Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition

Anne Cutler and Dennis Pasveer

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
anne.cutler@mpi.nl, dennis.pasveer@mpi.nl

Abstract

Experiments have revealed differences across languages in listeners' use of stress information in recognising spoken words. Previous comparisons of the vocabulary of Spanish and English had suggested that the explanation of this asymmetry might lie in the extent to which considering stress in spoken-word recognition allows rejection of unwanted competition from words embedded in other words. This hypothesis was tested on the vocabularies of Dutch and German, for which word recognition results resemble those from Spanish more than those from English. The vocabulary statistics likewise revealed that in each language, the reduction of embeddings resulting from taking stress into account is more similar to the reduction achieved in Spanish than in English.

1. Introduction: Stress in spoken-word recognition

Many languages distinguish words on the basis of stress. Thus English *FOREarm* refers to a part of the body, while *foreARM* is a verb meaning to prepare oneself in advance of trouble. There are no segmental differences between these two words; the stress placement is the sole phonetic distinction.

Stress languages are curiously reticent in exploiting this possibility for making distinctions between words. Minimal pairs differing only in stress are few in English (less than two dozen of them exist), and the same is true of other European stress languages such as Russian, German, Dutch or Spanish.

Nonetheless listeners can often use stress to distinguish words. Soto-Faraco, Sebastián-Gallés and Cutler [1] showed this in a study in Castilian Spanish. Listeners heard spoken sentences such as *He did not know how to write the word ...* ending with a word fragment which fully matched one of two potential words and differed from the other just in stress pattern. For instance, *prinCI-* (stressed on the second syllable) matches the first two syllables of the Spanish word *prinCIpio* ('beginning') and differs only in stress from *PRINcipe* ('prince'). Immediately after hearing the fragment, listeners saw a string of letters on a screen and had to decide whether this string was a real word. Their responses were significantly faster after fragments matching the visually presented word (e.g. to PRINCIPIO after *prinCI-* etc.) than after control fragments (e.g. *manti-*); responses after fragments which minimally mismatched and favoured another word (e.g., to PRINCIPIO after *PRINci-* etc.) were significantly slower than those after control fragments. Soto-Faraco et al. compared the use of stress information in this task to the use of vocalic and consonantal minimal mismatches. All three types of mismatch (vocalic, consonantal, stress) produced the same pattern of results. Thus the listeners were using stress information in the same way as phonemic information to ascertain which word they were hearing and to reject alternatives.

Rejection of alternatives is a particularly important component of spoken-word recognition. Current word-recognition models all assume that speech activates multiple candidate words which are either fully supported by the input, or temporarily receive partial support. These words compete with one another; the more one word is supported, the better it can suppress its rivals. The slower responses after mismatching fragments in Soto-Faraco et al.'s experiment are consistent with this competition-based explanation: the information supporting one of the alternatives (e.g., *principio*) simultaneously allowed it to suppress its rival (e.g., *principe*).

Not all stress-language users exploit stress information to the same extent as these Spanish listeners, however. Even stress languages which are closely related differ in the degree to which users exploit stress in word recognition. Dutch, German and English are closely related languages and lexical stress operates similarly in their phonologies [2]. But the use which listeners make of stress in these three languages differs. Results from experiments with Dutch and German listeners closely resemble the results from Spanish [3,4,5,6], but results from English suggest that listeners make significantly less use of stress information [7]. Especially the effects of mismatch in stress differ; Figure 1 shows the inhibitory effect of mismatch from the three very similar experiments conducted in Spanish [1], Dutch [4] and English [7] (mismatch data from German is not available, since the German studies [5,6] used a different experimental design). The effects in Spanish and Dutch are robust and statistically significant; the inhibitory effect in English is much smaller, and not statistically significant.

