Frequency Effects in Second Language Acquisition: An Annotated Survey

Galip Kartal¹, Ece Sarigul¹

¹Ahmet Kelesoglu Faculty of Education, Necmettin Erbakan University, Konya, Turkey

Correspondence: Galip Kartal, Ahmet Kelesoglu Faculty of Education, Necmettin Erbakan University, Konya, Turkey. E-mail: kartalgalip@gmail.com

Received: April 5, 2017   Accepted: April 20, 2017   Online Published: April 23, 2017
doi:10.11114/jets.v5i6.2327   URL: https://doi.org/10.11114/jets.v5i6.2327

Abstract

The aim of this study is to investigate the relationship between frequency and language acquisition from many perspectives including implicit and explicit instruction, frequency effects on morpheme acquisition in L2, the relationship between frequency and multi-word constructions, frequency effects on phonetics, vocabulary, gerund and infinitive constructions, bitransitive constructions, and so on. In his impressive article on frequency, Ellis (2002a) asked this crucial question about the relationship between frequency and SLA: “How exactly does the frequency of patterns in the input affect acquisition?” Another fundamental question that needs to be answered is: “What purpose does frequency serve in language acquisition?” In addition, frequency and range are said to be the most salient criteria for determining the general usefulness of a word (Koprowski, 2005). Moreover, both corpus and experimental data suggest that frequency has significant effects on second language acquisition. These questions and arguments, which constitute the foundation of frequency studies in second language acquisition, led and guided this study. This annotated survey, summarizing the studies on frequency effects on L2 learning, shows that there are still lots of ways to go with research on the relationship between frequency and second language acquisition in the literature. Some pedagogical implications are discussed with recommendations for the stakeholders of foreign language contexts (i.e. curriculum designers, coursebook writers, language teachers and learners).

Keywords: frequency, SLA, language learning, language teaching, annotated survey

1. Introduction

The effects of frequency on the acquisition of a second/foreign language have caught the attention of many researchers. Many research findings revealed that acquisition is closely related to frequency (e.g. Ellis, 2002a; Koprowski, 2005). In his impressive article on frequency, Ellis (2002a) asked this crucial question about the relationship between frequency and Second Language Acquisition (SLA): “How exactly does the frequency of patterns in the input affect acquisition?” (p. 165). Another fundamental question that needs to be answered is: “What purpose does frequency serve in language acquisition?” These questions, which constitute the foundation of frequency studies in the second language (L2) acquisition, led and guided this annotated survey.

This survey on frequency effects in second language education was prepared to investigate the relationship between frequency and language acquisition from many different perspectives. First, a brief introduction to the topic is provided by focusing on previous research findings. Second, the studies are grouped under general research aims and discussed in terms of where gaps occur and what future studies should investigate. Lastly, some pedagogical implications are discussed with some recommendations.

Second language acquisition studies have been affected by many other disciplines. Linguistics and Psychology are the two of these disciplines that have significant contributions to the field. Gass and Mackey (2002) have pointed out that “SLA has seen an increased emphasis on psycholinguistic issues.” They also discuss a connection between Swain’s (1995, 1998) claims about output and second language learning from frequency-based accounts. According to them, “it is possible that output functions as a priming mechanism for future input.” This is also related to probabilistic processing, a process given by Ellis (2002a) in order to account for the cognitive processes that are used by second language learners. Probabilistic processing refers to “the remarkable sensitivity that learners have regarding the relative frequency with which certain forms are used in particular contexts in the input they receive, and their ability to match
their output according to what they think might be appropriate” (Littlemore, 2009).

How morphemes are acquired by second language (L2) learners is a debatable issue among SLA researchers. First, Larsen-Freeman (1975) drew attention to frequency as a critical factor determining the acquisition of morphemes. Before this discovery, it was the dominant view that inner factors played the key role in the L2 acquisition process. So, what does frequency refer to? Frequency, more specifically frequency in the input, “refers to the relative frequency of formal features in the language that people hear or read (Vanpatten & Benati, 2010). According to Ellis (2002a, 2002b, 2008), Larsen-Freeman (1997), MacWhinney (1999), and (Schwartz & Causarano, 2007), frequency has a major role in second language acquisition. Ellis (2002a,2002b, 2008), for instance, believes that second language learning is input driven. Memory functions need frequency for storage of information. R. Ellis (1994) approaches the issue of frequency effects with a more watchful viewpoint:

Overall there is very little evidence to support the claim that input frequency affects L2 acquisition but there is also very little evidence to refute it. Perhaps the safest conclusion is that input frequency serves as one of the factors influencing development, often in association with other factors such as L1 transfer and communicative need (p. 273).

