Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Word accents and phonological neighbourhood as predictive cues in spoken language comprehension

Pelle Söderström [1], Merle Horne[1], Mikael Roll[1]

[1] Lund University, Sweden

pelle.soderstrom@ling.lu.se, merle.horne@ling.lu.se, mikael.roll@ling.lu.se

Abstract

The present contribution summarises event-related potential (ERP) and functional magnetic resonance imaging (fMRI) findings related to the processing of Swedish word accents in speech comprehension. It has previously been seen that word accents – either a low tone (accent 1) or a high tone (accent 2) on a word stem – can be used to pre-activate suffixes. Furthermore, it has been found that accent 1 seems to be a stronger "predictor" of upcoming suffixes as compared to accent 2. It has been proposed that accent 1 stems give rise to a pre-activation negativity (PRAN) brain potential, which is related to their high predictive weight as regards associated suffixes. We suggest that the processing differences between accent 1 and 2 stems can partly be explained by the difference in the number of word activations elicited by accent 1 and accent 2 word stems. This idea is tested by means of a regression analysis, which found that word stems in denser phonological neighbourhoods – i.e. which occur in more lexical items – elicit smaller PRAN effects. The results point to the importance of word accents in Swedish word comprehension.

Index Terms: word accents, Swedish, pre-activation, morphology, grammar, ERP, PRAN, fMRI, phonological neighbourhood

1. Introduction

Previous studies have found that Swedish word accents – 'accent 1' and 'accent 2' – are used as prosodic cues by listeners to rapidly pre-activate upcoming suffixes ([1], [2], [3], [4], [5], [6]). A possible reason for this rapid pre-activation can be tied to the parameter of frequency, which is known to constitute important information in language and speech processing ([7], [8]). Another important factor is competition between similar-sounding lexical items. The aim of the present article is to review findings from behavioural, event-related potential (ERP) and functional magnetic resonance (fMRI) data and to propose that word frequency and lexical competition play a role in triggering an early ERP[1] brain response that has been observed to occur when processing word stems in speech comprehension. Since the response differs depending on the word accent occurring on the stem, it can be thought to be related to lexical access processes. Thus, we performed a regression analysis to investigate whether differences in the ERP response for accent 1 and accent 2 stems can be explained – at least in part – by the number of words activated by a particular word stem.

1.1. Swedish word accents

Word accent is a prosodic phenomenon which exists in most varieties of Swedish and Norwegian. In Central Swedish – the variety spoken in the areas around the capital Stockholm – accent 1 (in unfocused position) can be analysed as a low tone at the beginning of the stressed vowel of a word (HL*). Accent 2, on the other hand, can be analysed as a high tone (H*L) followed by a fall through the stressed vowel. In focused position, a H-focal tone is added after the word accent ([9]). Figure 1 can be used as an illustration of the word accent difference in unfocused position. The phonetic realisation of the word accents varies in different dialects of Swedish. For example, in South Swedish, the situation is virtually the opposite: the stressed syllable begins with a higher tone for accent 1 as compared to accent 2. Unless otherwise stated, the present article will focus on Central Swedish word accents.

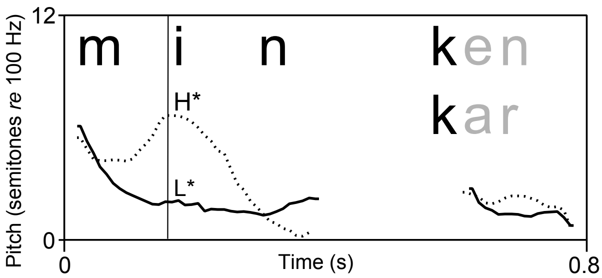

Figure 1: *F0 curves for the accent 1 (HL*) word* minken *'the mink' (solid line) and the accent 2 (H*L) word* minkar *'minks' (dotted line). The vertical line represents vowel onset.*

