purely suprasegmental, and orthogonal to segmental structure, while in others (e.g. English or Dutch) suprasegmental and segmental structure are interdependent. This is because the latter language type allows vowel reduction, and reduced vowels can only occur in unstressed position. Thus in English, bellow is stressed on the first syllable, below on the second; trusty on the first and trustee on the second. However, the words bellow and below also differ in the vowel sound in the first syllable: the unstressed first syllable of below contains the reduced vowel [ə] - schwa, while the stressed first syllable of bellow contains the full vowel [ɛ], as in bed. Vowel reduction frequently accompanies stress variation, but it is not a necessary feature of it; for instance, trusty and trustee do not differ in vowel sounds, and are distinguished only by the stress difference. Dutch contains similar variety: *VOORnaam* ('first name'; upper case indicates stress) and *voorNAAM* ('respectable') differ only in stress, while *REgent* ('is raining') and *reGENT* ('regent') differ both in stress and in vowel quality: the second syllables show exactly the same opposition between schwa in the unstressed case and [ɛ] in the stressed case as is found in the first syllables of English *beLOW/BELlow*.

Psycholinguistic studies have examined the relative contributions of segmental and suprasegmental information to the recognition of spoken words, in many languages. One standard way of assessing the role of stress in word recognition is to see what happens when listeners hear mis-stressed words. Studies in English suggest that mis-stressing has little adverse effect on spoken-word recognition unless vowel quality also alters. Thus of all the ways one can slightly alter a word's pronunciation, alteration of vowels in stressed syllables inhibits successful recognition most [1]; mis-stressing of words has no adverse effect on word recognition in noise unless vowel quality is also changed [2]; but mis-stressing with vowel quality change renders word recognition, even without noise masking, very difficult [3] or indeed impossible [4].

In Dutch, however, segmental and suprasegmental mispronunciation have closely parallel inhibitory effects on word recognition. Koster and Cutler [5] demonstrated this using a semantic judgement task, in which listeners see a prime word on a screen, then hear a spoken word, and their task is to judge whether the two are related. Changes in stress (not affecting vowel quality) and changes in vowel quality (not affecting stress) had almost exactly the same deleterious effect on response time (RT) in this task.

In most mis-stressing experiments (as in [1-5]), each stimulus corresponds to only one word candidate, correctly or incorrectly pronounced. Another way to consider mis-stressing, however, is to assess its effect as a mismatch which rules out one word candidate in favour of a competing candidate. Current models of word recognition [6] assume that speech input simultaneously activates multiple word candidates compatible with any part of the input, and a process of inter-word competition ensues to rule out spuriously activated word-forms and ensure that the correct sequence is recognised. Experiments in Spanish [7] show that a stress mismatch is exactly as influential as a segmental mismatch in distinguishing between two potential words; in Dutch, similarly, stress information influences strength of lexical activation [8]: the fragment *okTO-* facilitates recognition of the matching word *okTOber* but not of the mis-matching word *OCtopus*.

The present study addressed this same issue of the effect of mis-stressing on lexical activation, using Koster and Cutler's [5] semantic judgement task.

2. METHOD

2.1. Materials

Thirty-two sets of target words were selected, all monomorphemic bisyllabic Dutch nouns. The words of a set contained the same sounds in the first syllable plus at least the first sound of the second syllable, and they had an opposite stress pattern, e.g. *ATlas* ('atlas'), *atLEET* ('athlete'). All vowels in both stressed and unstressed syllables were full, so that mis-stressing (*atLAS*, *ATleet*) involved only a suprasegmental change. The target words of a set had no obvious semantic relation. For every target word a semantically related prime word was chosen; sometimes a synonym (noun), but mostly an associate. The primes for *atlas* and *atleet* were WERELD ('world') and SPORT ('sport'), respectively. A prime belonging to one target word of a set was never closely related to the other target of the set. The associative strength of the prime-target pairs was rated by 20 subjects in a pre-test; the mean rating across all pairs was 1.62 (on a scale from 1, "strongly related", to 5, "unrelated"), and no pair received a mean rating greater than the scale midpoint. The full list of experimental primes and targets may be found at www.mpi.nl/world/persons/private/ anne/materials.htm

In addition, 64 fillers were constructed; 32 semantically related prime-target pairs, half of them with a correct stress pattern, half of them with and incorrect stress pattern, and in the same way 32 unrelated prime-target pairs. Eight experimental lists were constructed, each containing 16 practice items, 18 warmup items (half at the beginning and half in the middle of the list), 32 experimental prime-target pairs, and 64 filler pairs, the latter two sets separately randomized for each list. There were eight possible permutations of any one experimental set (2 prime words x 2 target words x 2 stress conditions: see Table 1); these were counterbalanced across lists such that only one occurred in each list. All target words, with correct and incorrect stress patterns, were recorded on Digital Audio Tape by a female native speaker of Dutch. Word duration was measured, and segmental uniqueness point (UP) of each word determined.

	target	
	correct	incorrect
prime		
related		
WERELD	*ATlas*	*atLAS*
SPORT	*atLEET*	*ATleet*
non-related		
SPORT	*ATlas*	*atLAS*
WERELD	*atLEET*	*ATleet*

Table 1: Example prime-target combinations in the experiment.

