Salience, Awareness and SLA

S. E. Carroll
University of Calgary

1. Introduction

Input is an important, if ill understood, component of a theory of second language acquisition (SLA). We can detect two approaches to input in the broad range of current SLA research: (i) on one approach, input to language acquisition will be some form of mental representation, ultimately linked to Universal Grammar (UG). This mental representation may just have been computed de novo at the instant of acquisition, or may have been activated from long term memory to be internally altered in some way. In either case, it will combine with some other mental representation to create a novel mental representation. On this approach, input is as “internal” as the result of acquisition (see Carroll 2001, 2002a, b for elaboration). For generativists, in particular, the relationship between what is objectively available in the speech stream and what is acquired is indirect, and not mediated by conscious awareness, explicit instruction, feedback or correction. This approach to input to language acquisition is perhaps clearest in discussions of parameter-(re)setting whereby analysis of some cue to a parameter, possibly pre-determined, triggers the setting of a pre-determined specification (Epstein, Flynn & Martohardjono, 1996; Meisel, 1997, 1998; White, 1989, 2003; Dresher, 2003).1 (ii) On a different approach, the word input refers to purely environmentally available stimuli—sounds in the speech stream, marks on a page, or pixels on a computer screen, etc. For functionalists, the complexity of linguistic cognition emerges from the organisation and structure of speech itself, mediated by the properties of linguistic interactions among learners and native speakers. In such approaches, explaining how learners detect and analyse properties of the input is crucial for there is no UG to guide the language-making faculty. The Noticing Hypothesis (Schmidt 1990, 1995, 2001) is an interesting proposal in this regard. Schmidt has stated that learners must attend to and, moreover, notice something in order to acquire it.

(1) The Noticing Hypothesis (NH)

The noticing hypothesis claims that awareness at the point of learning (Time 1) is required for all learning. (Schmidt, 1995, p.27)

The “noticing hypothesis” states that what learners notice in input is what becomes intake for learning. (Schmidt, 1995, p. 20)

“SLA is largely driven by what learners pay attention to and notice in target language input and what they understand the significance of noticed input to be.” (Schmidt, 2001, p. 3–4)

[what is noticed are] “elements of the surface structure of utterances in the input, instances of language, rather than any abstract rules or principles of which such instances may be exemplars” (Schmidt, 2001, p.5)

---

1 This research has been supported by the Social Sciences & Humanities Research Council of Canada through the Canada Research Chairs program (ref. 905-202408). I gratefully acknowledge the Council's generous support.

This is so even if generative researchers habitually—and erroneously—talk about input as if it were stimuli in the speech signal.

The Noticing Hypothesis (NH) is designed in part to explain the “what” of language acquisition (some properties of the L2 are noticed while others are not), in part to explain the time course of acquisition (learners notice some aspects of language before they notice others). The NH ultimately requires, however, that we have a story to tell about why learners notice what they notice. It is commonplace, therefore, to see the NH coupled with suppositions that if some aspects of language are noticed before others, or indeed noticed at all, it is because they are “salient” in their context. In order to investigate the NH, we must therefore investigate the notion of salience.

In the rest of my paper I will discuss the NH in terms of noticing a sound unit in the signal for the purposes of (prior to) segmenting it and storing a representation of a sound form in long term memory. The fourth quote from Schmidt in (1) suggests that noticing is relevant to segmentation. One may construe the first three quotes as saying that noticing is a prerequisite for learning to segment prosodic forms from the signal. In my view (following Jackendoff, 1987) noticing results when prosodic representations are “projected into awareness.” Noticing forms is, therefore, a result of segmentation, not an input into it. Because the cues to segmentation are language-specific (see Section 2) they must be acquired. There is, consequently, a considerable amount of linguistic acquisition which must occur before the learner can successfully segment the L2 speech stream. This kind of language learning provides a prima facie case against the NH because it occurs well below the threshold of awareness (see Carroll, 2006 for learner reports to this effect).

2. Salience, Prosodic Prominence, Information Structure and Segmentation

It is frequently asserted that learners notice phenomena in the signal because they are salient. Salience is a psycho-perceptual effect which correlates with a number of psycho-perceptual properties. Here we will discuss one only: prosodic prominence. Words which are perceptually salient, by definition, have been noticed by the listener. Words which are prosodically prominent are, by definition, perceptually salient. Not just any word will be perceived as being prosodically prominent. This is because prominence is the psycho-perceptual effect of ways of marking focus, itself an element in an abstract representation called Information Structure.