1.2. Relationship between prosody, morphology and prediction

One of the most salient features of word accents – making them particularly interesting for the purposes of language processing research – is their association with morphology. The word accent attached to a stem is highly conditioned by morphology ([9], [10], [11]). For example, if the singular noun suffix *-en* is attached to the word stem *bil* ('car'), the resulting word is *bil*$_1$*-en* ('the car') where the subscript indicates which word accent is attached. However, if the plural suffix *-ar* is attached to the stem, the word will be associated with accent 2 (*bil*$_2$*-ar*, 'cars'). The same is true for loan words: the word *chatt* 'chat room' takes accent 1 together with the singular suffix (*chatt*$_1$*-en*) and accent 2 with the plural suffix (*chatt*$_2$*-ar*). Furthermore, in Central Swedish, accent 2 is associated

[1] Commonly used in neurolinguistic studies, ERP is a technique which measures electrical activity in the brain (please refer to [6], section 2.4 for a detailed description).

 doi: 10.21437/SpeechProsody.2016-10

with compound words. Compounding is highly productive in Swedish, allowing the formation of novel compounds such as *trädkrig*$_2$ ('tree war') or *forskningsministerskola*$_2$ ('research minister school'). Even if all constituents making up a compound would individually be accent 1 words, the compound will still have accent 2. For example, although *neder*$_1$ 'downwards' and *länderna*$_1$ 'the lands' are accent 1 words, the compound *Nederländerna*$_2$ ('the Netherlands', example taken from [12]) has accent 2. What is interesting from the point of view of predictive models of language processing is that initial word stems with accent 2 (as opposed to stems with accent 1) could then potentially cue related accent 2 suffixes (e.g. plural, past tense) as well as a potentially infinite number of compounds. As an illustration of this, the Swedish Academy Lexicon, *SAOL13* (http://spraakdata.gu.se/saolhist/) reveals that the stem *ned*$_2$- is the initial syllable of 186 compound noun entries.

1.3. Phonological neighbourhoods of lexical items

Many models of spoken word recognition assume that words are activated to different degrees and compete with each other during speech processing (e.g. [13], [14]). The activation of different numbers of candidates clustered into "neighbourhoods" of similar-sounding words can be thought to be reflected in different ERP effects. In experimental contexts, the most commonly used definition of a member of a phonological neighbourhood is "any [phonological] transcription that could be converted to the transcription of the stimulus word or non-word by a one phoneme substitution, deletion, or addition in any position" ([15], p. 326). A phonological neighbourhood with few competing words is known as a "sparse neighbourhood", while one with many candidates is known as a "dense neighbourhood". However, it is not immediately clear how Swedish word accents should be accommodated in such a definition of neighbourhood since their distinctive (phonemic) function is minimal ([10]). They are mainly morphologically conditioned, i.e. associated with different affixes, but are associated with stressed vowels and as the results of [5] indicate, are connected with segments very early in word processing.

An example of a phonological neighbourhood in Swedish is constituted by the form *bil* /bi:l/ 'car' and the 11 neighbouring forms arrived at by substitution, addition or deletion of one phoneme in /bi:l/: *bi* 'bee', *il* 'gust', *pil,* 'arrow' *fil* 'file', *sil* 'sieve', *mil* '10 kilometres', *kil* 'wedge', *bit* 'bit', *bila* 'go by car', *bal* 'ball' and *bål* 'torso' giving a total neighbourhood size of 12. This neighbourhood includes only one accent 2 word (*bila* 'to go by car') which is also the only multi-syllabic word. Were prosodic information included, the simple application of the definition of phonological neighbourhood becomes more complicated since a given stem (e.g. *bil*) can potentially be associated with both accents depending on what follows the stem. If we assume that the typical accent 1 stem (which cues a small number of continuations) resides in sparse neighbourhoods while accent 2 stems are part of denser neighbourhoods, it could be argued that the effect of phonological neighbourhood density on ERP amplitudes should tell us something about the way in which lexical candidates are differentially activated by stem phonological information. Thus, the hypothesis is that accent 1 stems elicit larger negativities due to their residing in sparse phonological neighbourhoods with less lexical competition, while the opposite is hypothesised for accent 2, i.e. smaller negativities due to relatively denser phonological neighbourhoods for accent 2 lexical items.