2.2. Subjects and Procedure

88 subjects (19 male, 69 female; age range 18 to 33) of the Max Planck Institute's subject pool participated in the experiment in return for a small payment. They were tested individually in a sound-attenuated booth. They saw a prime word displayed (for 1300 ms) on a computer screen, and then heard a spoken word; their task was to decide whether the two words were similar in meaning, and to signal their decision as rapidly as possible by pressing one of two response keys labelled YES and NO. The subjects were informed that some words would be mispronounced. Subjects had a maximum of two seconds to respond after word offset; however, their instructions emphasised speed of response. The RTs, from a timing mark aligned with spoken-word onset, were collected by a computer running the experimental control program NESU. We also recorded the error rate (proportion of missed or erroneous responses). There was a rest break in the middle of the experiment (which was followed by the second set of warmup items).

3. RESULTS

Error rates and RTs were subjected to analyses of variance across subjects (F1) and items (F2), separately for the YES and NO responses. Separate analyses were conducted for RTs from word onset, word offset, and UP. One item (*modus*) received a hitrate below 50 %, and hence was excluded from all analyses, along with its pair (*model*). A 2.5 SD outlier correction was applied to the RTs, which resulted in the loss of 1-3% of responses in the different RT analyses. In the RTs from word offset, two more word sets (*circus/circuit* and *tapir/tapijt*) were excluded, on the basis that most responses were made before word offset, i.e. the corrected RTs were mostly below zero. Additionally, five subjects in the YES response analyses, and three subjects in the NO response analyses, were automatically excluded by the statistical programme, based on too many missing cases (i.e. they had responded very often before word offset).

There were clear effects of mis-stressing: error rates were significantly raised, and correct YES and NO responses were both significantly slowed ($p < 0.001$) when spoken words were mis-stressed, and this effect held whether RT was measured from spoken word onset, offset or UP. We will report here the effects from UP, as these most accurately reflect the course of word recognition. Figures 1 and 2 show mean YES and NO responses measured from UP, and Figures 3 and 4 show mean error rates for related pairs (i.e. missed YES responses) and unrelated pairs (i.e. false alarms instead of NO responses).

3.1. Related Pairs

The RTs from UP (see Figure 1) were longer for mis-stressed than correctly stressed words [$F1\ (1,75) = 46.31$, $p < 0.001$; $F2(1,56) = 30.16$, $p < 0.001$], and were also longer for words with (correctly) initial as opposed to final stress [$F1\ (1,75) = 63.12$, $p < 0.001$; $F2(1,56) = 19.87$, $p < 0.001$]. The interaction between the two factors was not significant.

The error rates (see Figure 3) likewise were higher for mis-stressed than correctly stressed words [$F1\ (1,80) = 11.65$, $p < 0.001$; $F2\ (1,60) = 13.42$, $p < 0.001$]. The difference between words with initial vs. final stress was significant in the subjects but not in the items analysis [$F1\ (1,80) = 10.10$, $p < 0.002$; F2
(
i

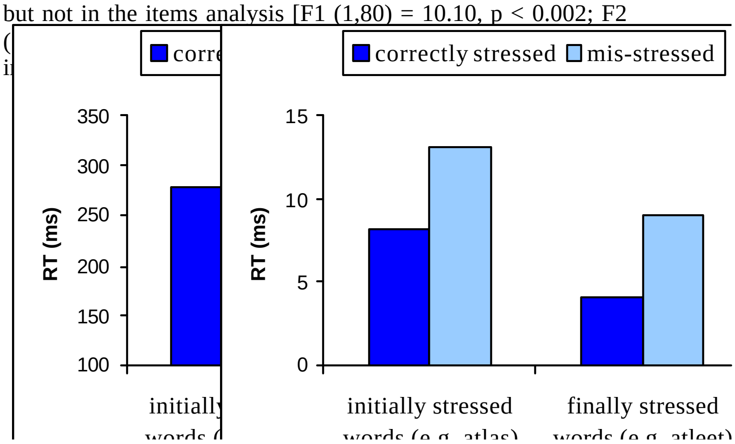


Figure 1: Mean correct YES responses to related pairs, measured from Uniqueness Point.