Ellis (2002a, 2002b) argues that there is a close relationship between language acquisition and frequency. He bases this relationship on the following assumption: “Humans are sensitive to the frequencies of events in their experience” (2002a, p. 145). There is a positive relationship between frequency and usefulness of words as well (Koprowski, 2005). Of course, frequency is not the only predictor of the usefulness of a word, but it is an important one. In order to determine the usefulness of a word, scholars have developed specific criteria, which include range, frequency, availability, opportunism, learnability, and coverage (White, 1988). However, frequency and range are regarded as the most salient criteria for determining the general usefulness of a word (Koprowski, 2005). In his analysis of multi-word constructions, Koprowski (2005) analyzed the usefulness of lexical phrases (e.g. phrasal verbs, idioms, collocations, binomials, and compound nouns) used in three contemporary coursebooks published in England. Frequency and range turned out to be the most significant criteria for indicating the ‘usefulness’ of a lexical phrase.

With a special link to implicit and explicit language acquisition, Ellis (2002a) discussed the effects of frequency with regard to many areas in his review of frequency effects in language learning. In other words, he emphasized that input frequency has effects on spelling, reading, phonology, phonotactics, formulaic language, lexis, morphosyntax, grammaticality, language comprehension, and syntax (sentence production). For example, Ellis (2002a) discussed the experimental studies on implicit knowledge on phonology and phonotactics and reached the assumption that: “higher frequency words get more activation from the same evidence than do low-frequency words” (p. 151).

According to Ellis and Collins (2009), some of the determinants of construction learning are input frequency, form, function, and. Redundancy. The terms ‘Frequency’ and ‘input frequency’ are used exchangeably. There are two types of frequency: token and type. The former one “is how often particular words or specific phrases appear in the input” while the latter one “is how many different lexical items can be applied to a certain pattern, paradigm, or construction” (Ellis, 2002a). Larsen-Freeman (2002) explains the connection between connectionist models and frequency with these words: “In these models of neural networks, items that are frequent in the input increase the connection weights between nodes. As the input is processed, certain connection weights are strengthened, while others atrophy, simulating the plasticity of synaptic connections.”

All these studies show that second language acquisition cannot be isolated from the frequency in the input. This is why corpus-based vocabulary teaching research has gained a momentum in the past decades. Gardner and Davies (2007), for instance, pointed out the 100 most frequent phrasal verbs in the British National Corpus. Trebits (2009), similarly, handled the phrasal verbs from English for specific perspectives and identified the most frequent phrasal verb combinations in European Union documents. Most of the scholars and authors who study language give a lot of importance to phrasal verbs in their studies since they are inseparable parts of a language

2. Frequency Effects in Second Language Acquisition

2.1 Implicit and Explicit Instruction

The role of frequency in language acquisition can also be discussed in regard to implicit and explicit instruction. According to Ellis (2002b), explicit memory is about the situations in which people remember prior events with a conscious process. In implicit memory, on the other hand, there is no conscious recall of the previous event, and remembering occurs with a related stimulus. Both Ellis (2002a, 2002b), and Larsen-Freeman (2002) find explicit instruction valuable as explicit instruction provides helpful negative evidence. Gass and Selinker, (2008) define negative evidence as the information to a learner that informs that the use is not appropriate to the norms of the language they are learning. When implicit and explicit instructions are discussed with regard to frequency effects Larsen-Freeman’s (2002) suggestions should be taken into account. She believes that the cognition of humans should
be explained by not ignoring the fact that we must measure the frequency with both focusing on form and function. Generative capacity requires considering human agency and attention as well. Teachers handle learners’ sensitivity to frequency issue with specific respect to taking advantage of this sensitivity by creating meaningful activities.

2.2 Frequency Effects on Morpheme Acquisition in L2

The previous studies on the relationship between frequency and morpheme acquisition in L2 (Larsen-Freeman, 1975; Ellis & Schmidt, 1997; Goldschneider & DeKeyser, 2001) found evidence that there is a significant relationship between frequency and morpheme acquisition order. More specifically, input frequency has a crucial role in second learners’ morpheme acquisition order.