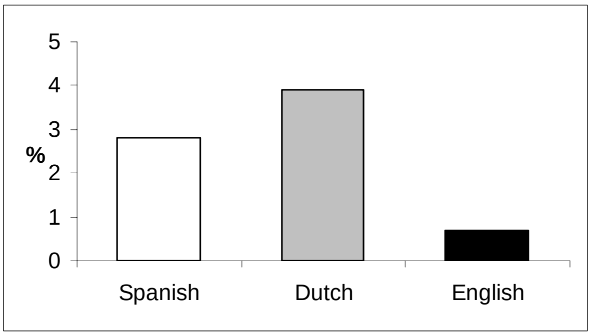

Figure 1: *Inhibitory effect of mismatching stress (RT given mismatching prime minus RT given control prime, expressed as % of control condition RT) in comparable cross-modal priming studies in Spanish, Dutch and English.*

Why should stress languages differ in the extent to which listeners exploit stress cues in word recognition, and in particular why should there be differences between such closely related languages as Dutch and English?

2. Vocabulary Structure

An answer was suggested in cross-linguistic analyses of vocabulary structure by Cutler, Norris and Sebastián-Gallés [8]. Exploiting a newly available lexical database of Spanish [9], they examined the characteristics of the Spanish lexicon, and compared it with English (using the CELEX database for that language [10]). Their interest concerned the implications of phonemic repertoire, and in particular the vowel:consonant ratio, for vocabulary makeup and hence for the competition process during word recognition in the two languages. Spanish has 25 phonemes and a skewed vowel:consonant ratio (four times as many consonants as vowels), whereas the English repertoire is larger (44 phonemes) but the vowel:consonant ratio is very nearly balanced. Indeed, the differences had far-reaching consequences: Spanish words were longer, and had more other words embedded within them, than English words.

An interesting effect of stress emerged, however, in the course of these analyses. Cutler et al. calculated the number of embedded words in each language (note that embeddings of short words within long ones are an inevitable consequence of vocabularies of hundreds of thousands of words constructed from repertoires of just a few dozen phonemes. All languages have many embedded words, but the competition process in spoken-word recognition efficiently solves this problem for the listener. Thus one would not expect languages to differ widely on this score). If embeddings are defined as any string of segments corresponding to a lexical form embedded within any other such string of segments – i.e., taking no account of stress – then Spanish has more than 2.2 times as many embedded words as English. This was a shocking finding, suggesting that the word recognition process could be plagued by much more competition in Spanish than in English.

However, the pattern changed considerably if stress was taken into account. (As an example: *sea* would be considered to be embedded in *secret, seniority, lessee* and *lassie* if stress is not considerd, but only in *secret* and *lessee* – in which [si] is the stressed syllable – if it is.) In English, the reduction due to conderation of stress was just over one-third: weighted by word frequency, matrix words of two to six syllables contain on average 0.94 embedded other words when stress is ignored, but only 0.59 when stress is taken into account. In Spanish, the reduction was at least twice as substantial. There were on average 2.32 words per carrier word if stress was not considered, but only 1.19 words per carrier word if content words were marked for stress, and only 0.73 words if all monosyllabic words, including function words, are considered to be stressed. In fact the true value is probably somewhere between the latter two – somewhere just under one embedded word on average. Either way, the gain delivered by stress information is strikingly greater for Spanish than for English.

Thus it appeared that paying attention to stress in word recognition produced a clear payoff in reduction of alternative word candidates for Spanish listeners. The payoff for English is very much smaller. This, according to Cutler et al. [8], could explain why Spanish listeners use stress in word recognition more effectively than English listeners do. If English listeners paid no attention to stress in word recognition at all, the amount of competition that they would be dealing with would be, on average, just under one spuriously activated word for every actually intended word. If Spanish listeners took full account of stress, the amount of competition for them would also average out at just under one spuriously activated word for every actually intended word. Thus the average embedding counts in the two languages are comparable when we assume the stress-adjusted figures for Spanish, but the unadjusted figures for English. The reduction of embedding in Spanish by considering stress thus removes the cross-language asymmetry in probability of embedding. We speculate that the resulting amount of competition – less than one spurious word per real word – may be effectively the acceptable level in spoken-word recognition.