The size of the initial stretch of speech used to activate possible candidates has been proposed to correspond roughly to two to three segments/phonemes ([16], [17], [18]), often "an initial consonantal phoneme or phoneme cluster and part of the following vowel" ([16], p. 52). Also, it has been proposed that stressed syllables are more likely to initiate a lexical search ([19]).

2. Behavioural and neurophysiological correlates of word accent processing

Results from behavioural experiments (response/reaction times) have shown that accent 2 words take longer to process as compared to accent 1 words ([2]). Furthermore, invalidly cued suffixes (suffixes cued by the wrong word accent) also tend to take longer to process, suggesting some type of revision/re-analysis mechanism. In ERP data, accent 1 word stems have consistently been found to give rise to a negatively charged effect which occurs ca. 200-300 ms following F0-onset[2]. An illustration of the stem negativity – which we will refer to as the pre-activation negativity (PRAN) – can be seen in Figure 2.

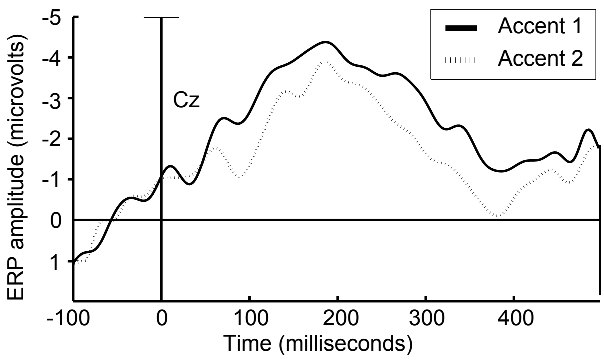

Figure 2: *ERPs measured from F0 onset, averaged over items (data from Roll et al. ([5])). Accent 1 stems (solid line) elicit more negative ERP amplitudes as compared to accent 2 stems (dotted line).*

2.1. ERP negativity effects – possible sources

Earlier studies have shown that the differential effects found for accent 1 and 2 words are not simply due to e.g. acoustic factors related to the two word accent patterns, which are obviously phonetically different (L* vs H* at the beginning of the stressed syllable). Roll et al. ([3]) included a condition in which test words were de-lexicalised (using the Hum function in Praat ([20]), leaving the tonal contour intact. If the PRAN had reflected differences in the phonetic realisation of accent 1 and 2, it should have been present for the de-lexicalised words. However, accent 1 did not produce greater negativity in these stimuli. The effect was rather an increased N1 component for

[2] This has been previously analysed as a positivity induced by accent 2 suffix activation ([1]). However, accent 1 stems have since been found to elicit increased neural activity (BOLD and gRMS ([5])), suggesting that the effect is rather a negativity driven by accent 1 stems during lexical access.

accent 2, reflecting the increased auditory salience of the high tone. The following negativity is thus more likely related in some way to lexical access processes when accents are already associated with stem vowels.

As already mentioned, in South Swedish, accent 1 can be analysed as a high tone while accent 2 is a low tone at the beginning of the stressed vowel. Thus, it is virtually the acoustic phonetic mirror image of Central Swedish. Roll ([6]) took advantage of this difference in an ERP study which found that accent 1 stems in South Swedish elicited a negatively charged effect just as in Central Swedish. Taken together, these findings can be used to rule out purely acoustic explanations for the ERP negativity differences found in response to accent 1 and 2 words.