Larsen-Freeman (1975) investigated the effects of frequency on the order of acquisition of morphemes. The subjects of the study were 24 elementary level Arabic, Japanese, Persian and Spanish adult ESL learners. There were six participants from each background. Reading, writing, listening, imitating and speaking tasks were administered to the participants. The results of the study yielded that the participants’ performance on English morpheme acquisition followed the first language morpheme acquisition order. The following table represents the first language morpheme acquisition order reported by Brown (1973). It is concluded that frequency is a major determinant of the order of acquisition of morphemes for adult ESL learners. Dulay and Burt (1973) and Bailey, Madden, and Krashen (1974) also claim that there exists an acquisition order of English morphemes no matter what their and language backgrounds and ages are.

Table 1. Grammatical Morpheme Acquisition Order in Brown (1973)

<table>
<thead>
<tr>
<th>Order</th>
<th>Morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Progressive -ing</td>
</tr>
<tr>
<td>2</td>
<td>Plural -s</td>
</tr>
<tr>
<td>3</td>
<td>Past irregular</td>
</tr>
<tr>
<td>4</td>
<td>Possessive’s</td>
</tr>
<tr>
<td>5</td>
<td>Uncontractible copula</td>
</tr>
<tr>
<td>6</td>
<td>Articles (a/an, the)</td>
</tr>
<tr>
<td>7</td>
<td>Past regular-ed</td>
</tr>
<tr>
<td>8</td>
<td>Third-person singular -s</td>
</tr>
<tr>
<td>9</td>
<td>Third-person irregular</td>
</tr>
<tr>
<td>10</td>
<td>Uncontractible auxiliary</td>
</tr>
<tr>
<td>11</td>
<td>Contractible copula</td>
</tr>
<tr>
<td>12</td>
<td>Contractible auxiliary</td>
</tr>
</tbody>
</table>

Ellis and Schmidt (1997) found evidence to support Larsen-Freeman’s suggestions about the role of frequency in L2 morpheme acquisition order. They found a statistically significant effect of frequency on stem learning. In a more recent study, Goldschneider and DeKeyser (2001) investigated how input frequency, phonological salience, semantic complexity, morphophonological regularity, and syntactic category account for morpheme order in L2. The results showed that 71% of the variation in accuracy scores for grammatical morphemes is explained by these five variables altogether. Above and beyond the other four variables, input frequency itself explained 29% of the variation in the accuracy scores for grammatical morphemes. Also, except for morphophonological regularity, input frequency was found to be statistically and significantly related to the other four variables. More specifically, it is found that there is a significant positive correlation between the input frequency and phonological salience. On the other hand, there was a significant negative relationship between the input frequency and semantic complexity, and input frequency and syntactic category. All these findings led Goldschneider and DeKeyser (2001) to conclude that input frequency, which is one of the causal variables that explain L2 English morpheme acquisition order, has positive effects on morphological learning in a second language.

After Ellis’s salient emphasis about the importance of frequency in SLA, numerous empirical studies have been conducted to investigate the role of frequency in the field. These studies mainly focused on multi-word constructions (Arnon & Snider, 2010; Durrant & Doherty, 2010; Gardner & Davies, 2007), vocabulary (Rott, 1999), morpheme order studies (Goldschneider & DeKeyser, 2001), gerund and infinitive constructions (Schwartz & Causarano, 2007), bitransitive constructions (Year & Gondon, 2009), and phonetics-syllable contraction (Myers & Li, 2009). The details of these studies are given below.

2.3 Multi-word Constructions and Frequency

There are numerous studies that investigated the impact of frequency on the multi-word phrases. One of them, conducted by Arnon and Snider (2010), tested the sensitivity of adult learners to frequency effects in L2 via two experiments. A meta-analysis was conducted to predict processing latencies and to compare its effectiveness with a binary measure. 26 native English participants took part in the first experiment. A twenty-eight-item test was administered to the participants and they were asked to work in a computer laboratory. The twenty-eight item pairs,
fifty-six experimental items, were answered by all participants. The results of Experiment 1 indicated that participants were faster to respond to high frequent phrases than less frequent phrases. It is also found that, there is an effect of phrase-frequency on recognition times for phrases of varying frequency. In Experiment 2, the number of the participants were 23 and 17 target items. This time, the cutoff point distinguishing high and less frequent items was set to five million words. The procedure of experiment two is similar to the one in Experiment 1. It was also found in Experiment 2 that participants were faster to respond to higher frequent phrases than the less frequent ones. Participants sat in a computer laboratory and completed a phrasal decision task. Each participant saw twenty-eight item pairs. The outcomes of the study revealed that participants are sensitive to phrase frequency. In other words, participants responded to high frequent phrases quicker than low frequent phrases. In Arnon and Snider’s study, all the phrases were constituents ones. What is more, the focus was only on content words. Function words were excluded.