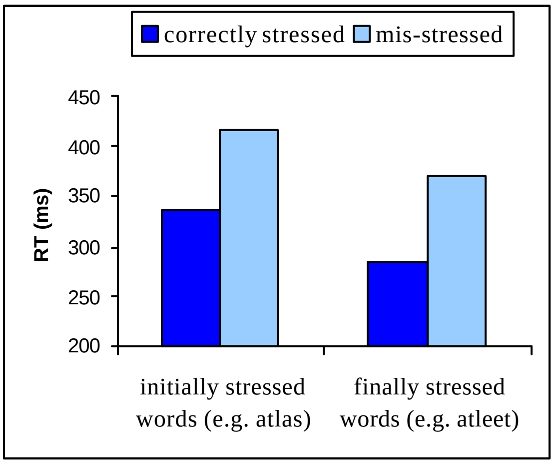


Figure 2: Mean correct NO responses to unrelated pairs, measured from Uniqueness Point.

3.2. Unrelated Pairs

The RTs from UP (see Figure 2) were again longer for mis-stressed than correctly stressed words [F1 (1,77) = 70.58, $p < 0.001$; F2 (1,56) = 50.20, $p < 0.001$]. RTs were again also longer for words with initial as opposed to final stress [F1 (1,77) = 27.06, $p < 0.001$; F2 (1,56) = 4.84, $p < 0.035$]. The interaction was not significant.

The false alarms (see Figure 4) showed a marginal advantage for correctly stressed over mis-stressed words [F1 (1,80) = 3.61, $p < 0.061$; F2 (1,60) = 3.78, $p < 0.056$]. Lexical stress pattern had no effect, and there was no interaction.

3.3. Word Frequency

The repeated finding of longer RTs for words with initial as opposed to final stress led us to investigate whether this could be an effect of word frequency. In fact, words with initial stress had lower (average) frequency in CELEX [9] (LogLemmaFreq=

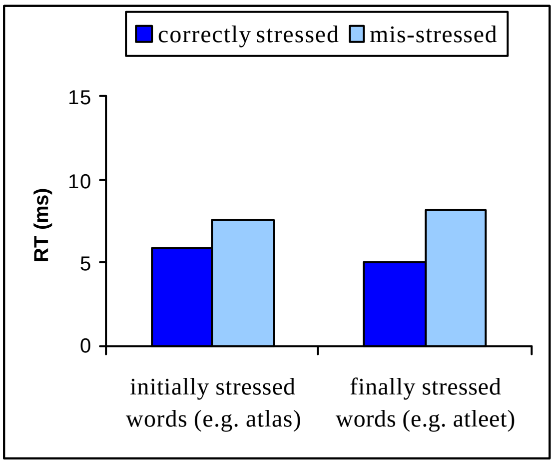
Figure 3: Mean error rate for related pairs.

Figure 4: Mean error rate for unrelated pairs.

.62) than words with final stress (LogLemmaFreq=.88). We performed an analysis of covariance on the YES-response item means, with log frequency as a covariate. Frequency could account for part of the variance [F (1,55) = 7.12, $p < 0.01$; Beta = -.29], but the effect of lexical stress pattern remained significant [F2 (1,55) = 18.34, $p < 0.001$].

4. DISCUSSION

The principal finding was that mis-stressing slowed RTs and raised error rates for judgements of semantic relatedness between a visually presented prime word and a subsequent spoken word. This is confirming evidence that, as we previously observed with this task [5], recognition of a spoken Dutch word



is adversely affected by mis-stressing even when vowel quality remains unchanged. Adverse effects of mis-stressing have also been reported in other word recognition tasks with Dutch materials [10, 11]. Although other studies [5, 10, 11] have reported different patterns of effects in bisyllables with initial stress (*atlas*) and final stress (*atleet*), here initially- and finally-stressed words produced closely similar patterns.

The mean effect on "YES" RTs (72 ms) and "NO" RTs (82 ms, for RTs from UP) was closely comparable. The effect of mis-stressing on error rate was also significant, and also equivalent for the two word types, but it was relatively small. Both misses (SPORT-*atleet*: "NO") and false alarms (SPORT-*atlas*: "YES") were increased by mis-stressing, but the highest error rate in any condition was 13% and the greatest inhibitory effect of mis-stressing only 5%. This result is not indicative of a fully determinative role for stress mismatch in lexical activation. In particular, note that the mean false alarm rate for mis-stressed words in unrelated pairs (e.g. SPORT - *atLAS*) was just 8%, which was only marginally greater than the 5.5% false alarm rate for the same condition with correct stress (e.g. SPORT - *ATlas*). Although subjects were instructed to respond as fast as possible, they did not tend to produce false alarms based on just the first syllable in this condition.