Speakers package their utterances to facilitate identification of referents, properties, locations, and other semantic constructs which they want to talk about (Haviland & Clark, 1974; Clark & Haviland, 1977; Prince, 1981; Vallduví, 1992; Lambrecht, 1994; see Van Valin & Lapolla, 1997, p.199 for a cursory introduction to a voluminous literature). The purpose of information packaging is to structure one’s speech so that the first concepts that the listener thinks of in a conversational exchange are the ones the speaker intended to convey. Central to a theory of Information Structure are constructs like old (given) information and new information, topic and comment, and focus. When a referent is introduced into a discourse for the first time, it is new information and may be anchored in that it is linked to some more identifiable referent (Prince, 1981). Once in the discourse, a referent becomes identifiable (Chafe, 1995). In Chafe’s model, if a referent is identifiable, it is said to be active if it is the current focus of attention and awareness. If not active, it may be accessible if it is inferable from the situational or linguistic context or inactive if it is represented in long term memory (LTM) but not in short term memory (STM). In this way, Information Structure can be integrated into a set of assumptions about the processing of sentences in discourse, attention, and memory functions. Information Structure is, consequently, a grammatical system with potentially significant processing effects, namely the psycho-perceptual effect of finding some part of an utterance salient.

Focus Structure is the name given to the conventional organisation of the distribution of information in a conceptual structure such that some bit of information is to be asserted against a set of

---

2 There is a tendency among supporters of the NH to sweep all phonetic (and possibly much phonological) acquisition of language under a rug labelled “perception” (Robinson, 2003 can be so construed). It is important, therefore, to stress that phonetic knowledge is language-specific and quite distinct from other kinds of auditory knowledge (such as that which permits us to discriminate barking dogs from barking seals, Canadian police sirens from French police sirens, etc.). In particular, the cues for focal accent are not relevant to discriminating police sirens or barking dogs.
presuppositions which the listener is assumed to know or believe (Lambrecht, 1994, p.213). Focus Structure can be realised syntactically, lexically or phonologically through *focal accent* (Jackendoff, 1972; 1990). The sentences of (2) illustrate narrow focus (focal accent is expressed conventionally in writing through capital letters):

2) Question: WHO ate the cookie that was on the kitchen counter?

   Presupposition 1: There was a cookie on the kitchen counter.
   Presupposition 2: Someone ate that cookie.

   a. JOHN did. (focal accent on the subject NP + verb ellipse + deaccenting of did)
   b. It was JOHN who ate the cookie. (cleft construction + focal accent on the clefted NP + deaccenting of following sequence up to cookie)
   c. Only JOHN would do it. (focus particle only, focal accent on subject NP + deaccenting of following sequence)

Focal accent is a grammatical construct. Focally accented words are perceived by users of a language which marks focus in this way to be prominent and speakers may make certain words prosodically prominent in planning their utterances. It is important to keep in mind that prosodic prominence and focal accent are distinct constructs, which can be seen in the fact that prosodic prominence can result not only from focal accent (as in (2)) but also emphasis, as when a speaker repeats some part of an utterance that a listener did not hear or understand, as in (3):

3) Speaker A: I just put the CAT out.
   Speaker B: What?
   Speaker A: I just put the CAT OUT.

   (deaccentuation of the first part of the sentence, prosodic prominence on the last two words which are not, N.B., a constituent)

In languages like English and Dutch, Anglophones and Dutch listeners are actively searching for focus when processing speech and the identification of focus facilitates the rapid comprehension of the message in that accented words are identified more rapidly than unaccented words (Cutler, 1976a, b, 1997; Cutler & Foss, 1977; Cutler & Fodor, 1979). This facilitation effect is not due, however, to the objective salience of focal accent. This is because the emergence of the psycholinguistic effect requires a “fit” between prosodic cue and Information Structure. This has been shown by the fact that the facilitation of accented words disappears if accented words correspond to old information and unaccented words correspond to new information (Terken & Nooteboom, 1987; Bock & Mazzella, 1983; von Donselaar, 1992). Regardless of how “obvious” and “in the signal” prosodic prominence may appear to us when we listen to speech, it is an *effect* of speech and Information Structure processing. It is not a universal *input* to speech processing.

And what of focal accent? Focal accent is not directly instantiated in the speech signal. Instead, the relationship between focal accent and segmentation strategies must be mediated by phonetic and phonological representations. Prosodic prominence is perceived when a coalition of phonetic properties is detected in the signal and associated with focal accent as a consequence of having acquired the markers of Information Structure of a given language. We must learn the phonetic cues to focal accent. Note that it is not the case that prosodic prominence arises solely from the processing of the physical properties of the signal. When listeners are asked to mark what they perceive to be prosodically prominent expressions in an utterance, a variety of prosodic, lexical and morpho-syntactic features of an utterance correlate to that perception (Streefkerk, 2002).