In order to test whether the connection between word accents and suffixes remains even when the stem has no meaning (i.e. pseudo-stems), Söderström et al. ([4]) created stimuli composed of pseudo-stems (nouns) plus existing singular or plural suffixes (forming pairs such as *tväk*$_1$-*en*/*tväk*$_2$-*ar*, 'the tväk/tväks'). A PRAN was still found for accent 1 stems, while invalidly cued suffixes elicited a P600 effect, suggesting that the connection between word accents and morphology is not dependent on the lexical meaning or status of the word stem.

2.2. ERP effects and fMRI activations for word stems

Increased negativities for accent 1 stems have been found in all ERP studies on Swedish word accents, including Roll's ([6]) study on South Swedish and the Söderström et al. ([4]) study on pseudo-stems. As noted above, this shows that the negativity cannot be explained by acoustic differences between the word accents, or by the lexical status or meaning of the stem. However, it has been suggested that the negativity could reflect in some way the greater predictive status of accent 1 stems: accent 1 cues a smaller number of possible continuations, and it should be easier to predict how a word is going to end if there are fewer possibilities to choose from. Accent 2 stems cue a larger number of suffixes as well as potentially many compound words, and thus prediction of continuations based on these stems will be more difficult.

In order to investigate the PRAN more closely, a number of suffixes in the Söderström et al. ([4]) study were replaced with coughs, while the task was as previously to judge whether the word was singular or plural. Response accuracy was relatively high (88% for accent 1 and 72% for accent 2 stems) for both word accent patterns, suggesting that the word accent can be used to access the meaning of an omitted suffix. However, perhaps the most interesting finding was that there was a correlation between accent 1 stem ERP amplitude and response accuracy: a larger PRAN was associated with greater response accuracy. This was taken to suggest that the negativity is indeed an effect reflecting pre-activation of the upcoming suffix.

Findings in the first combined ERP/fMRI investigation of Swedish word accents ([5]) suggested that the stem tone (the word accent) is analysed in left auditory cortex (A1) and activates a phonological representation in superior temporal gyrus (STG) (a process similar to e.g. Chinese or Thai word tones). The tone then immediately activates a suffix representation in the left inferior frontal gyrus (LIFG) – an area associated with e.g. morphological processing – as early as 136 ms after F0 onset, and morphological processing peaks 120 ms later in the LIFG. The study thus supports the assumption of a strong tone-suffix connection in the brain.

3. Word accents and neighbourhoods of activated words

All neurolinguistic studies on Swedish word accents reported on here have pointed to a mechanism whereby the stem tones are used to pre-activate upcoming suffixes. In addition – while both word accents can be said to cue possible continuations – accent 1 seems to be a more robust predictor than accent 2, as evidenced by the accuracy data from Söderström et al. ([4]) and the fact that accent 1 stems seem to activate their suffixes more strongly ([1], [2], [5]). One suggested explanation for these results is that accent 1 stems are always part of sparser neighbourhoods of competing words, while accent 2 stems cue words which are part of denser neighbourhoods. In support of the suggestion that accent 2 stems are part of denser neighbourhoods, Hunter ([21]) found that increased neighbourhood density leads to decreased ERP negativities, just as has been found for accent 2 stems.

3.1. Testing the effect of phonological neighbourhood density on ERP amplitudes

Tonal information is not customarily included in phonological neighbourhood analyses. However, in a language like Swedish, where accents have a certain degree of distinctive function, it can be assumed that they constitute part of the phonological neighbourhood. Thus, since we hypothesise that the PRAN found for accent 1 stems is related to the fact that a smaller number of words are activated by these stems, we expect ERP amplitudes to be more negative with decreasing phonological neighbourhood size.

