When is input salient? An exploratory study of sentence location and word length effects on input processing

SUSANNE E. CARROLL

Abstract

Sentence position and word length have been claimed to contribute to the perceptual salience of words. The perceptual salience of words in turn is said to predict L2 developmental sequences. Data for such claims come from sentence repetition tasks that required perceptual re-encoding of input and that did not control for focal accent. We used a more direct measure of input processing, comparing 60 German learners of English to native speakers on a probe detection task. We also controlled for focal accent. With this task, we found no effect of sentence position or word length on word learning. Our results suggest that the relationship between perceptual salience, segmentation and word learning is more complex than previous research suggests.

1. Introduction

1.1. *Input and input processing*

Input and input processing are a corner stone of second language acquisition (SLA) theorising.[1] Despite its importance, there is no consensus on what input

1. Thanks are due to the Social Sciences & Humanities Research Council of Canada (SSHRCC) who funded the writing of this paper through its Canada Research Chairs program (grant no. 950-202408). The data reported on here were gathered through funding from the Deutsche Forschungsgemeinschaft to Project C4 "Prosody and Information Structure as forms of 'input' in second language acquisition", Collaborative Research Centre 632 *Information Structure: Linguistic markers of the organisation of utterance, sentence and text*. The principal investigators of this project were myself and Dr. Ruben van de Vijver. I acknowledge and thank the following people for completing various essential tasks: Ruben van de Vijver did the programming and stimulus editing and was responsible for the final form of the stimuli; Anne Zimmer-Stahl created and recorded the German stimuli and provided the native speaker common/uncommon classification of the German bird names; Anke Sennema collected the data. Various people have offered advice and encouragement including Görla Bruce, Anne

IRAL 50 (2012), 39–67 0019042X/2012/050-039
DOI 10.1515/iral-2012-0002

consists of. Should it be equated with objective and measurable physical properties of the signal, as in the Competition Model (Bates and MacWhinney 1989: 157)? Or does it consist of mental representations that, on previous occasions of learning, have been abstracted and that will, on a new occasion of learning, provide essential information for the construction of novel representations (Carroll 2001)? Alternatively, does it consist of abstract mental representations that are innate and therefore not acquired at all (Lightfoot 2006)? There is fundamental disagreement among researchers as to which answer is appropriate. Similarly, there is no consensus among researchers as to how L2ers process input in order to begin to encode items or rules of the target language. Discussions of perceptual salience and its role in SLA constitute a prime example of such important theoretical disagreements.

1.2. *Input and perceptual salience*

Many researchers have asserted that the perceptual salience of input plays a critical role in early word learning (Larsen-Freeman 1975; Henderson and Nelms 1980; Hatch 1983; Klein 1986: 69; Bardovi-Harlig 1987; Sharwood Smith 1991; Hawkins 2001: 07; Ellis 2003: 65, 2008: 372). By hypothesis, a learner's attention during speech processing will naturally be drawn to some parts of the input, namely parts that are perceived to be prominent.[2] Because of this interaction between attention and prominence, these stretches of speech will be learnt first (Barcroft and VanPatten 1997; Goldschneider and DeKeyser 2001; Rast 2008: 143).

This perspective is widespread, not because it has a solid foundation in studies determining what beginner L2 learners perceive as salient, in acoustic studies of the input causing such perceptions, and/or L2 developmental studies of the emergence of word forms. Instead, salience has been invoked post hoc to explain developmental sequences that have emerged from longi-

---

Cutler, Vincent van Heuven and Caroline Féry, co-ordinator of SFB 632. Audiences at various conferences and colloquia in Banff, Calgary, Pisa, and Thessaloniki have heard bits and pieces of these ideas and commented and criticised. Perspicuous comments by reviewers of this journal were especially helpful in improving the argumentation and form of this paper. Errors of interpretation and fact are the author's alone.

2. There has been no discussion in the L2 literature as to why certain stretches of speech might be perceptually prominent. There are two options: either these stretches of speech are inherently prominent for humans because we have an a priori predisposition to perceive them that way, or else some people perceive the stretches of speech to be prominent because their knowledge systems cause this effect when they process the input. With respect to the usual properties of language mentioned in the works cited, e.g., "stress", "focal accent", "intonation", the second option is the correct one: perceived salience is an output of speech processing entirely dependent on an acquired phonology.

tudinal or cross-sectional studies analysing *production* data. Thus, perceptual salience has been called upon to explain the well documented fact that L2 learners produce lexical roots and stems before they produce inflectional suffixes and functional categories like determiners, auxiliaries, and so on (Klein 1986: 69). Perceptual salience has also been invoked post hoc to explain the order of emergence within the set of functional categories (Larsen-Freeman 1975; Goldschneider and DeKeyser 2001). Such research provides no data on what the learners who produced the L2 data analysed actually perceived as salient in the input, indeed, there are no perceptual data at all in this research.

It may strike native speakers of English as just common sense that lexical expressions are more salient than functional categories. After all, *we* hear the difference between focally accented lexical expressions and functional categories that are de-accented.[3] However, lexical roots and stems differ from functional categories at every level of representation: in the kinds of concepts they encode, in their selectional and co-occurrence properties, in the positions they occupy within phrases and sentences, as well as in their prosodic properties. That it is the prosodic properties of words on which the developmental facts hinge, i.e., the only properties that are "in the input", remains to be demonstrated. To the best of my knowledge, we do not know yet if L2 learners process unaccented, unstressed and reduced syllables contained within a prosodic word, e.g. [johane] *Johanna*, differently than they process unaccented, unstressed and reduced syllables that precede another prosodic word, e.g., [ajnəfʁaʊ] *eine Frau* 'a woman'.[4] The same arguments can be extended to the analysis of development sequences within the set of functional categories.

To move beyond post hoc accounts of developmental sequences, we need to understand "input" as a theoretical construct and to empirically investigate how learners process speech. Consider, as an example, that there are conceiv-

3. That functional categories are typically de-accented is a consequence of information structure and the kinds of topics that we choose to talk about. It is well-known from the literature on focal accent that in the right context, where a speaker is clarifying what she said, even grammatical functors can be accented:

   (i) I said DON'T do it.

4. Research from my laboratory shows that even ab initio learners (Anglophones exposed to German) can, on the basis of minimal input, segment, represent and reproduce unstressed syllables that are contained within prosodic words, either finally or word-internally. In the text, it is suggested that functional categories positioned to the *right* of a lexical expression may be prosodically parsed as the right edge of a prosodic word, so that, e.g., *das ist Sonja* 'that is Sonja' will be prosodically parsed as (dasıst) (zonja) If this line of thinking is correct, we predict quite different learning trajectories for functional categories that are sentence-initial, or that precede lexical expressions within intonational phrases versus those that follow lexical expressions.

ably many distinct ways that attention and perceptual processes analysing the signal might interact. If one continuous stretch of speech is perceptually more salient than another, it might be more rapidly segmented from the signal (perhaps because the acoustic cues to boundaries with neighbours become clearer). On this story, salience helps the learner to create initial sound forms (exemplars) of a word. These exemplars could be acoustic representations, reproducing many of the properties of the signal. Alternatively, if a sequence of sound units, say a suite of syllables, is perceptually more salient than another sequence of syllables, the phonological content of these units might be more accurately abstracted from the acoustic variants the input actually provides on different occasions. On this story, perceptual salience leads to better identification of an abstract phonological unit (here a sequence of syllables that stand for a class of acoustic exemplars). If it turns out to be the case (the matter is one of considerable debate among psycholinguists, see Frauenfelder and Lahiri 1989; Lahiri and Marslen-Wilson 1991; Norris et al. 2006) that the lexical access codes used in word recognition depend on the creation of such abstract phonological forms, learning such a form from a variety of phonetic exemplars might subsequently lead to more accurate recognition of the word under variable pronunciations. Surely these kinds of processing are part and parcel of what is meant by "learning a word". Each hypothesis is compatible with the idea that attending to the input leads to more accurate processing, but each makes different predictions about the course of early word learning. Moreover, each one makes quite different claims about exactly what it is the learner is attending to. In the first case, the learner is attending to the physical signal, and a causal story involving attending to the input can go through. In the second case, however, either the learner isn't attending to anything at all in the signal (he is eliminating from his representation-in-long-term-memory exemplar-specific detail to create a phonetically underspecified prosodic structure), or else the learner is attending to a mentally represented prosodic unit (syllables) in an on-line computation. However, in order to attend to an *already* mentally represented prosodic unit, we can conclude both that that unit has already been acquired (it is mentally represented) and that attention is not being directed to something in the input. Rather, it is directed to something *in the mind*. Either way, the second scenario vitiates the perceptual salience story as a causal account of language development.