If this explanation is correct, it also offers a potential solution to the puzzling asymmetry in the use of stress information in word recognition in the closely related languages Dutch, German and English. The solution would be that consideration of stress reduces competition in Dutch and German more than in English. From this, obvious predictions may be derived concerning vocabulary structure in the two languages; the Dutch (and German) embedding counts should more closely resemble the Spanish than the English counts. In particular, it should be the case that an average of less than one embedded word per carrier word should be achievable in these two languages only by taking stress into consideration in computing the embeddings. These predictions were tested in the present study. Note that currently available results from Dutch experiments strongly motivate these predictions; the results for German are as yet not fully conclusive, though those which exist so far suggest a pattern resembling that of Dutch. The analyses will thus speak further to the question of whether German indeed patterns like Dutch in this respect.

3. Lexical Analyses

The analyses we report were based on the CELEX database [10], which contains all words of Dutch and of German as well as of English. For the present study, the embedding calculations which Cutler et al. [8] had carried out for Spanish and English were repeated for Dutch and German. Note that the Dutch lexicon in CELEX is much larger than the English and German lexicons. The Dutch lexicon contains more than 124,000 lemmas, and frequency statistics based on a 42 million word corpus; the English lexicon contains 52,000 lemmas, and frequency statistics based on a 17.9 million word corpus; the German lexicon contains 51,000 lemmas, and frequency statistics based on a five million word corpus. These differences – especially in the corpora providing frequency statistics – are of course due in the most part to differences in the original sources from which the database was compiled. However, the differences in lexicon size are also in part due to reasons discussed further in section 5, below. Because of the size differences, absolute totals are obviously not directly comparable. Therefore only proportions are reported below.

For each corpus, the initial analysis involved checking each word in the corpus against all shorter words in the corpus to find possible embeddings. Homophonic forms (*peil*/*pijl*; *Bund*/*bunt*) were represented as single items, and entries with zero frequency were excluded. The size of the Dutch lexicon thereby became 67,500 items, of English 52,500, and of German 34,500. Syllable boundary matches were required for an embedding to be counted (thus *zee* 'sea' would be counted in *mu.se.um* 'museum' but not in *zeep* 'soap'; periods here give the syllable boundaries). Once these vocabulary totals had been ascertained, they were translated into an estimate of actual word availability (i.e., lexical competition) in natural listening by weighting the counts according to matrix word frequency, according to the frequency of occurrence statistics provided by CELEX.