A regression analysis was carried out on stimulus words from a previous experiment ([5]) to see whether differences observed in the amplitude of the stem negativity could in part be attributed to lexical activation or competition. There were 120 words in total (60 accent 1, 60 accent 2 such as *fisk*$_1$-*en*/*fisk*$_2$-*ar*, 'the fish/fishes'). Neighbourhood size was calculated using *SAOL13* for each word stem, mean = 7.4, *SD* = 3.7. A neighbour was defined as an existing Swedish word that was obtained by adding, replacing or deleting one phoneme from the original word. Given the definition where neighbourhood is calculated on the basis of phonemes, there were no segmental differences between accent 1 and 2 stems (i.e. *fisk*$_1$ and *fisk*$_2$ differ only in word accent). One test item was excluded from the analysis due to outliers found in the ERP data. The time window from Roll et al. ([5]) – 136-280 ms from word F0 onset – was split into an "early" and "late" window, and the analysis focused on the late window (208-280) due to it being closer to the time window in which Hunter ([21]) found effects of neighbourhood density (240-300 ms measured from word onset). Also, the analysis focused on a region of interest (ROI) where the ERP effect had been greatest in Roll et al. ([5]): electrodes Cz, C3, Fz and F3. Thus, ERP amplitude in the 208-280 ms time window and in the chosen ROI was used as dependent variable, while neighbourhood density was included as a predictor.

3.2. Results

A significant regression equation was found ($F(1,114)$ = 7.659, p = 0.007), with an R^2 of 0.063. The equation for the predicted ERP amplitude is given in Example 1.

ERP amplitude = -3.426 + 0.098(density) (1)

Thus, the negative ERP effect decreased 0.098 µV for each unit increase of neighbourhood density, meaning that word stems with denser phonological neighbourhoods gave rise to less negative ERP amplitudes (see Figure 3).

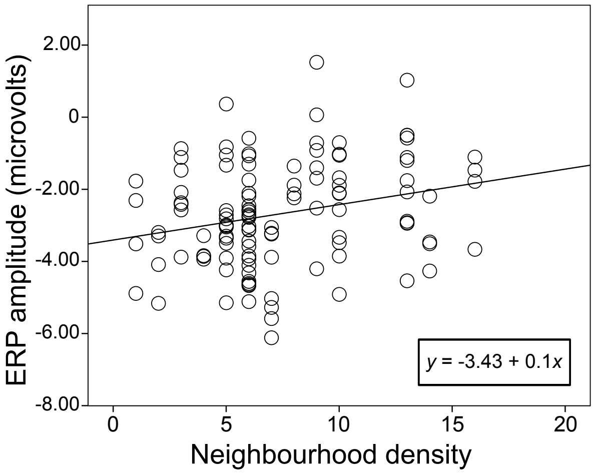

Figure 3: *Relationship between phonological neighbourhood density and ERP negativity amplitude for test items in Roll et al. ([5]). The regression line shows that the amplitude became more positive as neighbourhood density increased. Thus, the ERPs for test items were more negative when the stem was part of a sparse neighbourhood.*

3.3. Discussion

The results indicate that increased phonological neighbourhood density leads to decreased stem ERP negativities. In other words, the fewer neighbours a word stem has, the more negative its ERP amplitude will be in the time window 208-280 ms. This is in line with Dufour et al. ([18]) and Hunter ([21]), who also found more positive ERPs for words in dense neighbourhoods in similar time windows. Hunter suggests that this reflects increased lexical activation and competition demands for words with many neighbours and Dufour et al. similarly propose that word processing is facilitated when a lexical item is part of a sparser neighbourhood.

4. Conclusion

Brain imaging studies on speech comprehension in Swedish have led to results pointing to an ERP effect (PRAN) ca. 200-300 ms after F0-onset on word stems. The effect occurs after word accent activation and has been believed to reflect lexical access processes in which highly predictable material can be activated ahead of time. Furthermore, correlations of PRAN with phonological neighbourhood density measurements indicate that the PRAN shows sensitivity to lexical competition effects. Thus the results provide support for viewing PRAN as an index of lexical access in speech comprehension (see e.g. [22]), and it is currently the object of a follow-up study.

5. Acknowledgements

The work was supported by the Swedish Research Council (grant number 2011–2284), Knut and Alice Wallenberg Foundation (grant number 2014.0139), and Marcus and Amalia Wallenberg Foundation (grant number 2014.0039).