1.3. *Salience and sentence position*

When we examine what properties of input have actually been investigated with respect to perceptual salience, we see that they involve phenomena that are possibly many processing steps removed from the signal. Barcroft and Van-

Patten (1997) have claimed that the position of a morpheme in a sentence can make it more or less perceptually salient. They formulate the Sentence Location Principle in (1):

(1) Sentence Location Principle (SLP)
Learners tend to process items in sentence initial position before those in final position and those in medial position. (VanPatten 1996: 32, 2004: 14, 2007: 127)

The Sentence Location Principle (SLP) is embedded in a theory of second language acquisition. The temporal preposition "before" in (1) is therefore not making a claim about the time course of speech processing, which would be a trivial claim given the incremental nature of speaking. Rather it predicts that morphemes will emerge in developmental sequences according to where they typically occur in the sentence. In their study of English learners of Spanish, Barcroft and VanPatten find support for the SLP. Rast (2008: 156–157) also finds sentence positional effects on word learning in a study of ab initio Francophone tutored learners of Polish. There are, however, empirical, theoretical and methodological reasons for scepticism about such positional effects based on the data provided.

On the empirical front, taken at face value, the SLP predicts that English-speaking learners of French should learn words like *le*, *un* or *des* or words like *et*, *puis* or *et puis* before they learn anything else. Not only do these words typically begin utterances, they are also very frequent in conversation. This prediction is patently false. To rescue the claim, it is critical, therefore, to know what the term "items" in (1) refers to. There may be independent reasons, having to do precisely with the nature of speech processing and the prosodic properties of words why functional categories like *le, un*, or *des* might be learned later than particular nouns, verbs and adjectives like *homme*, *achète* or *pain* (see Note 4). Functional categories are often instantiated as weak syllables gathered into a neighbouring prosodic word or prosodic phrase. Such patterns define the typical patterns of rhythm in stress-timed languages like English. Suppose that the perception of abstract rhythmic domains by English-speaking L2 learners helps to define the input to language learning mechanisms. This might mean that a learner starts carving up the signal at the first strong syllable of an utterance, in which case she might not represent the acoustic material furnished by an utterance-initial *un*, *le*, or *des*, contra the SLP. Alternatively, suppose that English-speakers represent all of the acoustic material in an utterance, even at initial stages of learning, but are critically sensitive to strong syllables as cues to word boundaries (Cutler et al., 1983; Curtin, Mintz and Christiansen, 2005). In this scenario, the issue is that some words are more likely than others to be correctly represented as prosodic words. In particular, functional categories that occur to the right of stressed syllables might be segmented as part of the prosodic word containing the morpheme *to the left*. In this case the point is not that prosodic parsing has eliminated prosodic material from the input to the language learning mechanisms but rather that the segmental identity of the morpheme is contained within a prosodic structure that disguises its status as an independent word. How the learner comes to identify two separate morphemes only one of which can be a prosodic word is an important question.[5]

Theoretically, positional effects are supposed to arise because of the nature of working memory. Thus, VanPatten (1996) invokes studies of recall of memorised word lists that show primacy effects (the words at the beginning of the list are recalled better than those in the middle) and recency effects (the words at the end of the list are recalled better than those in the middle). VanPatten's use of this literature to buttress his claims about L2 developmental stages has, however, been critically examined and found wanting in several respects, in particular, in his reliance on a limited-capacity, single-resource model of attention (DeKeyser et al. 2002). Also problematic is the fact that claims about positional salience are not interpretable without a clear definition of the positions involved. However, "initial position", "medial position" and "final position" are not well-defined in any of the studies cited. Results will differ considerably depending on whether we define position acoustically, prosodically, or syntactically.

On the methodological front, neither Barcroft and VanPatten (1997) nor Rast (2008) controlled for information structure in their studies. Sentence position is known to interact with information structure both in languages with focal accent and in those without (Erteschik-Shir 1997, 2006). Thus, in declarative sentences in English, focal accent will fall on the last accentable syllable in the intonation phrase, which will often occur at or just before the end of the sentence (Ladd 1996). A second locus of prosodic prominence may occur at or near the beginning of a simple declarative sentence if it contains a heavy subject, viz. one bearing its own intonational phrase. It is therefore a critical question whether we perceive different parts of a sentence to be prominent just because of their position, or because we are using our knowledge of the prosody of a language to actively search the input to locate cues to focal accents (Cutler 1976a, b, 1988) – and these focal accents are more likely to occur at the beginning or at the end of a utterance.

Claims about positional salience have all come from repetition tasks (Klein 1986; Barcroft and VanPatten 1996; Rast 2008). A cautious interpretation of the data would assert no more than that these studies provide evidence that learners repeat sentence-initial and sentence-final words in a repetition task

5. A question that is not elucidated by labelling prosodic analysis as "chunking", see, e.g., Ellis (2003), MacWhinney (2005).

more readily than words in the middle of sentences. Caution is indeed required because repetition tasks demand that a learner transform what she perceives (either an acoustic or a more abstract prosodic representation) into a motor-articulatory code. This task may well require of learners that they rehearse the input in order to retain it in phonological memory (Baddeley et al. 1998). If a listener begins rehearsing what she has heard at the beginning of an utterance, this might interfere with the perceptual processing of the rest of the stimulus. Or, if she waits to the end of the utterance to begin rehearsing what she has heard, she may have forgotten the initial and medial sections of the utterance (Papagno et al. 1991). Thus, the effects observed may say more about the nature of phonological memory and rehearsal strategies on repetition tasks than they do about processing input for word learning purposes. It is therefore important to test claims about the perceptual prominence of sentence position, using tasks that do not require re-encoding of the input for speech production. The methodology of our study meets that requirement.

1.4. *Salience and word length*

Word length has also been claimed to affect perceptual prominence (Goldschneider and DeKeyser 2001). The Goldschneider and DeKeyser study was a meta-analysis of existing L2 studies and did not directly test perception. In a perceptual study, Streefkerk (2002) showed that Dutch listeners treat Dutch polysyllabic words as more prosodically prominent than single-syllable words and this preference remains even when the grammatical category (i.e., lexical versus functional category status) of words is controlled for.[6] Thus, native speakers found longer words perceptually more salient than shorter words. On the other hand, we should be careful before extrapolating from studies with expert native listeners to conclusions about learners. The learner must construct novel representations and automatise the processing of the input until she is skilled and efficient at it (Segalowitz et al. 1998; DeKeyser 2001). Phonetic studies of speech show that stressed syllables are more likely to be pronounced canonically (de Jong 1995). If a syllable is pronounced canonically, it will

6. Streefkerk (2002) had subjects listen to, e.g., two playings of *De vliegtuigkaping werd tijdens de vlucht opgelost* 'The airplane hijacking was resolved during the flight'. Subjects saw written versions of the sentences on a computer screen. Under each word on the monitor was placed a button and subjects clicked on any button when they perceived the word as prominent. Scores were summed for all subjects and each presentation of a sentence. Lexemes were typically longer than functional categories but the study included both monosyllabic lexemes and polysyllabic functional categories.

exhibit less phonetic variability.[7] Thus, a word consisting of a single stressed syllable may be pronounced less variably in the input to learners and they might find it easier to accurately encode a representation of the word's consonants and vowels. In short, L2 learners might find it easier to map acoustic input into a phonological form if a word consists of a single stressed syllable.

Rast (2008: 146) tested word length in her study of tutored Francophones learning Polish. In this study, learners performed a repetition task prior to any instruction, after four hours of instruction and again after eight hours of instruction. She found no effect of word length in any of her participants (Rast 2008: 151). However, as noted above, the matter should be investigated with a task that does not require phonological re-encoding of what the learner has segmented from the input.

1.5. *Hypotheses of this study*

Three distinct positional hypotheses can be formulated:

(2) a. Sentence Position Hypothesis 1:
Subjects will segment and recognise novel words least accurately when they occur in sentence medial position.
b. Sentence Position Hypothesis 2:
Subjects will segment and recognise novel words best when they occur in sentence initial position (due to primacy effects).
c. Sentence Position Hypothesis 3:
Subjects will segment and recognise novel words better in sentence final position than in sentence medial position (due to recency effects).