Durrant and Doherty (2010) designed another study in order to investigate the collocations, one of the problematic multi-word construction types for EFL learners. They tried to determine whether the frequency of co-occurrence in a large general-purpose corpus is a reliable indicator of psychological priming between words. The study included two experiments. The aim of the first experiment was to determine whether there was a priming between high-frequency collocations. 32 native British English speakers took part in the study. Two groups were constituted. Participants received ten practice trials and were allowed a self-timed break before commencing the experiment. There were four groups of pairings: collocations (moderate, frequent, and associated frequent) and low-frequency combinations. The materials of the first experiment were 16 collocating prime-target pairs. The participants of the second experiment were 38 native speakers. The only difference in the procedure is that a “pattern mask” before the prime was shown. According to the results of both Experiment 1 and 2, significant facilitation was found for associated frequent collocations in the low frequency or moderate collocation conditions. This study investigated the collocation priming in a well-organized way. The study provided remarkably detailed and comprehensive appendices. The frequency measures of the study were taken from the British National Corpus. Its size, with over 90 million words, as well as its representatives and contemporaneity, made the results of the BNC study quite reliable.

2.4 Frequency Effects on Phonetics, Vocabulary, Gerund and Infinitive Constructions, and Bitransitive Constructions

The effects of frequency on phonetics, vocabulary, gerund and infinitive constructions and bitransitive constructions were also investigated by the researchers. Rott (1999), for instance, investigated whether intermediate language learners’ incidental vocabulary acquisition and retention are affected by exposure frequency while reading the texts. Another aim of the study was determining the effectiveness of vocabulary growth through reading on receptive and productive vocabulary learning. The former one (receptive) was found to be more affected. As for the frequency, it is found that six encounters resulted in significantly more receptive as well as productive word knowledge than the two and four encounters.

Schwartz and Causarano (2007) designed a study in order to investigate whether the frequency of constructions (gerund and infinitives) has any effects on students’ target language input and output. It was hypothesized that “high-frequency constructions provide more exemplars for L2 learners to make generalizations than low-frequency constructions and that this directly relates to the number and kind of L2 learner errors” (p. 43). The participants of the study were 13 native Spanish speaking students studying at a university in the USA. The British National Corpus was used to determine the frequency of gerunds and infinitives. Type frequency was used for the frequency count. For example, all possibilities of the verb ‘want’: “wants, wanted, and wanting” were calculated. The findings of the study yielded a significant difference with regard to frequency use of gerunds and infinitives. So, the hypothesis was confirmed.

Pinker (1999) argued that coding of high-frequent regular inflected words into the long-term memory as whole units would be easier due to their frequency. Lehtonen et al. (2006) investigated the effects of word frequency on the processing of monomorphemic versus inflected words. In other words, morphological processing in Swedish was investigated from a frequency perspective. There were two experiments in the study. 17 university students took part in the first experiment. Swedish was their only language before they started the school. So, they were all monolinguals. In the procedure, students were shown some letter strings on the computer screen and asked to decide whether the letter strings are real Swedish words or not. The results yielded a difference with regard to the processing of the words. More specifically students processed the higher frequency items faster than, the lower frequency ones. By utilizing the same materials and the experimental procedure, the second experiment was conducted with 18 bilingual undergraduate students. The results were similar to the results of Experiment 1. To sum up, it was found that, for both monolinguals and bilinguals, frequency affected the subjects’ processing way of inflected. This study was put here because Swedish and English are similar with regard to morphological limitedness.