6. References

[1] M. Roll, M. Horne and M. Lindgren, “Word accents and morphology—ERPs of Swedish word processing,” *Brain Research*, vol. 1330, pp. 114–123, 2010.

[2] P. Söderström, M. Roll and M. Horne, “Processing morphologically conditioned word accents,” *Mental Lexicon*, vol. 7, pp. 77–89, 2012.

[3] M. Roll, P. Söderström and M. Horne, “Word-stem tones cue suffixes in the brain.” *Brain Research*, vol. 1520, pp. 116–120, 2013.

[4] P. Söderström, M. Horne and M. Roll, “Stem tones pre-activate suffixes in the brain,” (submitted).

[5] M. Roll, P. Söderström, P. Mannfolk, Y. Shtyrov, M. Johansson, D. van Westen and M. Horne, “Word tones cueing morphosyntactic structure: neuroanatomical substrates and activation time course assessed by EEG and fMRI,” *Brain & Language*, vol. 150, pp. 14–21, 2015.

[6] M. Roll, “A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs,” *Nordic Journal of Linguistics*, vol. 38, no. 2, pp. 149–162, 2015.

[7] J.L. Bybee, “From Usage to Grammar: The Mind's Response to Repetition,” *Language*, vol. 82, no. 4, pp. 711–733, 2006.

[8] J.B. Pierrehumbert, “The Dynamic Lexicon,” in A. Cohn, M. Huffman and C. Fougeron (Eds.) *Handbook of Laboratory Phonology*. Oxford: Oxford University Press, pp. 173-183, 2012.

[9] G. Bruce, *Swedish word accents in sentence perspective*. Lund: Gleerup, 1977.

[10] T. Riad, *The Phonology of Swedish*. Oxford: Oxford University Press, 2014.

[11] J. Rischel, “Morphemic Tone and Word Tone in Eastern Norwegian,” *Phonetica*, vol. 10, no. 3–4, pp. 154–164, 1963.

[12] T. Riad, “Scandinavian accent typology,” *Sprachtypologie und Universalienforschung (STUF)*, vol. 59, no. 1, pp. 36–55, 2006.

[13] P.A. Luce, *Neighborhoods of Words in the Mental Lexicon (Research on Speech Perception, Technical Report No.6)*. Bloomington: Speech Research Laboratory, Department of Psychology, Indiana University. 1986.

[14] W.D. Marslen-Wilson, “Functional parallelism in spoken word-recognition,” *Cognition*, vol. 25, pp. 71–102, 1987.

[15] M.S. Vitevitch and P.A. Luce, “When Words Compete: Levels of Processing in Perception of Spoken Words,” *Psychological Science*, vol. 9, no. 4, pp. 325–329, 1998.

[16] W.D. Marslen-Wilson and A. Welsh, “Processing Interactions and Lexical Access during Word Recognition in Continuous Speech,” *Cognitive Psychology*, vol. 10, pp. 29–63, 1978.

[17] D. Norris, “Shortlist: a connectionist model of continuous speech recognition,” *Cognition*, Vol. 52, pp. 189–234, 1994.

[18] S. Dufour, A. Brunellière and U.H. Frauenfelder, “Tracking the Time Course of Word-Frequency Effects in Auditory Word Recognition With Event-Related Potentials,” *Cognitive Science*, vol. 34, pp. 489–507, 2013.

[19] F. Grosjean and J.P. Gee, “Prosodic structure and spoken word recognition,” *Cognition*, vol. 25, pp. 135–155, 1987.

[20] P. Boersma and D. Weenink, “Praat: doing phonetics by computer [Computer program],” 2012.

[21] C.R. Hunter, “Early effects of neighborhood density and phonotactic probability of spoken words on event-related potentials,” *Brain & Language*, vol. 127, pp. 463–474, 2013.

[22] F. Pulvermüller, Y. Shtyrov and O. Hauk, “Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain,” *Brain & Language*, vol. 110, pp.81-94, 2009.