Our discussion of word length leads to two distinct and contradictory Word Length Hypotheses:

(3) a. Word Length Hypothesis 1:
Subjects will segment and recognise novel polysyllabic words sooner than novel monosyllabic words.
b. Word Length Hypothesis 2:
Subjects will segment and recognise novel monosyllabic words sooner than novel polysyllabic words.

---

7. That variability of pronunciation is an issue is well known to both phoneticians and sociolinguists and has been shown dramatically by Johnson (2004), who demonstrated in an analysis of conversational American English that most words are not canonically pronounced.

We compared L2ers to native speakers; we predict that the native speakers will outperform the L2 learners due to their greater linguistic competence and better processing abilities. We also used our L2ers as their own controls. For the same reasons (greater competence, more skilled processing abilities), we predict that the L2ers will perform better in their L1 on the task than in their L2.

2. Experiment

2.1. *Design*

60 German learners of English were tested in their L1 and L2; a comparison group of 22 Anglophones were tested in English only. Participants were divided into three sub-groups depending on whether they heard focally accented target items or not.[8] The experiment thus has a mixed design, with focal accent (focally accented target words presented to Group1, de-accented target words presented to Group2 and Group3), learner status (L2 learner vs. native speaker), and language (English vs. German) serving as between-group variables.[9] Word length (monosyllabic versus polysyllabic), position of the word in the sentence (initial, medial, final), position of the sentence in the presentation block (first, second, third, last) were the within-subjects variables.

2.2. *The task*

2.2.1. *A forced-choice receptive task.* We developed a forced-choice probe identification task in which participants were required to listen to a block of

8. This was done because the primary goal of the project from which this data are drawn was to ask: Do learners learn focally accented words better than de-accented words? The question is rooted in the same developmental logic as that discussed in the text, namely that learners will learn better words that are perceptually salient. Analysis of the accuracy data by accent group revealed that the Anglophones tested on focally accented stimuli actually scored *lower* than one group exposed to de-accented stimuli (Group2) and only marginally higher than the other group (Group3). Among the Germans, the group exposed to focally accented stimuli scored about the same as the Group3 and both of these groups were higher than the Group2. Separate one-way ANOVAs revealed these differences to be not significant (Anglophones: $F1(2,19) = 1.58$, $p = 0.23$, n.s.; Germans: $F1(2,56) = 1.80$, $p = 0.17$, n.s.).
9. Two-types of de-accented stimuli were constructed. In one case, focal accent appeared elsewhere in the same phrase, e.g., [*The orange BEAK of the puffin*] *has made it a famous character in children's books*; in the other case, focal accent appeared elsewhere in the sentence, e.g., *PHOTOGRAPHY has helped enormously to encode species like* [*the fillmore*] *which are shy*.

four sentences and then to a single word heard in isolation.[10] Participants had to decide if the single word (the probe) had appeared in the previous block of sentences. The probe was pronounced in citation form, meaning that it was focally accented and surrounded by pauses.

(4) A sample block
   a. Albatrosses are blown from their breeding grounds by more frequent, raging storms.
   b. Once pets on large estates, mandarins have escaped into the wild.
   c. Because they swarm in very large flocks, we cannot count the starling.
   d. Shelducks have long necks and can be confused with swans when counting birds.

   MANDARINS

To perform the judgement correctly, participants presumably had to compare a mental representation of this form with a newly created representation of that word heard several seconds before in a quite distinct phonetic context (exhibiting co-articulation effects from neighbouring words and pronounced with or without focal accent). We assume that this necessitated segmenting the novel form from its acoustic environment.[11]

2.2.2. *Preparation of the stimuli.* Our task was intended to be a word learning task, not a word recognition task. We therefore needed to control for prior

---

10. This is a novel task, whose properties are not familiar from previous research on this topic. Altenberg's (2005) study was designed to study transfer of L1 allophones and L1 phonotactic constraints as well as L1 acoustic cues to the segmentation and recognition of juncture pairs occurring in a constant phonetic context: "Say ____ again". For obvious reasons, this task does not lend itself to a study of the perceptual salience of sentence position. Similarly, I rejected the use of a phoneme-monitoring task, such as that used in Cutler's many studies of segmentation (e.g., Cutler 1976a, b) because it is designed to shed light on whether word boundaries become available *before* or *after* word recognition. Thus, they presuppose that the listener is listening to known words and not constructing novel phonological representations in real time.
11. It is logically possible that participants might have responded by holding acoustic representations of all four sentences in the short-term phonological memory (STM) and compared the representation of the isolated word with those. Given the length of the utterances used (mean length English = 15.2 morpho-syntactic words and 22.1 syllables; mean length German = 13.0 morpho-syntactic words and 23.5 syllables), this is an implausible alternative. This is because Williams and Lovatt (2003: 72) report that there is consensus among researchers on memory that information about exact forms (the exact words used, their order, and so on) can be retained in STM for approximately 2 seconds. Our procedures would clearly exhaust this capacity. See Baddeley (1986) for further discussion.

knowledge of words without exposing our subjects to the forms we wanted to use. One way to do this is to present subjects with nonce forms. However, studies that use nonce forms may not be able to adequately control the semantic and syntactic properties of the input. Consider, for example, studies where learners are presented with pictures of novel objects having unfamiliar shapes or colours, and are asked to learn the names of the objects, perhaps having been told that the objects are "alien objects". Typically, such pictures will not involve the object in use (to avoid activating an L1 translation equivalent or an L1 superordinate word). Yet, in learning the vocabulary of a new language, learners typically make use of the context to map novel forms onto existing lexical entries because the new form is deemed to mean "the same thing" as an existing L1 word (Cammarota and Giacobbe 1986; Ringbom 1987; Odlin 1989). Moreover, grammatical frames that a novel word occurs in help learners to identify the semantic field or basic semantic category that a word belongs to (Brown 1957; Fisher et al. 1991). If learners cannot make use of an acquired theory of entities to classify a depicted thing, precisely because the depicted thing is completely novel and presented out of context, the study is unlikely to shed light on normal L2 word learning processes. We therefore decided to make use of real words from an open-ended class where most listeners would be likely to have limited real-world knowledge of the referents (namely bird names). This permitted us to control both the syntactic and the semantic class of the items. We decided to use both morphologically complex and simple English and German names.

Bird names are highly infrequent nouns. (See Appendices 1 and 2 for frequency counts from the CELEX database, Baayen et al. 1995) Still, frequency in the case of bird names is not a good criterion for determining whether a word is likely to be known or not. The English are likely to be thoroughly familiar with culturally common bird names like *sparrow* or *swan*; Germans will be familiar with names like *Adler* and *Meise*. Galbraith and Underwood (1973) demonstrated that university students can classify concrete words according to their subjective frequency in a remarkably accurate way. This seemed like a reasonable point of departure to develop an initial list of stimuli. The author therefore selected English items to be either culturally familiar or unfamiliar, using her intuitions. A native-speaker of German did the same classification for the German names. This classification gave sets of *common* and *uncommon names*. We hypothesised that while learners might already know some of the common bird names, they were unlikely to know uncommon names unfamiliar to even native speakers.[12]

12. A subset of our written lists of bird names were subsequently submitted to separate samples of English ($n = 23$) and German ($n = 23$) native speakers, asking them to circle any words

Several hundred English and German bird names were devised by consulting specialist literature (Ede et al. 1965; Flegg et al. 1993) and websites (Birds of Britain; Multilingual Birdsearch Engine). We deviated from this practice only to create a very small number of distractor names to complete lists requiring specific properties, for example, that they be monosyllabic. See Appendices.

From these lists, 48 items were selected: 24 names served as target items; 24 names served as distracters. 23 English items were common names and 25 were uncommon names. 25 German items were common and 23 were uncommon. Target items were selected to be monosyllabic and polysyllabic words and were morphologically simple.

Carrier sentences were then constructed in each language, making use of information about the birds from the same sources. An effort was made to create sentences with a certain degree of semantic coherence. These sentences contained information about habitat, ecology, bird feeding, breeding and migratory habits.[13] Sentences were so constructed that the reference to the birds was generic. In both English and German, generic reference can be conveyed by using plural nouns, e.g., *Bay ducks are more likely to be seen in open estuaries* or by the use of a definite article and a singular noun, e.g., *Modern farming has destroyed the grasslands where we find the bustard*. We made use of both kinds of generic in order to be able to have monosyllabic nouns in German.[14] All the names were thus morphologically and/or syntactically cued in the stimulus sentences as to word class.

