

Saliency, Focus and Second Language Acquisition

S.E. Carroll & Christine Shea

Department of Linguistics
University of Calgary
ceshea@ucalgary.ca

Abstract

In the literature on second language acquisition, saliency is assumed to automatically render input available for processing. The precise nature of "saliency" is rarely clarified; it is simply assumed that prosodic constructs such as focal accent or stress are universally perceptually salient. We examined this relationship between intrinsic saliency and second language learning by using focal accent to highlight familiar and unfamiliar lexical items in sentences presented to adult second language learners. If learners perceive unfamiliar focally accented words as perceptually salient, they will recognize them with greater accuracy than items without focal accent.

1. Introduction

The research reported on here is part of a larger program examining the complex relationships between input, language processing and second language acquisition. It investigates the commonplace assumption in second language acquisition research that parts of the signal are inherently more prominent or salient than others and that a learner's attention during speech processing will necessarily be drawn to those prominent parts of the input. In other words, the learner will be better able to process (i.e., analyze faster and with more accuracy) the parts of the signal that are marked in ways which result in perceptual saliency, to the detriment of those which are not so marked.

Upon close examination, however, many of the common claims regarding input do not stand up to scrutiny. Specifically, perceptual saliency is not an objective, acoustic property of the speech signal. Instead, saliency is a perceptual effect resulting from speech perception and language processing. There are many phenomena which can lead to perceptions of prosodic prominence, most of which can only be explained in terms of abstract mental representations, i.e., phenomena which are not in the signal. Following this argument to its logical conclusion, the input to speech processing cannot be the same as the input to language learning because prominent elements

extracted from the speech stream during speech processing must be first somehow translated into the abstract mental representations that give them meaning. In other words, language acquisition *precedes* perceptual effects unless it can be demonstrated that the perceptual effects are universal and a priori. Given that the cues to constructs like focal accent are language specific, we think that the latter is unlikely. The question that emerges, therefore, is what initial sensitivity listeners have to prosodic prominence as an acoustic phenomenon, extracted from the speech stream, and what role does that initial sensitivity play in segmenting sound units during language acquisition. In other words, how do learners go from the acoustic cues that exist in the speech stream, which may render a particular unit prominent, to linking this prominent unit with higher prosodic units that can subsequently be used for segmentation of the incoming signal?

We propose that one way of approaching this question is to examine if second language learners are sensitive to focal accent in the segmentation of novel words in their target language.

2. Focus and Prosodic Prominence

Listeners naturally search for the most informative and relevant part of an utterance. Second language learners are no different. However, they are different from native speakers in the sense that they have lower vocabulary levels, less efficient lexical recognition processes and possibly different segmentation strategies (Akker & Cutler 2003; Cutler et al. 1983, 1992). Native language speech perception research shows that this search for information in utterances leads listeners to process accented syllables more rapidly than unaccented syllables when accented syllables correspond to focus and unaccented syllables correspond to old information (Cutler 1988; Terken & Nootboom 1987). Accented syllables are often acoustically clearer and easier to process than unaccented syllables.

The present study examines how L2 learners go from prominence at the sentence level, or the auditory salience of a phonetic or a linguistic unit in an utterance, to focus, a linguistic notion that encodes information structure. In addition, subjects are more likely to remember words if they were in focus position. Akker and Cutler (2003) call this first effect the “predicted-accent effect” and maintain that it is accompanied by an effect for semantic focus.

It must be emphasized, however, that the relationship between focal accent and stretches of the signal are mediated by phonetic and phonological learning. Neither focal accent nor prominence are physical properties of the signal. Like stress, they do not constitute

objective cues to word boundaries. Rather, prosodic prominence is perceived when a coalition of phonetic properties are detected in the signal and associated with focal accent as a consequence of the language user having acquired the markers of information structure of a given language. It must also be pointed out that prosodic prominence is not a universal marker of focus. And in those languages in which it does occur, phonetic exponents vary (Eady et al. 1986; Fowler & Housum 1987). Thus, Wells (1986) investigating focal accent in British English found that listeners react to a complex coalition of phonetic factors, including maximum pitch range, kinetic tone, loudness peaks, crescendo and decrescendo. There may not be any simple relation between phonetic properties of the signal and the perception of prosodic prominence. It is highly unlikely, following this reasoning, that adults will universally be initially sensitive to focal accent when exposed to an L2 for the first time. Beginner learners of a new language may need to acquire the language specific coalitions of cues to focal accent *before* they will be able to locate word boundaries in the signal. The ability to direct one's attention to prosodically prominent syllables may be something which we can do as L2 learners as a *consequence of learning* rather than it being a prerequisite for it. In short, segmentation ability may only emerge subsequent to an initial phase of phonetic acquisition and mappings between phonetic patterns and phonological units. There are good reasons, therefore, to investigate the role of prosodic prominence among L2 learners.