---

3 This is the standard view. Herburger (2000), in contrast, rejects the view that focus is part of Information Structure and argues instead that focus re-organises in Logical Form (LF) the quantificational structure of unary quantifiers. Focus is thus part of the standard semantics and has no necessary implications for language processing or attention. See also Portner & Yabushita (1998) who also reject the hypothesis that we compute an independent Information Structure in computing meanings.
Focal accent is also not universal; it is not observed in tone languages (Hartmann, 2003), where kinetic tone is marshalled for lexical purposes. Moreover, even in closely related languages whose prosody of sentence accents is very similar, such as Dutch, German or English (Gussenhoven, 1983; Eefting, 1991), English speakers make far more extensive use of focal accent than German speakers do (Féry, p.c. 2004) and focal accent will interact with morpho-syntactic ways of marking focus, which vary according to other syntactic variables (such as presence or absence of V2 effects). Finally, the phonetic exponents of focal accent vary in those languages in which it does occur (Eady, Cooper, Klouda, Müller & Lotts, 1986; Koopmans-van Beinum, 1990; Hirschberg, 1993; Bartels & Kingston, 1994; Ayers, 1995). 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. The minimal conclusion is that there will not be any simple relation between phonetic properties of the signal and the perception of prosodic prominence, which means that there will not be any simple story to tell about how the perception of prosodic prominence aids and abets segmentation in SLA. It certainly seems unlikely that the story will involve pure “bottom-up”, signal-based processing. The fact, already mentioned, that accenting given information and de-accenting new information can actually slow down the perception of words, shows that listeners’ expectation of what is given or new information is reaching deep down into the perception of speech to guide the processing of cues to focal accent. To put the matter differently, we are not noticing sound forms because they are objectively salient, they are perceptually salient because we expect them to be filling certain informational functions.

For all of these reasons, it is highly unlikely that adults will universally be initially sensitive to focal accent when exposed to an L2 for the first time. Beginning 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. The upshot of this for a theory of input is that a learner could make use of prosodic prominence as a cue to word boundaries and as a mechanism to segmentation even if we were to grant—which we are not obliged to do—that prosodic prominence is something that a learner could make use of on-line while acquiring the L2. In short, there is no direct route from attention/prosodic prominence to word boundaries. Nonetheless, I think one might weave a more complex structural connection between focal accent and Information Structure, via the pitch accents which instantiate focal accent in English, Dutch or German, and the phonetic cues to pitch accent, if we could show a universal sensitivity to kinetic tone. Achieving this demonstration is part of the current research activities of my lab.

3. Possible Phonetic Cues to Prosodic Unit/Word Boundaries

Wells (1986) showed that kinetic tone—shifts in the fundamental frequency—is important to the perception of focal accent in British English. Is there any evidence for an initial sensitivity to kinetic tone which might serve as a route into segmentation? The answer is “yes, but very limited evidence.” Henderson and Nelms (1980) exposed Anglophones to Czech auditory stimuli for the very first time. They tested both pause and intonation contours as prosodic cues to sentence boundaries (see Cooper & Sorenson, 1977, 1981, and Cooper & Paccia-Cooper, 1980). Henderson and Nelms showed that their participants revealed no sensitivity to pause as a marker of clause boundaries but were sensitive to kinetic tone.

Indirect evidence comes from a study by Wakefield, Doughtie and Yom (1974), who showed that even 27 minutes of exposure to a language is sufficient for novice learners to begin to discern the prosodic boundaries of clauses. They presented 2 groups of Anglophones with auditory stimuli either from Korean or Mandarin Chinese, consisting of a 3-minute story (repeated 9 times) in which a 1-second pause was inserted in some of the stimuli in such a way as to interrupt a word or a phrase or to

---

4 Hartmann has made this claim in a stronger form for Chadic languages (as here) and in a weaker form (as in the project description). As is frequently the case, the details are complex.
occur between words. Participants were to indicate which kind of pause sounded “more natural”. Subjects were better able to perceive phrase boundaries as opposed to word boundaries. So the subjects of Wakefield et al. did not behave as if pause were a natural cue to word boundaries but did appear to analyse it as a cue to phrase boundaries. Now pause does not normally mark word boundaries in English but it can mark the edges of intonational phrases. The Wakefield et al. results are compatible with the hypothesis that Anglophones are transferring prosodic parsing procedures of the L1. Pilon (1981) conducted a similar study, again with Anglophones listening to Korean stimuli, with comparable results. In other words, Anglophones were sensitive to pause in particular locations, namely at the edges of syntactic (probably intonational) phrases.