In other words, although stress information constrained lexical activation in this task, the constraint was not strong enough to completely rule out segmentally matching but suprasegmentally mismatching words. This pattern of results is compatible with models of lexical selection based on competition between words which receive greater and lesser activation, rather than on an all-or-none matching between input and lexical representation. Studies in Spanish [7], using a fragment priming task, showed that responses to words mismatched in stress by a spoken input were inhibited compared with a control condition, which again is evidence of competition between simultaneously activated words. In the Spanish experiments, too, error rates were low; also, the effects on competitor activation of a stress mismatch or a segmental mismatch were essentially identical.

Whether this is also the case in a competition environment in Dutch has not yet been established. A fragment-priming experiment in Dutch by Cutler and Donselaar [12] produced results suggesting that a segmental mismatch exercises greater constraint on activation than does a stress mismatch; although the mis-stressed fragment *ROton-*, for instance, produced significantly less facilitation of the visual target word ROTONDE than did the correctly stressed fragment *roTON-*, the segmentally mismatching fragment *woTON-* produced even less facilitation. However, in Cutler and Donselaar's experiment the input did not activate a minimally different competitor word (there are no Dutch words beginning *ROton-* or *woTON-*). Competitor words were manipulated in Donselaar and Cutler's *octopus/oktober* experiment [8], but in this case no comparison between stress and segmental mismatch was made.

Lexical stress languages differ widely in the extent to which stress and segmental structure covary. It is now becoming increasingly clear that closely similar word processing experiments may produce different patterns of results in different stress languages (see [13] for a review). Although, as described above, many findings suggest that in English it is the case that stress information plays a much weaker role in lexical activation than does segmental information, the role that stress plays in lexical activation seems to be substantially stronger in Dutch.

5. ACKNOWLEDGEMENTS

We are grateful to Marlies Wassenaar for technical assistance.

6. REFERENCES

1. Bond, Z.S. "Listening to elliptical speech: Pay attention to stressed vowels," *Journal of Phonetics 9*: 89-96, 1981.

2. Slowiacek, L.M. "Effects of lexical stress in auditory word recognition," *Language and Speech 33*: 47-68, 1990.

3. Cutler, A., and Clifton, C.E. "The use of prosodic information in word recognition," in H. Bouma, and D.G. Bouwhuis (Eds.), *Attention and Performance X*. Hillsdale, N.J.: Erlbaum, 1984.

4. Bond, Z.S., and Small, L.H. "Voicing, vowel and stress mispronunciations in continuous speech," *Perception & Psychophysics 34*: 470-474, 1983.

5. Koster, M., and Cutler, A. "Segmental and suprasegmental contributions to spoken-word recognition in Dutch," *Proceedings of EUROSPEECH 97*, Rhodes; pp. 2167-2170, 1997.

6. Frauenfelder, U.H., and Floccia, C. "The recognition of spoken words," in A. Friederici (Ed.), *Language Comprehension: A Biological Perspective*. Springer; Heidelberg, 1-40, 1999.

7. Soto, S., Sebastian-Galles, N., and Cutler, A. "Segmental and suprasegmental cues for lexical access in Spanish," *Journal of Memory and Language*, submitted for publication.

8. Donselaar, W. van, and Cutler, A. "Exploitation of stress information in spoken-word recognition in Dutch," *Journal of the Acoustical Society of America 102*: s3136, 1997.

9. Baayen, R.H., Piepenbrock, R., and Rijn, H. van, *The CELEX lexical database* [CD-ROM]. University of Pennsylvania, Linguistic Data Consortium Philadelphia, 1993.

10. Heuven, V.J. van "Perception of stress pattern and word recognition: Recognition of Dutch words with incorrect stress position," *Journal of the Acoustical Society of America 78*: s21, 1985.

11. Leyden, K. van, and Heuven, V.J. van "Lexical stress and spoken word recognition: Dutch vs. English," in C. Cremers, and M. den Dikken (Eds.), *Linguistics in the Netherlands 1996*. John Benjamins, Amsterdam, 1996.

12. Cutler, A., and Donselaar, W. van "*Voornaam* is not a homophone: Lexical prosody and lexical access in Dutch," *Language and Speech*, submitted for publication.

13. Cutler, A., Dahan, D., and Donselaar, W. van "Prosody in the comprehension of spoken language: A literature review," *Language and Speech 40*: 141-201, 1997.