In half of the blocks ($n = 24$), a target item appeared in one of the four sentences and then appeared after the block in citation form. Participants were

---

that they had not heard or read before. Our instructions allowed subjects to treat as familiar compound names one of whose component words they recognised. Thus, a subject could treat, e.g., *bay duck*, as familiar on the basis of her recognition of the words *bay* and *duck* although she had not previously heard the string *bay duck* used to refer to a particular bird. This way of proceeding should overestimate subjects' knowledge of bird names. Results are shown in the Appendices. On the whole, target words that we initially treated as likely to be unknown, were indeed not known to most of our judges, while words that we had assumed as likely to be known were indeed known to our judges. These results provide some additional support to our claim that the uncommon words were unlikely to be known by the subjects of the experiments reported on here.

13. Nonetheless, the sentences were recorded as individual sentences. Time constraints did not permit smoothing of the intonation contours to create a 'paragraph intonation', a step which would have increased the naturalness of the materials and hence the ecological validity of the task. We thank Vincent van Heuven for pointing out this weakness in our materials and its effects on perception.
14. Both English and German can create plural nouns through affixation of an /s/ consonant, e.g., *kite/kites* and *Drongo/Drongos*, as well as through affixation which creates an additional syllable, e.g., *thrush/thrushes* and *Huhn/Hühner*. Affixation which creates an additional syllable is common in German, which forced on us the use of the definite article + singular noun generic. See Lieber (1980), Köpcke (1988) and Eisenberg (2004) on German plural formation.

Table 1. *Distribution of English target and distracter items by length of word, position of word in sentence position and position of sentence in block*

$n = 48$	Monosyllabic ($n = 21$)						Polysyllabic ($n = 27$)					
Distracters ($n = 24$)	Common 3			Uncommon 6			Common 9			Uncommon 6		
Targets ($n = 24$)	I	M	F	I	M	F	I	M	F	I	M	F
1st sentence ($n = 8$)			1		2	1	1	1	1	1		
2nd sentence ($n = 5$)	1			1				1			1	1
3rd sentence ($n = 5$)	1				1	1			1	1		
4th sentence ($n = 6$)		1	1	1			1				1	1
Totals	2	1	2	2	3	2	2	2	2	2	2	2

I = Initial position; M = Medial position; F = Final position

expected to respond "yes" to these target blocks. In half of the blocks ($n = 24$) a bird name was pronounced in citation form after the block and was unrelated to the bird names in the prior four sentences. Participants were expected to respond "no" to these distracter blocks. Target blocks contained one target sentence and three fillers, while distracter blocks contained four fillers. Blocks were so constructed that each block 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. The distribution of the word and sentence properties for the experiments is shown in Tables 1 (English) and 2 (German).

The order of sentences was fully randomised within each block and the order of blocks was pseudo-randomised. To control for episodic effects due to the order of the items, two lists of stimuli were prepared. The order of items in List 2 was the reverse of the order of items in List 1.

The carrier sentences in which the bird names appeared were designed to express different foci (naturally elicited) and were written in minimal contexts (question/answer pairs) prior to being recorded by our talkers. For the focally-accented sentences, questions were constructed to elicit an answer with narrow focus on the target item. To elicit the de-accented sentences, we asked broad focus "What happened?" questions. In the case of sentences for Group2, focus occurred within the phrase containing the bird name but the target bird name was itself de-accented. In the case of sentences for Group3, focus occurred on another constituent of the utterance altogether.

English stimuli were recorded in standard southern British English by a female talker. This variety of English is officially the target language in the Ger-

Table 2. *Distribution of German target and distracter items by length of word, position of word in sentence position and position of sentence in block*

$n = 48$	Monosyllabic ($n = 18$)						Polysyllabic ($n = 30$)					
Distracters ($n = 24$)	Common 3			Uncommon 6			Common 10			Uncommon 5		
Targets ($n = 24$)	I	M	F	I	M	F	I	M	F	I	M	F
1st sentence ($n = 7$)	1				1	1	1	1	1	1		
2nd sentence ($n = 7$)					1	1	1	2	1	1		
3rd sentence ($n = 5$)		1		1					1		1	1
4th sentence ($n = 5$)			1	1			1				1	1
Totals	1	1	1	2	2	2	2	3	4	2	2	2

I = Initial position; M = Medial position; F = Final position

man school system, so it was presumed that all of the L2 learners would be familiar with this accent. German stimuli were recorded in standard German by a female talker. Recordings were made separately of target and distracter words, filler sentences and target sentences. In particular, to elicit the target sentences, the talker had to read the prepared question/answer pairs. She read the question to herself silently and then pronounced the appropriate response out loud. All recordings were done using a type 4033 microphone, at an audio sampling rate of 22.05 kHz in the lab of the Linguistics Department of the University Potsdam.

2.3. *Partipants*

60 adults (42 females and 18 males), all students of linguistics, psychology or English studies at the Universität Potsdam participated in the study and constitute the L2 learners. They were all native speakers of German and all had studied English in the school system (starting in grade 5 or grade 7). These particular students are required to be competent users of English in order to succeed in their university studies and were not linguistically naive. Some of the students participated in the study to meet a program requirement while others were paid for their participation. Also participating were 22 adults (18 females, 4 males) who were students either at University College London or at the University of Essex, all native speakers of English. They constituted our English native-speaking group. These subjects were paid for their participation. Within each language group, subjects were randomly assigned to the accent groups.

2.4. *Running the experiment*

2.4.1. *Apparatus and software.* Three different computers were used for presentation of stimuli: two laptops and a personal computer. The personal computer and one laptop were used with the German L2 subjects; a second laptop was used to collect the data in Great Britain. The experiment was controlled by a program created using the DMDX experimental software (version 3.0.2.4), software programmed by Jonathan Forster at the University of Arizona.[15] It manipulated WAV files which were created with PRAAT (version 4303).

2.4.2. *Procedures.* The tasks were run with the German L2 learners in two sessions, approximately one week apart. In the first session, subjects answered a basic questionnaire that included a self-estimate of their English language skills. In this session, they also did the Oxford English Placement Test (Allen 1992). In the second session, subjects performed the main experimental task listening to stimuli presented binaurally through studio-quality headphones. All participants were tested individually while seated in a quiet room. Half of the L2 subjects performed the main task first in English; the other half performed the main task first in German. Procedures were the same for the Anglophones with the difference that they did not take the Oxford English Placement Test and performed the judgment task in English only.

Training trials preceded the main experiment (3 blocks, 12 sentences). Written and verbal instructions instructed subjects to listen carefully to a block of four sentences and to indicate as quickly as possible whether a word pronounced in isolation afterwards had occurred in the block. Subjects were told they would have 2500 ms to press one of two labelled keys on the computer keyboard once the individual name was presented, after which a screen-centred fixation cross appeared to indicate that the next trial would begin. Feedback as to accuracy was provided in the training blocks. No feedback was provided following responses to the experimental stimuli. All trials were experimenter initiated. Two short pauses occurred after approximately one-third and two-thirds of the material had been presented. The judgment task lasted about 35 minutes.

15. DMDX is a member of the DMASTR family of experimental software developed at Monash University and at the University of Arizona by K. I. Forster and J. C. Forster.

3. Results

3.1. *Measures*

Two measures were gathered: accuracy rates and reaction times (RT). We report here on the accuracy rates. The RT data will be discussed only insofar as they shed light on whether the learners processed the common and uncommon nouns differently, i.e., to help motivate our choice of stimuli.

3.2. *Scoring*

The Oxford English Placement Test (OEPT) is a standardised test of English, commonly used to sort learners of English into proficiency groups. It consists of three independent tests: a listening comprehension test (OLT) and two grammar tests (OGT). Means scores were calculated for the combined results (max. possible = 200) and for each independent test. Responses on the forced-choice probe task were automatically scored as "1" if correct or as "0" if incorrect.

3.3. *Statistical tests*

Correct responses were summed for both participants and items. Data were submitted to analyses of variance (ANOVA). Alpha was set in all analyses at 0.05.