In our research, we are working on the assumption that learners may show a preferred mapping between shifts in F0 and segmentable prosodic units. In other words, kinetic tone may be a preferred cue to the “edges” of prosodic units such that even beginners might be sensitive to it. To assess the validity of this hypothesis, we are using a task which involves exposing L2 learners to sentences of the L2 in a block. Subjects subsequently hear a word in isolation and must make a judgement as to whether they have heard the sentence in the block of sentences or not. The words to be detected are all novel items—bird names—embedded in a semantic context dealing with pollution and animal conservation. In the experiments which are currently underway, we have artificially created test items which are prosodically prominent or not. The sentences appear in blocks of 6 sentences as in (4).

4) a. The rain-forest supports carrion-eating caracaras.
   b. Puffins are quite numerous still in Newfoundland.
   c. Predators make Antarctica less safe for penguins.
   d. Accentors are threatened in ski resorts by hikers.
   e. Blackcaps prefer to live in tall trees in woodlots.
   f. Buntings are no longer protected by their drab colour.

ACCENTORS
(expected response “Yes”)

We will use a within–subjects designs with Canadian learners of German performing the task in both English and German, and German learners of English also performing the task in both English and German.

In earlier experiments, which I conducted with Dr. Ruben van de Vijver, testing German-speaking students and a comparison group of English students, we used a between-subjects design. We found that the German subjects receiving prosodically prominent stimuli in English performed significantly better than the German subjects who were tested with non-prosodically prominent stimuli in English. Unfortunately, when I examined the German students on the Oxford English Proficiency Test (Allan, 1992), I found that there was a significant difference in the groups’ scores on the Listening Test. As it

5 A native speaker pronounces the bird name in a modified NP in response to an appropriate question as in (i):

i) Elicited pronunciations
   Accentors are threatened in ski resorts by hikers.
   a. Which birds?
      the fat ACCENTORS
   b. Which accentors?
      the FAT accentors

Our “voices” read the target sentence first silently and then read a question creating either a narrow focus on the head noun (ia) or a narrow focus on the adjective, with ensuing de-accentuation of the head noun (ib). In the strings, the prosodic prominence is perfectly detectable as indicated. We then splice the nouns back into the target sentences to create focally accented and non-accented words.

6 In previous experiments which I conducted in Potsdam, we used blocks of four sentences. The task was very easy, in fact too easy, hence the changes to work currently in progress. We have an equal number of distracter blocks for each target block.
turns out, 70% of the items which this particular listening test tests were cases of prosodic prominent words. So, we had no choice but to assume a pre-treatment difference in our groups and set these results aside, to await the results of the replication study. However, because we tested both our learners and our English comparison group in their L1, we were able to look at the role of prosodic prominence in a case where proficiency differences are assumed not to play a role. We found in each case no effect of prosodic prominence on the task. From this we can conclude that if prosodic prominence is facilitating the segmentation of prosodic words, it is not doing so among monolingual native speakers or L2 users in their L1. This result leads me to expect either that focal accent (and its ensuring perception of prosodic prominence) plays no role in any group of word learners (contrary to speculation on the matter), or else that the nature of processing (viz. the way in which prosodic prominence is used in word segmentation) is changing as linguistic competence improves such that more knowledgeable and more proficient listeners do not need prosodic prominence to identify new words. Such a result would not be entirely surprising (see Akker & Cutler, 2003).

4. Conclusions

Possibly the first step in learning a language is learning to segment the speech signal so that the continuous sound stream is perceived as a linear sequence of sound forms. To reach a phase where one can do this, L2 learners must acquire a complex coalition of cues to prosodic units, including prosodic words. In some languages, cues to focally accented units might be acquired before other kinds of cues to prosodic words. However, this acquisition takes place below the threshold of awareness and presents a prima facie case against the Noticing Hypothesis. Of course, we have not shown that the NH is wrong but a review of the literature defending it seems to presuppose that learners can already segment the speech stream. It certainly is worth exploring what occurs in the very initial phase of language acquisition. Segmentation is a fascinating and complex topic in its own right, and has yet to be taken up empirically in a serious way in SLA research. It is a particularly useful topic to examine to make headway in the development of a theory of input for SLA acquisition.

References