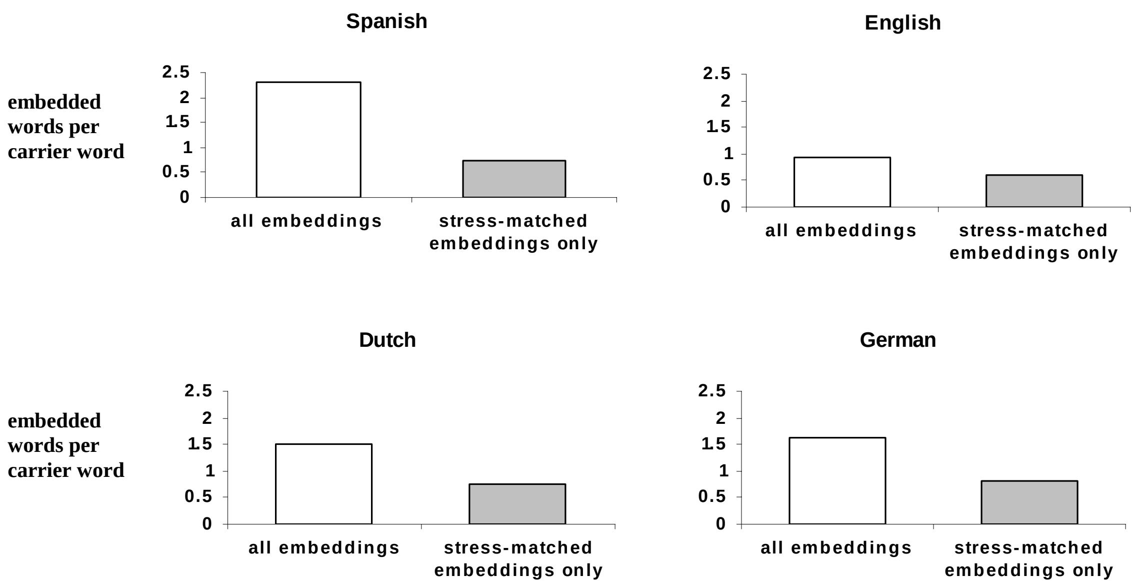

Figure 2: *Mean number of embedded words per word of spoken language in Spanish, English (data from [1]), Dutch and German, separately for all embeddings without regard to stress (left bars), and only embeddings in which stress levels of each syllable in embedded word and carrier word match (right bars).*

4. Results

4.1 Dutch

Weighted by word frequency, Dutch carrier words of two to six syllables contain on average 1.52 embedded other words when stress is ignored, but only 0.74 when stress is taken into account. Consideration of stress here thus results in a reduction of just over one half in the spurious activation count. With stress considered, the count thus falls below the level of one embedding per real word in the speech input. With stress not considered, it is above this figure.

4.2. German

Weighted by word frequency, German carrier words contain on average 1.62 embedded other words when stress is ignored, but only .80 when stress is taken into account. Again, carriers contain on average more than one word in the first case, but less than one in the second. The reduction achieved by taking stress into account is 51%.

Figure 2 shows the results for each language, along with the same calculation applied retrospectively to the results reported for Spanish and English by Cutler et al. [8]. It can easily be seen that English differs from all the other three languages in the low level of embeddings in the lefthand (all embeddings) column; that is, there is less benefit to be had from taking stress into account in English in the first place.

5. Discussion

From the results of the very similar word recognition experiments conducted in Dutch, Spanish and English, in which the Dutch listeners' results resembled the Spanish results more than they resembled the English results, we had predicted that we would find that the statistics of the Dutch vocabulary would resemble the Spanish more than the English statistics. This was a fairly remarkable prediction, because in all other respects – historical, structural – Dutch is of course far closer to English than to Spanish. Nevertheless, the predictions were warranted if we wish to make the case that stress information is used in word recognition if and only if its use pays off in terms of competition reduction.

The results of the vocabulary analyses indeed revealed that the reduction of embedding due to stress is greater in Dutch and in German than it is in English. In both languages, as in Spanish, the spurious activation count when stress is not taken into account is above one per real word, while taking stress into account reduces the count to below one.

Thus the lexical payoff in terms of competition reduction offers a simple explanation for why the experiments produced varying patterns. In Spanish, listeners use stress exactly as they use phonemic information in rejecting unwanted lexical competitors [1]. The results from comparable experiments in Dutch show that Dutch listeners are just as adept at using stress [3,4]. In fact, Dutch listeners can actually outperform English native speakers in making stress distinctions on isolated English syllables (e.g., does a fragment *mus-* come from *MUSic* or *muSEum*? [7]). Although the available results from German [5,6] are based on studies with a somewhat different experimental design, the present analyses suggest that were exact analogues of the Dutch experiments [3,4] to be conducted in German, the results would be as in Dutch. In German too, exploitation of stress information pays off in competition reduction.

Only in English is the payoff relatively small, and apparently too small to warrant full use of stress in lexical processing. We suggest that the magic number of one embedded word per carrier word is the clue; it is presumably impossible to get rid of embeddings entirely, but reducing them to a minority is achievable. Once they are in the minority, word recognition may not be seriously impeded. Without considering stress, embeddings are already in the minority for English. The further reduction to be achieved by taking stress into consideration is on this account simply not worth it for English. Note that it is not the case that English profits less from other options for embedding reduction. Thus rejection of any embedding which does not match the syllable boundaries of the carrier produces a huge reduction in the embeddings total, for English as for other languages [11].