3. Hypotheses

The main hypotheses tested in this study were:

- (1) Focal accent will attract a listener's attention to some part of the signal, making it salient.
- (2) The salience created by focal accent will facilitate segmentation of syllables and prosodic words from the speech stream.

The cumulative effect of these hypotheses is that novel words which are focally accented will be acquired sooner than words which are not. However, attention alone will not adequately account for how segmentation occurs, since there must be a description of how continuous acoustic properties align with prosodic units.

4. Experiment

The objective of this experiment was to see whether prosodic prominence affects extraction (segmentation) and recall of lexical items in a second language.

4.1. Subjects

Participants were 26 adults, all University of Calgary students. Results of two participants were discarded due to experimenter error and a further two were discarded when their German results revealed only 1 correct response, leaving 22 subjects in total, 10 males and 12 females, all native speakers of Canadian English.

Results of a questionnaire eliciting a description of subjects revealed that five participants were exposed to German before the age of six and 17 had late exposure, i.e., after the age of 17.

4.1.1. Proficiency in German

In the questionnaire, subjects were asked to estimate their proficiency in German. Median self-ranking, on a scale of 1 (highly proficient) -5 (low proficiency) was 2.75. Participants took the on-line Goethe *Einstufungs* or placement test which is a grammar test. Results ranged from 10%-97%, the mean was 41%. This corresponds to a low intermediate proficiency level (B1 on the Common European Frame of Reference scale). We also had subjects carry out the *Zum Zertifikat Deutsch* (Klett) listening test. Results ranged from 13% to 100%, mean 55%. The grammar test scores correlated positively with the listening test scores at $r = 0.783$.

4.2. Stimuli

We wanted to make use of an auditory judgment task using unknown words. We decided to use bird names as our target lexical items because they provided an open-ended class of which our learners would have little knowledge. In addition, by using real words, we avoided problems inherent to the creation of nonce words, and also assured that our participants would be mapping to nominal lexical categories with a familiar semantic meaning (even if they did not know the specific bird, all adults have a notion at least of what a prototypical bird is).

In the creation of our stimuli, we manipulated two factors: whether or not the bird names were culturally common or uncommon and whether or not the bird names received prosodic prominence. These factors were balanced across blocks. The common/uncommon distinction was made using native speaker impressionistic judgments. 32 items were selected, 16 targets and 16 distracter items, with an equal combination of common and uncommon names in each language. Carrier sentences were then constructed for each word, in each language, containing information about bird habitat, feeding, migratory habits and ecology. There were 288 sentences in total. The

stimuli were recorded by a female speaker of standard Canadian English.

In order to create our stimuli in the prosodically prominent (PP) and non-prosodically prominent (NPP) contexts, the bird names were elicited through question and answer pairs, requiring narrow focus answers:

- (3) Prosodically Prominent Question:
Which birds are moving further and further into the Arctic?
The fat *ACCENTORS*
- (4) Non-Prosodically Prominent Question:
Which accentors are moving further and further into the Arctic?
The FAT *accentors*

In both (3) and (4), the target words were uttered as the response to a narrow focus question with focus on the bird name. In the non-prosodically prominent case, the target words were uttered as a response to a narrow focus question with focus on the immediately preceding adjective. This had the consequence that the bird name was deaccented. The bird names were in each case spliced out and inserted into carrier sentences. Across the full set of stimuli, bird names occurred sentence initially, medially or finally, but all target names were sentence initial. German stimuli were recorded by a female speaker of standard German who recorded sentences, bird names and the question-answer pairs in a sound-attenuated booth in a laboratory of the Language Research Centre of the University of Calgary. Recordings were digitised, adjusted using Audition/CoolEdit 2, and saved as wav files.

The experiment was created using the DMDX program. 8 target sentences were prosodically prominent while 8 were non-prosodically prominent. Each target sentence appeared with 5 filler sentences to comprise a target block. Order of sentences in each block was initially random with the proviso that no target sentence could appear in the 5th or 6th position. Order of sentences in each distracter block was fully randomized. Once each target block and distracter block was set up, its order of presentation of sentences was fixed. Order of blocks was randomized during the experiment.

4.2.1. Acoustic characteristics of the stimuli

Given that the major acoustic correlates of focus are duration, intensity, and pitch maxima and minima, we took the relevant measurements of our prosodically prominent (narrow focus) and non-prosodically prominent (NPP) items. T-tests were conducted, which revealed significance between the two sets of stimuli on duration ($t=2.57$, $p<0.05$) and intensity ($t=9.245$, $p<0.001$). There was an outlier in the pitch max-min difference measurements that, once eliminated, gave us significance ($t= 2.56$, $p<0.051$).