3.4. *Language proficiency*

The mean combined score on the OEPT was 157.6/200 or 78.8 % ($s.d. =$ 19.55). A combined score of 150 is defined as "post intermediate competent user". There was, however, considerable variability among the scores, with a range from 110 to 190. A score of 110 is defined as "elementary: limited user", 160 to 170 is defined as "advanced: proficient user", while a score of 180 is defined as "very advanced: highly proficient user". These classificatory labels are comparable to the self-ratings that the German students gave themselves on a five-point scale (1 = very good, 5 = very poor); the typical rating was 2.6 or "good". Thus both on the standardised proficiency test and on the self-rating, the German students appear to be advanced learners of English.

To ensure homogeneity of English language proficiency across the experimental groups, OEPT test scores were analysed by stimulus group. Group1 obtained a combined mean score of 161.3/200 or 80.7 %, $s.d. = 17.27$ (range = 132 to 186, or 66 %–93 %), Group2 obtained a combined mean score of

Table 3. *Frequency of scores by range of the L2 treatment groups on the sub-parts of the Oxford English Proficiency Test*

	Listening test			Grammar test		
Scores range	Group1	Group2	Group3	Group1	Group2	Group3
21–30	–	–		–	1	
31–40	–	–		–	–	
41–50	–	–		1	3	
51–60	–	–		4	3	5
61–70	–	2	1	3	2	6
71–80	1	6	1	5	5	2
81–90	10	7	10	7	6	4
91–100	9	5	8	–	–	3
Total	20	20	20	20	20	20

Table 4. *Mean scores of the L2 learners by group on the sub-parts of the Oxford English Language Proficiency Test*

Accent Groups	Listening Test max. = 100 (s.d.)	Grammar Test max. = 100 (s.d.)	Combined scores max. = 200 (s.d.)
Group1	.88 (5.7) Range: 76–99	.72 (13.1) Range: 48–90	161.3 (17.2) Range: 132–196
Group2	.82 (8.6) Range: 61–93	.67 (16.7) Range: 30–90	150 (21.2) Range: 110–181
Group3	.84 (7.4) Range: 64–96	.73 (13.1) Range: 54–96	161.5 (18.6) Range: 118–190
Total	.86 (7.7)	.71 (14.4)	157.6 (19.5)

150/200 or 75 %, s.d. = 21.21 (range = 110 to 181 or 55 %–90.5 %), and Group3 obtained a combined mean score of 161.55 or 80.7 %, $s.d. = 18.65$ (range = 118 to 190). Group1 and Group3 can be defined as at the very low end of the "advanced: proficient user" classification; Group2 was somewhat lower in proficiency. Table 3 shows the range of scores by treatment group. Table 4 shows the mean scores.

A one-way ANOVA revealed no significant differences among the groups ($F(2,57) = 2.46$, $p = 0.09$, n.s.).

The Oxford Listening Test (OLT) merits a closer look. It requires test-takers to listen to a series of sentences and to choose for each test sentence from a pair of target items offered as "what was heard". Target items consist of words or lexical strings which have been shown to give rise to word or phrase recognition problems. Items consist of rhymes and other confusable pairs. One of the project staff listened to the OLT and circled all of the items she heard as prosodically prominent, attempting to replicate the method used in Streefkerk's (2002) study. Analysis of the results showed that over 70 % of the test items were perceived to be prosodically prominent, either because they instantiate the focal accent of the sentence or because they are the locus of a secondary source of prominence. The OLT, therefore, bears a strong resemblance to our task in that the test-taker often has to listen for a prosodically prominent word embedded in a sentence.

As Table 3 shows, subjects clustered at the higher range of scores on the OLT but exhibited a broad range of scores on the grammar test (OLG). Mean scores were much higher on the OLT than on the OGT. Both Group1 and Group3 scored higher on both parts of the test than Group2. Separate ANOVAs on the two test sub-parts revealed no significant difference among the groups on the OGT ($F(2,57) = 1.10$, $p = 0.3406$, n.s.), but a significant difference among the groups on the OLT ($F(2,57) = 4.20$, $p < .05$). Comparison of means (with Bonferroni correction) indicated that both Group1 and Group3 were significantly different from Group2 on the OLT.

Inspection of the data revealed that this difference was due to one subject in Group2 whose score on the OLT was more than 2 standard deviations from the mean. This subject's data were removed from the data set and scores re-calculated. The adjusted mean of Group2 then equalled 83.7, $s.d. = 7.19$ (range: 69–93). Adjusted OLT data were submitted to a one-way ANOVA: ($F(2,56) = 3.08$, $p = 0.054$), which is just above the chosen alpha. Although the subject whose data were depressing Group2's OLT scores actually performed extremely well on the English experiment, it was decided to remove her data from the English data set.

3.5. *English language results: Comparison of native speakers to L2 learners*

3.5.1. *Data cleaning.* Mean accuracy rates were calculated for Anglophone and German groups for each of the 48 blocks, first by participant and then by items. The combined mean was 81.8 % ($n = 3888$; $s.d. = 0.898$, range: 62.5 % to 100 %). Among the 22 Anglophone participants the mean was 86.7 % ($n = 1056$; $s.d. = .0982$; range: 65.48 % to 100 %). Tests of skew and kurtosis revealed a normal distribution. When Anglophone responses were analysed in terms of items, the mean was 86.2 % ($s.d. = .1150$, range: 54 % to 100 %). This distribution was not normal, two items having means more than 2 standard deviations from the mean. These items were adjusted upwards to within 2 standard deviations of the mean (= 63.2%).

Among the German participants, the mean was 80 % ($n = 2832$; $s.d. =$ .0799; range = 62.5 % to 93.75 %). Tests of skew and kurtosis on the participants' data revealed a normal distribution. Analysis of the Germans' responses by item (mean = 79 %, s.d. = .1731, range: 37 % to 100 %), however, again revealed a non-normal distribution, one item being more than 2 standard deviations from the mean (= 42.5 %). Its mean was adjusted upwards.

Separate two-tailed t-tests conducted on the two lists of stimuli showed the differences in the means not to be significant (Anglophones: $t(20) = -0.5145$, $p = 0.61$, n.s.; Germans: $t(57) = -0.9467$, $p = 0.34$, n.s.). Responses from the two lists were collapsed for further analysis. German subjects had been tested on one of two computers; a two-tailed t-test with unequal variance revealed differences in means not to be significant ($t(30.9) = .8134$, n.s.). Data were collapsed for further analysis.

In addition to collecting accuracy responses, we also timed the responses. Four responses that were less than 300 ms were removed by hand from the data set. Responses on the probe task were longer than 2500 ms were automatically counted as errors. The total frequency of such errors was 239 out of a total of 3884 responses, an error rate of 6.15 %. They constitute 34 % of the errors.

3.5.2. *Target and distracter blocks combined.* Responses by L1 group (Anglophones vs. Germans) are shown in Table 5.

Accuracy on the English word identification task was high: mean = 81.88 % ($s.d. =$ .0898), with the Anglophones performing better (86.7 %, $s.d. =$ .0981) than the Germans (80 %, $s.d. =$ .0799). Accuracy rates were submitted to a one-way ANOVA with L1 group (Anglophones vs. Germans) as fixed factor. As predicted, this difference was significant ($F_1(1,79) = 9.97$, $p < 0.01$; $F_2(1,94) = 4.89$, $p < 0.05$). Among the Anglophones, hit rates at 85.6 % were equal to correct rejections at 86.9 % (false alarm rate = 13.1 %). A d' statistic of -0.669 showed that they were relatively insensitive to the distinction. In contrast, the Germans had a lower hit rate of 79.69 % and a higher false alarm rate of 21 %. The d' statistic on the Germans' scores was 1.80, suggesting that the L2ers were sensitive to the distinction between targets and distractors and responded differently to the two types of block. Pair-wise correlations were computed on the responses of the English test and the component tests of the OEPT: correlations were high on the listening test (0.45) but even higher on the grammar test (0.75).

3.5.3. *Target blocks.* Only target blocks were relevant for the analysis of the stimulus properties. Table 5 presents accuracy rates organised by word status (uncommon vs. common), word length (number of syllables), sentence position (initial, medial, final), and block position (1st, 2nd, 3rd, and 4th sentence in the blocks).