It is also not the case that the differences simply reflect spurious differences in what is included in each lexicon. In Dutch and in German, compound words are written together, while in English they are written as separate words. Part of the greater size of the Dutch lexicon in CELEX is due to a very large number of compounded forms. However, the Dutch lexicon far outstrips the German lexicon in this respect too; but the embedding statistics for Dutch and German were, as we saw, very alike. Further, the constraints we applied for lexical inclusion resulted in lexicon sizes which were more closely comparable than the sizes in the full CELEX database.

Finally, the differences are not due to the main difference which concerned Cutler et al. [8], namely phoneme repertoire. In this respect, the three Germanic languages are extremely similar, and in no case is the vowel:consonant ratio skewed in the way that it is in Spanish.

Why, then, is the payoff of stress information so much poorer for English? Although stress placement patterns similarly in the three Germanic languages we have compared [2], the three differ in the phonological consequences of stress placement, in particular in the extent of vowel reduction. Consider the words *gala, cola, scala*, which exist (with mostly the same meaning) in all three of the Germanic languages compared here, and in all three cases have the same stress pattern (initial stress). The form *la* also exists as an independent syllable (sometimes even as a word with the same meaning) in all three languages. But *la* is simply not an embedding in the English words *gala, cola, scala*, because all three of them are pronounced with schwa in the second syllable. In Dutch *gala, cola, scala* and German *Gala, Cola, Skala*, however, the second syllable is not reduced. Thus it contains the same vowel as in *la*, making *la* an embedded word; but because the second syllable is in each case unstressed, the spurious embedding can be easily discarded once stress is taken into account in the activation of lexical candidates for recognition.

Vowel reduction, in other words, ensures that the vowels of any strong syllable mismatch the vowel in most unstressed word-internal syllables in English. The far more widespread use of vowel reduction in the phonology of English is responsible for the low level of embedding when stress is not taken into account. The lesser occurrence of vowel reduction in Dutch and German paves the way for more embeddings. Thus the phonological patterns in closely related languages may differ in such a way that the task confronting the listener may in consequence be radically affected (see [12] for further detail of this argument).

6. References

[1] Soto, S.; Sebastián-Gallés, N.; Cutler, A., 2001. Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language* 45, 412-432.

[2] Hulst, H. van der, 1999. *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter.

[3] Cutler, A.; Donselaar, W. van, 2001. *Voornaam* is not a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech* 44, 171-195.

[4] Donselaar, W. van; Koster, M.; Cutler, A., 2005. Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology* 58A, 251-273.

[5] Friedrich, C., 2002. *Prosody and spoken word recognition: Behavioral and ERP correlates.* Ph.D. dissertation. University of Leipzig.

[6] Friedrich, C.; Kotz, S. A.; Gunter, T.C., 2001. Event-related evidence of word fragment priming: A new correlate for language processing? *Psychophysiology* Suppl., 42.

[7] Cooper, N.; Cutler, A.; Wales, R., 2002. Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech* 45, 207-228.

[8] Cutler, A.; Norris, D.; Sebastián-Gallés, N., 2004. Phonemic repertoire and similarity within the vocabulary. In *Proceedings of the 8th International Conference on Spoken Language Processing,* S.H. Kin & M. Jin Bae (eds.). Seoul: Sunjin Printing Co., Vol. 1, 65-68. (CD-ROM).

[9] Sebastián-Gallés, N.; Martí, M. A.; Carreiras, M.; Cuetos, F., 2000. *LEXESP: Léxico informatizado del español*. Barcelona: Edicions Universitat de Barcelona.

[10] Baayen, H.; Piepenbrock, R.; van Rijn, H., 1993. *The CELEX Lexical database*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania (CD-ROM).

[11] Cutler, A.; McQueen, J.M.; Jansonius, M.; Bayerl, S., 2002. The lexical statistics of competitor activation in spoken-word recognition. *Proceedings of the 9th Australian International Conference on Speech Science and Technology*. Melbourne, 40-45.

[12] Cutler, A., 2005. Lexical stress. In *The Handbook of Speech Perception,* D.B. Pisoni & R.E. Remez (eds.). Oxford: Blackwell, 264-289.