4.3. *Procedures and Methodology*

During the judgment task, subjects listened to a block of six sentences, followed by a word in isolation. They were asked to judge whether the single word had appeared in the previous block of sentences or not. In the target blocks, a target item appeared in one of the six sentences and then appeared after the block in citation form. In the other half, a name appeared that had not been heard in the sentences. Blocks were so constructed that each contained at least one sentence with a common name (either as a filler sentence or as a target) to ensure that subjects could not distinguish target blocks and distracter blocks simply by the presence in the former of common names.

Target and distracter blocks were preceded by three training blocks with feedback. Subjects were told they had 2500 msec to press one of two labeled keys on the computer keyboard once the individual name was presented. The entire test lasted approximately 35 minutes.

Both accuracy measures and latencies were collected.

5. Results

5.1. *Results on German main task: accuracy rates*

The accuracy rates on the targets and the distracters were high and virtually identical: Targets = 73%, Distracters = 71%. A one-way ANOVA examining accuracy rates by the order of language (German < English vs English < German) was not significant. Group results were collapsed for further analysis.

The accuracy on the prosodically prominent and non-prosodically prominent targets was similar: PP = 74.3%, NPP = 71%. Finally, the accuracy results for the second factor, common vs. uncommon names were less similar: 78% and 67%, respectively. Accuracy rates on the target items were submitted to repeated measures ANOVAs with subjects and items as random factors. Both factors were significant: on target items, $F(15, 751515152) = 4.54$, $p = 0.0000$; on subjects, $F(21, .3068818182) = 1.85$, $p < 0.05$.

5.2. *Results on German main task: RTs*

There were 256 correct answers to be analyzed. 12 responses were mechanically excluded because the latency was less than 300 msec. Mean RT of the remaining 244 responses was 965.64 msec. A repeated measures ANOVA showed a main effect for subjects $F(21, 558429.281) = 4.15$, $p = 0.000$ and main effects for items, $F(15, 388017.508) = 2.88$, $p < 0.001$. The prosodically prominent/non-prosodically

prominent contrast was not significant, nor was the common/uncommon contrast.

The variance in accuracy and latency scores can be accounted for by subjects and items. There was no effect for the PP/NP factor, nor was there any effect for the common/uncommon factor. This indicates that focal accent plays no role in the success on our task, which required the segmentation of unknown words.

6. Discussion and Conclusions

Second Language Acquisition studies that consider how learners extract information from the input typically assume that learners will “notice” input items that are salient. It is supposed that saliency will facilitate the learning of these items. We tested this hypothesis by presenting Anglophone learners of German as a second language with sentences containing both familiar and unfamiliar lexical items, in prosodically prominent and non-prosodically prominent positions. Our results show that the relationship between cues in the signal and salience is much more complex than the second language literature appears to acknowledge. Recognizing focal accent in the input requires identifying both the phonetic cues (language specific) and cues to information structure (also language specific). Given this, our results show that learners do not automatically privilege items because they are prosodically prominent. Instead, learners must move through a process of acquisition that allows them to perceive language-specific cues and employ them in a linguistically appropriate manner. Our results demonstrate that the segmentation of a second language and acquiring its information structure requires more than just the perception of acoustic salience.

Bibliography

- AKKER, E. & A. CUTLER (2003), « Prosodic cues to semantic structure in native and nonnative listening », *Bilingualism : Language and Cognition*, 81-96.
- CUTLER, A. & J. D. FODOR (1979), « Semantic focus and sentence comprehension », *Cognition* 7, 49-59.
- CUTLER, A., J. MEHLER, D. NORRIS & J. SEGUI (1983), « A language-specific comprehension strategy », *Nature* 304, 159-160.
- CUTLER, A. & D. NORRIS (1988), « The role of strong syllables in segmentation for lexical access », *Journal of experimental psychology : Human perception and performance* 114, 113-121.
- CUTLER, A. & S. BUTTERFIELD (1992), « Rhythmic cues to speech segmentation : Evidence from juncture misperception », *Journal of Memory and Language* 31, 218-236.

- EADY, S.J., W.E. COPPER, G.V. KLOUDA, P.R. MULLER, D.W. LOTT (1986), « Acoustical characteristics of sentential focus : narrow vs. broad and single vs. dual focus environments », *Language and Speech* 29, 233-251.
- FOWLER, C.A. & J. HOUSUM (1987), « Talkers' signalling of new and old words in speech and listeners' perception and use of the distinction », *Journal of Memory and Language*, 26, 489-504.
- TERKEN, J. & S.G. NOTEBOOM (1987), « Opposite effects of accentuation and deaccentuation on verification latencies for Given and New information », *Language and Cognitive Processes* 2, 145-163.
- WELLS, W.H.G. (1986), « An experimental approach to the interpretation of focus in spoken English », in C. Johns Lewis (ed), *Intonation in Discourse*, London: Croom Helm, 53-75.
-