Table 5. *Accuracy results on probe detection tasks*

Group	English		German
	Anglophones $n = 22$ (s.d.)	Germans $n = 60$ (s.d.)	Germans $n = 60$ (s.d.)
Totals	.86 (.08)	.80 (.07)	.88 (.07)
Targets			
Hits	.85 (.34)	.79 (.07)	.86 (.33)
Misses	0.14	0.2	0.14
Distractors			
Correct rejections	.86 (.33)	.79 (.40)	.90 (.29)
False alarms	0.13	0.21	0.1
Target blocks only			
Word Status			
Common	.85 (.35)	.81 (.38)	.85 (.35)
Uncommon	.87 (.33)	.77 (.41)	.88 (.31)
Word length			
Monosyllabic	.88 (.31)	.80 (.39)	.85 (.34)
Polysyllabic	.84 (.36)	.78 (.40)	.87 (33)
Position in sentence			
Initial	.88 (.32)	.77 (.41)	.82 (.40)
Medial	.89 (.31)	.80 (.39)	.86 (.34)
Final	.81 (.38)	.81 (.38)	.92 (.26)
Block position			
1st	.80 (.40)	.65 (.47)	.78 (.41)
2nd	.81 (.38)	.74 (.43)	.85 (.34)
3rd	.93 (.24)	.86 (.34)	.92 (.25)
4th	.92 (.26)	.97 (.14)	.95 (.21)

Separate ANOVAs were computed on the variables. For the Anglophones, none of the variables were significant, although block position approached significance ($F2(3,27) = 2.55$, 0.0764, n.s.). For the Germans, only block position was significant ($F2(3,27) = 4.01$, $p < 0.05$. In other words, the commonality of the target words, their word length and sentence position had no effect on accuracy rates in either the native speaker nor the L2 learner groups. Figure 1 shows the mean scores of both groups by sentence position. Figure 2 shows the mean scores of both groups by block position.

3.5.4. *Reaction time data.* Correct responses were analysed to see if they could provide independent evidence that the uncommon items were generally

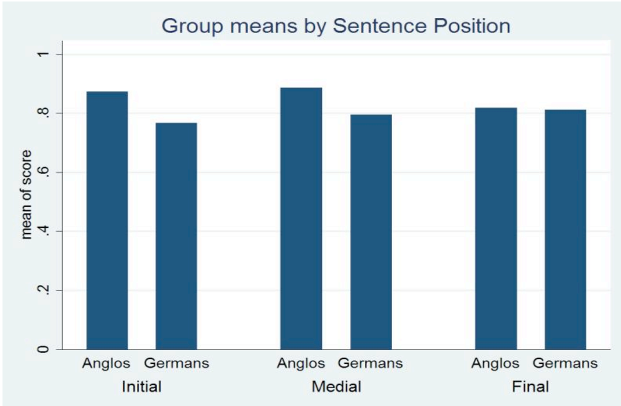

Figure 1. *English language group means by sentence position*

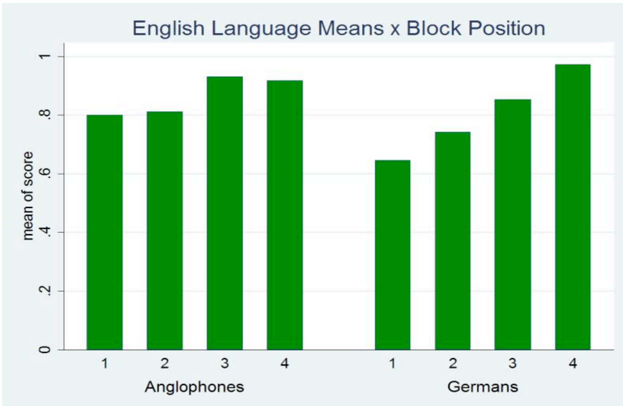

Figure 2. *English language group means by block position*

novel words. 706 incorrect responses were deleted from the data set, leaving 3178 responses from the original 3884 data points. These were log-transformed before being analysed. Table 6 presents the responses latencies, organised by language and commonness status.

Table 6 reveals that responses to common words were faster than responses to uncommon words. Two-tailed t-tests on the logRTs revealed that these differences were significant. In particular, participants in both groups responded much more slowly to uncommon distractors (mean = 1808 ms) than to common distractors (mean = 1423 ms). For the Anglophones, $t(458) = 10.6$, $p = 0.000$; for the Germans, $t(1135) = 17.7$, $p = 0.000$. While the more rapid re-

Table 6. *Response latencies in ms in English and German*

	English means (s.d.)		German means (s.d.)
Group	Anglophones	Germans	Germans
Targets	1410 (407) $n = 455$ Range: 533–2468	1482 (385) $n = 1129$ Range: 563–2496	1428 (393) $n = 1251$ Range: 508–2497
Common	1206 (361) $n = 225$ Range: 533–2347	1290 (333) $n = 580$ Range: 563–2455	1279 (386) $n = 614$ Range: 508–2497
Uncommon	1608 (347) $n = 230$ Range: 912–2468	1685 (328) $n = 549$ Range: 833–2496	1572 (344) $n = 637$, Range: 756–2493
Distractors	1548 (457) $n = 460$ Range: 519–2491	1632 (404) $n = 1137$ Range: 595–2956	1539 (422) $n = 1306$ Range: 668–2497
Common	1351 (445) $n = 238$ Range: 641–2454	1461 (394) $n = 593$ Range: 595–2501	1345 (394) $n = 682$ Range: 668–2497
Uncommon	1758 (368) $n = 222$ Range: 519–2491	1828 (314) $n = 544$ Range: 978–2956	1752 (343) $n = 624$ Range: 772–2485

sponses on the target items (Anglophones: $t(453) = 12.11$, $p = 0.000$; Germans: $t(1127) = 20.0$, $p = 0.000$) might be due to the relationship of the probe to its original stimulus sentence (priming), in the case of the distractors, these words were being presented for the first time. An obvious way to explain why the responses to the common distractors were significantly faster than the responses to the uncommon distractors is to hypothesise that participants knew some of the common words but did not know the uncommon words.[16] Reaction time data thus provide some independent evidence that the uncommon bird names were not familiar to our subjects.

16. I thank Dr. Robin Hörnig for pointing out the significance of this analysis.

3.6. *German language results: Comparison of L1 vs. L2 status*

Our L2 subjects performed the task in German (their L1). This permits us to compare their performance on the task in each language and to compare their L1 performance with that of the Anglophones.

3.6.1. *Data reduction and adjustment.* Mean accuracy rates were calculated, first by participant and then by items. The combined mean was 88.74 % ($n = 1104$; $s.d. = 0.104$, range: 66 % to 100 %). Tests of skew and kurtosis calculated on the responses by items revealed a normal distribution. Tests of skew and kurtosis calculated on the responses by subjects revealed skew ($p < 0.008$, $p > \chi^2 = 0.03$). Two subjects' scores were adjusted to within 2 standard deviations of the mean. The adjusted mean was 88.85 % ($s.d. = .0738$). Two-tailed t-tests performed on the two lists showed means not to be statistically different ($t = 1.20, df = 58$, $p = 0.23$, n.s.). Lists were collapsed for further analysis. As for the latency data, 138 responses were automatically timed out. These errors constituted 4.49 % of the total number of responses ($n = 2879$) and 42 % of the errors. They were removed from the data set. No responses were too short.

3.6.2. *Targets and distracters combined.* Subjects scored slightly higher on the distractor blocks at 90.62 % correct ($n = 1440, s.d. = .2915$) than on the targets (86.87 %, $n = 1440$, $s.d. = .3377$). A d' statistic was calculated on the hits and false alarms. At 3.05, it suggests that subjects were responding differently to the two types of items.

Recall that on the English task, German subjects scored 80.06 %. The difference in means of the Germans on the German and English tasks was significant ($F(1,141) = 84.75$, $p < 0.001$), with the subjects performing significantly better in their L1 than in their L2, as predicted.

3.6.3. *Target blocks.* Results for the different stimulus properties are shown in Table 5. Scores on the two word lengths were comparable: 85.7 % (monosyllables) vs. 87.8 % (polysyllables). Thus, word length had no effect on the scores. Scores by word status were also similar: 88.3 % (uncommon) 85.7 % (common). Scores increased from initial position, through medial position to final position: 82.2 %, 86.2 %, and 92.2 % respectively, meaning that medial position was not the most difficult position to process. Scores also increased through the block positions: 78.4 %, 85 %, 92.6 % and 95 %. Only block position was significant ($F(3,7) = 4.02$, $p < 0.05$).

Subjects responded faster to the common bird names in German at 1279 ms ($n = 614$, $s.d. = 386.14$, range 508–2497) than to the uncommon bird names at 1572 ms ($n = 637$, $s.d. = 344.90$, range 756–2493). See Table 6. This difference was significant ($F(1,1249) = 201.01$, $p < 0.001$). This again suggests that some common names may have been known whereas the uncommon names were novel names.

Finally, a comparison of the Anglophones in English and the Germans in German revealed that performance on the task was both high and very similar, both in the accuracy scores and in the response latencies.

3.7. *Summary*

The OEPT showed that our L2 subjects exhibited a range of proficiency levels from "elementary" to "very advanced", with higher scores on the listening test than on the grammar test. The Listening Test turned out to be particularly revealing for our purposes because many of its items are focally accented. Results on the probe task and on the OLT correlated well.

On the main probe task, German subjects performed very well both in their L1 and their L2. Anglophones performed very well in English. Separate ANOVAs showed that performance on the task was not affected by properties of the words. In particular, our data provide no evidence for either of the word length hypotheses (either 2a or 2b). This result confirms those of Rast (2008). We also found no support for our positional hypotheses. The one factor that did have an effect was block position.

4. Discussion and Conclusions

We conducted a probe detection test in order to test the hypotheses that the factors *position in sentence* and *word length* alter the salience of unfamiliar words in input processing. No effect of either hypothesis was found. Our study fails to confirm claims that words in medial position of a sentence are harder to process in the input than words at the ends of sentences because words at the ends of sentences exhibit primacy or recency effects. In contrast, we did see an effect of block position, namely the closer the probe was to the sentence containing it, the higher the score. This *is* a clear recency effect very likely attributable to the operation of short-term memory. The contrast between the effect within the block and the absence of an effect within the sentence raises obvious questions about VanPatten's (1996) attempts to attribute his sentence position effects to attention. In particular it raises questions about his assumptions about the limited-capacity of input processors:

> If a language-based L1 model [of speech- and language-processing] suggests that comprehension is capacity-robbing, we can only begin to imagine the drain on resources for the L2 learner attempting to comprehend in a new language and the impact this has on processing. (VanPatten 2002: 826).

The investigation of sentence position and input processing by Barwise and VanPatten (1997) is motivated by the same assumption. The Sentence Location Principle shown in (1) putatively arises because of capacity constraints on working memory that operate during sentence processing. Whatever the developmental story to tell about word-learning and sentence position,[17] our study does not support this particular interpretation of input processing. Recall, however, that sentence position and focal accent interact and that neither Barcroft and VanPatten (1997) nor Rast (2008) controlled for focal accent when presenting their subjects with sentences to repeat. We did. It is possible, therefore, that the effect they found is due to confounds between focal accent and sentence position. In other words, subjects might have repeated words better in sentence-initial and sentence-final position on a sentence repetition task because words in these positions were focally accented and focal accent is salient for L2 learners. This will need to be tested in future research. It is, however, implausible that focal accent will explain the differences among these studies. After all, the primary purpose of our project was to investigate if focally accented words are easier to segment and recognise than de-accented words. We failed to find any effect whatsoever of focal accent on our task. See note 8. Moreover, in a subsequent replication study using different stimuli and an improved methodology, we also failed to find an effect of focal accent on segmentation and word identification. It is possible that focal accent is perceived as prominent only *after* learners have acquired the relevant phonology. If this thinking is correct, listeners will perceive focally accented words as prominent when they have acquired the relevant knowledge of the target language intonation and its role in information structure.[18] On this story, the perception of prominence on words will be an *outcome* of language learning, not an input to it.

---

17. In experiments with English-speakers having no prior knowledge of German, we have been able to show that our subjects can segment and represent, that is to say "learn", unfamiliar names occurring in both medial and final positions of the sentence (Carroll et al. 2009). Our primary task was a receptive task (a forced choice task based on rather gross phonological differences between two German words, e.g. *Frank* versus *Frantz* but we also asked our subjects to repeat the names, which they could readily do. This research shows that even absolute beginners can process easily sounds in sentence-medial and sentence-final position.
18. Cutler's extensive work on focus processing shows that knowledgeable listeners use intonation contours *preceding* and *following* focal accent to predict where the focus will fall. This requires a dynamic integration of information from various sources. See Cutler (1988) and Cutler et al. (1997) on focus processing in natives, and Akker and Cutler (2003) for a study on focus processing by bilinguals.

If additional experimentation fails to find any effect of focal accent and word learning, we may conclude that positional effects found in previous work are not due to subjects' reacting to focal accent in sentence-initial position or in sentence-final position. Rather, we may attribute these effects to the use of a sentence repetition task. Sentence repetition tasks require subjects to re-encode what was segmented from the signal in order to pronounce it. This fact about speech production has been known from some time (Naiman 1974). In contrast, probe detection tasks do not requiring re-encoding of the input. If so, positional effects result from rehearsal strategies during speech production and are not a factor constraining input processing.

It also might be argued that once an L2 learner reaches a certain level of proficiency in the target language, such properties of the input become irrelevant to word learning. It could be argued that the perceptual salience of an input matters only at the very earliest stages of acquisition or for the lowest levels of proficiency. This argument is consistent with the fact that the native speakers (Anglophones in English, Germans in German) showed no effects of position or word length in their L1s. Some have suggested that our learners were "too advanced" and the task "too easy". Such comments presuppose that we already know that at some point in language proficiency, perceptual salience ceases to be relevant to segmentation and form learning. Since we do not know *that* this is so, and we do not know when this point occurs, we can only continue to investigate the issues further. Nevertheless, Rast's (2008) study addresses this question, at least as far as word length is concerned. She found no effects of word length in her subjects' results even *before* they had had instruction in Polish. Her results and the data here, taken together, suggest that word length (as measured by the contrast between monosyllabic and polysyllabic words) is simply not relevant for segmenting L2 words.

Such comments may also reflect an assumption that a "physicalist" story about input, input processing, perceptual salience, and word learning must be true, although there is very little real evidence for it. In other words, there is very little real evidence that shows that L2ers attend to properties of the speech stream that are inherently perceptually prominent. Indeed, there is very little evidence to suggest that there *are* properties of the signal that are inherently perceptually prominent. It is worth emphasising that what little discussion of segmentation and word learning in SLA there is adverts to stress, duration and pause as cues to segments (Hatch 1978, 1983; Barcroft and VanPatten 1997; Doughty 2003). These properties are acquired abstract properties of phonological systems (Lehiste 1970). There is no evidence yet that L2ers use such information (acquired or inherently processable) to segment words that have not yet been acquired.

In conclusion, if we want to make claims about how input is processed, it behoves us to adopt tasks that are not contaminated by constraints that may arise solely in speech production. Better yet, we should investigate what properties of the input learners at various proficiency levels actually find salient.[19] This perceptual data must then be linked to word learning. Only then can we determine empirically if perceptual salience is a causal element in L2 word learning or if it is, rather, a by-product of the acquisition of specific phonological systems emerging as a result of or concurrently to word learning.

University of Calgary
⟨susanne.carroll@ucalgary.ca⟩

Appendix A and B, see http://dx.doi.org/10.1515/iral-2012-0002_suppl_1 or email the author directly (susanne.carroll@ucalgary.ca).

References

Allen, Dave. 1992. *The Oxford English placement test.* Oxford: Oxford University Press.

Akker, Evelien, & Anne Cutler. 2003. Prosodic cues to semantic structure in native and non-native listening. *Bilingualism: Language & Cognition* 6(2). 81–96.

Altenberg, Evelyn. 2005. The perception of word boundaries in a second language. *Second Language Research* 21(1). 1–34.

Baddeley, Alan. 1986. *Working memory*. Oxford: Oxford University Press.

Baddeley, Alan, Susan Gathercole & Costanza Papagno. 1998. The phonological loop as a language learning device. *Psychological Review* 105(1). 158–73.

Barcroft, Joe & Bill VanPatten. 1997. Acoustic salience of grammatical forms: the effect of location, stress, and boundedness on Spanish L2 input processing. In William R. Glass & Ana Teresa Pérez-Leroux (eds.), *Contemporary perspectives on the acquisition of Spanish*, vol. 2: *Production, processing and comprehension*, 109–121. Somerville, MA: Cascadilla Press.

Bardovi-Harlig, Kathleen. 1987. Markedness and salience in second language acquisition. *Language Learning* 37(3). 385–407.

Bates, Elizabeth & Brian MacWhinney. 1989. Functionalism and the Competition Model. In Brian MacWhinney & Elizabeth Bates (eds.), *The cross-linguistic study of sentence processing*, 157–193. Cambridge: Cambridge University Press.

Birds of Britain. http://www.birdsofbritain.co.uk/index.htm. Accessed 2003 and 2004.

Brown, Roger W. 1957. Linguistic determinism and the part of speech. *Journal of Abnormal & Social Psychology* 55(1). 1–5.

Cammarota, Marie-Ange & Jorge Giacobbe. 1986. L'acquisition du lexique en français par des adultes hispanophones. *Langages* 84(1). 65–78.

Carroll, Susanne. 2001. *Input and evidence. The raw material of second language acquisition*. Amsterdam: Benjamins.

Carroll, Susanne, Susan Jackson & Danica MacDonald. 2009. Adults process phonology from first exposure. *Mind–Context Divide*. University of Iowa, April 30th–May 2nd 2009.

Curtin, Suzanne, Thomas H. Mintz & Morten H. Christiansen. 2005. Stress changes the representational landscape: Evidence from word segmentation. *Cognition* 96. 233–62.

19. We are currently conducting just such a study.

Cutler, Anne. 1976a. Beyond parsing and lexical look-up: An enriched description of auditory comprehension. In Roger J. Wales & Edward C. T. Walker (eds.), *New approaches to language mechanisms*, 133–149. Amsterdam: North Holland.

Cutler, Anne. 1976b. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics* 20(1). 55–60.

Cutler, Anne. 1988. Prosody in situations of communication: Salience and segmentation. *Proceedings of the 12th International Congress of Phonetic Sciences*, vol. 1, 264–70.

Cutler, Anne, Jacques Mehler, Dennis G. Norris & Juan Segui. 1983. A language-specific comprehension strategy. *Nature* 304: 159–160.

Cutler, Anne, Delphine Dahan & Wilma van Donselaar. 1997. Prosody in the comprehension of spoken language: A literature review. *Language & Speech* 40: 141–201.

DeKeyser, Robert W. 2001. Automaticity and automaticisation. In Peter Robinson (ed.), *Cognition and second language instruction*, 125–151. Cambridge: Cambridge University Press.

DeKeyser, Robert W., M. Rafael Salaberry, Peter J. Robinson & Michael Harrington. 2002. What gets processed in processing instruction: A response to Bill VanPatten's "Update". *Language Learning* 52: 805–23.

de Jong, Kenneth J. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97: 491–504.

Ede, Basil, Goetz Rheinwald & William D. Campbell. 1965. *Unsere schöne Vögelwelt*. Köln: Lingen Verlag.

Eisenberg, Peter. 2004. *Grundriß der deutschen Grammatik*, vol. 1, *Das Wort*, 3rd ed. Munich: J.B. Metzler.

Ellis, Nick C. 2003. Constructions, chunking and connectionism: The emergence of second language structure. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 63–103. Oxford: Basil Blackwells.

Ellis, Nick C. 2008. Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention and the limited L2 endstate. In Peter W. Robinson & Nick C. Ellis (eds.), *Handbook of cognitive linguistics and second language acquisition*, 372–405. London: Routledge.

Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure*. Cambridge: Cambridge University Press.

Erteschik-Shir, Nomi. 2006. *The syntax/discourse interface: Information structure*. Oxford: Oxford University Press.

Fisher, Cynthia, Henry Gleitman & Lila R. Gleitman. 1991. On the semantic content of subcategorisation frames. *Cognitive Psychology* 23. 331–392.

Frauenfelder, Ulrich & Aditi Lahiri. 1989. Understanding words and word recognition: Does phonology help? In William D. Marslen-Wilson (ed.), *Lexical representation and process*, 319–341. Cambridge, MA: MIT Press.

Galbraith, R., and B. Underwood. 1973. Perceived frequency of concrete and abstract words. *Memory & Cognition* 1(1). 56–60.

Goldschneider, Jennifer & Robert M. DeKeyser. 2001. Explaining the 'natural order of L2 acquisition' in English: A meta-analysis of multiple determinants. *Language Learning* 51(1). 1–50.

Hatch, Evelyn M. 1978. Discourse analysis and second language acquisition. In Evelyn M. Hatch (ed.), *Second language acquisition: A book of readings*, 401–474. Rowley, MA: Newbury House.

Hatch, Evelyn M. 1983. Simplified input and second language acquisition. In Roger W. Andersen (ed.), *Pidginization and creolization as language acquisition*, 64–88. Rowley, MA: Newbury House.

Hawkins, Roger. 2001. *Second language syntax: A generative introduction*. Oxford, U.K: Blackwells.

Henderson, A.I. & S. Nelms. 1980. Relative salience of intonation fall and pause as cues to the perceptual segmentation of speech in an unfamiliar language. *Journal of Psycholinguistic Research* 9. 147–159.

Hosking, David, Jim Flegg & Eric Hosking. 1993. *Jim Flegg's field guide to the birds of Britain and Europe.* Ithaca, N.Y: Cornell University Press.

Johnson, Keith. 2004. Massive reduction in conversational American English. In *Spontaneous speech: Data and analysis. Proceedings of the 1st Session of the International Symposium*, K. Yoneyama & K. Maekawa (eds.), 29–54. Tokyo, Japan: The National & International Institute for the Japanese Language.

Klein, Wolfgang. 1986. *Second language acquisition*. Cambridge: Cambridge University Press.

Köpcke, Klaus-Michael. 1988. Schemas in German plural formation. *Lingua* 74. 303–335.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Lahiri, Aditi & William D. Marslen-Wilson. 1991. The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition* 38. 245–294.

Larsen-Freeman, Diane. 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9. 409–430.

Lieber, Rochelle. 1980. *On the organisation of the lexicon*. Doctoral dissertation, MIT. Distributed by the Indiana University Linguistics Club, 1981. Published by Garland Press, 1990.

Lightfoot, David. 2006. *How new languages emerge*. Cambridge: Cambridge University Press.

MacWhinney, Brian. 2005. A unified model of language acquisition. In Judith F. Kroll & Annette M. B. de Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, 49–67. Oxford: Oxford University Press.

Multilingual Bird Search Engine. http://www.knutas.com/birdsearch. Accessed 2003 and 2004.

Naiman, Neal. 1974. The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism* 2(1). 1–37.

Norris, Dennis G., Anne Cutler, James M. McQueen & Sally Butterfield. 2006. Phonological and conceptual activation in speech comprehension. *Cognitive Psychology* 53(2). 146–93.

Odlin, Terence. 1989. *Language transfer: Cross-linguistic influence in language learning.* Cambridge: Cambridge University Press.

Papagno, Costanza, Tim Valentine & Alan Baddeley. 1991. Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory & Language* 30(3). 331–347.

Rast, Rebekah. 2008. *Foreign language input: Initial processing.* Clevedon: Multilingual Matters.

Ringbom, Håkan. 1987. *The role of the first language in foreign language learning*. **???**

Segalowitz, Sidney J., Norman S. Segalowitz & Anthony G. Wood. 1998. Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics* 19(1). 53–67.

Sharwood Smith, Mike. 1991. Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research* 7(2). 118–132.

Streefkerk, Barbetje. 2002. *Prominence. Acoustical and lexical/syntactic correlates.* Doctoral dissertation. LOT series 58. Utrecht: LOT/UCLC.

VanPatten, Bill. 1996. *Input processing and grammar instruction: Theory and research*. Norwood, N.J: Ablex.

VanPatten, Bill. 2004. Input processing in second language acquisition. In Bill VanPatten (ed.), *Processing instruction: Theory, research, and commentary.* Mahwah, N.J: Erlbaum, 5–31. Mahwah, N.J: Lawrence Erlbaum.

VanPatten, Bill. 2007. Input Processing in adult second language acquisition. In Bill VanPatten & Jessica Williams (eds.), *Theories in second language acquisition*, 115–135. Mahwah, N.J: Lawrence Erlbaum.

Williams, John N. & Peter Lovitt. 2003. Phonological memory and rule learning. *Language Learning* 53: 67–12.