The effect of frequency on phonetics with a focus on syllable contraction was investigated by Myers and Li (2009). The research had two phases. In the first one, frequency estimates were investigated. Forty native speakers of Taiwanese
took part in the study. The results showed that corpus frequency itself explained 21% variance in the uncontracted frequency. The statistically significant difference between contracted and uncontracted frequency suggests that phonetic details were taken into consideration by listeners while giving their frequency estimates. For the production task, the researchers worked with 20 university students in Taiwan. The material was 120 auditory recordings of disyllabic uncontracted forms. It is found that the effectiveness of frequency on the degree of syllable contraction is not stable. It was affected by many other factors and the unique contribution in explaining the degree of syllable contraction was ‘gradient.’

Year and Gordon (2009) designed experimental research in order to investigate the effects of frequency on the ditransitive construction. In this construction, examples are loaded in favor of the given prototype. The participants were a first-year class of a middle school in Korea. The researchers constructed frequency groups and provided participants with 40 ditransitive and 40 prepositional constructions. In order to provide participants with concrete visual images for the meaning of each sentence, short video clips were utilized. The results of this well-designed study yielded a limited effect of frequency on Korean learners’ acquisition of ditransitive constructions. The researchers attributed this limited effectiveness to the “nature of construction, implicit/explicit learning mode adopted during learning, the number of stimuli taught, exposure duration, the order in which stimuli are presented, and construction complexity” (p. 412). In addition to experimental researches on the effectiveness of frequency in second language acquisition, there are some studies which were designed to detect the frequencies of some particular multi-word constructions (Gardner & Davies, 2007). The study identified and analyzed phrasal verb frequencies in the British National Corpus.

Table 2. A summary of the studies on frequency effects on L2 learning.

<table>
<thead>
<tr>
<th>Author &amp; Year</th>
<th>What is Investigated?</th>
<th>What is Found?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larsen-Freeman, 1975</td>
<td>The sequence of acquisition of grammatical morphemes in L2.</td>
<td>Frequency has effects on morpheme acquisition order.</td>
</tr>
<tr>
<td>Ellis &amp; Schmidt, 1997</td>
<td>Relationship between frequency and morphological abilities.</td>
<td>Frequency has effects on morpheme order.</td>
</tr>
<tr>
<td>Rott, 1999</td>
<td>Effects of exposure frequency on vocabulary knowledge in reading.</td>
<td>More exposure frequency (6 times exposure) resulted in the more vocabulary knowledge than the fewer ones.</td>
</tr>
<tr>
<td>Goldschneider &amp; DeKeyser, 2001</td>
<td>The frequency of lexical phrases in coursebooks.</td>
<td>Frequency has effects on morpheme order.</td>
</tr>
<tr>
<td>Koprowski, 2005</td>
<td>The relationship between students’ TL input and output in terms of frequency effects.</td>
<td>Coursebooks with more lexical phrases were considered as more useful for L2 learners.</td>
</tr>
<tr>
<td>Schwartz &amp; Causarano, 2007</td>
<td>The effects of frequency on syllable contraction.</td>
<td>Frequency played a crucial role in SLA.</td>
</tr>
<tr>
<td>Myers &amp; Li, 2009</td>
<td>The effects of frequency on the acquisition of the English ditransitive constructions.</td>
<td>A gradient relationship was detected.</td>
</tr>
<tr>
<td>Year &amp; Gordon, 2009</td>
<td>Adults’ sensitivity to frequency.</td>
<td>A limited effect was found.</td>
</tr>
<tr>
<td>Arnon &amp; Snider, 2010</td>
<td>Adults respond to high frequent phrases quicker than low frequent phrases.</td>
<td>Adults respond to high frequent phrases quicker than low frequent phrases.</td>
</tr>
<tr>
<td>Durrant &amp; Doherty, 2010</td>
<td>Relationship between frequency and psychological priming between words.</td>
<td>Frequency is an indicator of psychological reality.</td>
</tr>
<tr>
<td>Diependaele, Lemböfer, &amp; Brysbaert (2013)</td>
<td>The differences in the word frequency effect between L1 and L2 speakers.</td>
<td>Language proficiency increases lexical entrenchment and thus reduces frequency effect.</td>
</tr>
</tbody>
</table>

3. Discussion and Conclusion

Although a few studies found that frequency has a limited effect on the acquisition of some certain patterns (Year & Gordon, 2009), both corpus and experimental data suggest that frequency has significant effects on second language acquisition. This annotated survey shows that despite numerous studies, there are still lots of ways to go with research on the relationship between frequency and second language acquisition. One important limitation of frequency related studies is that they have difficulties in excluding some other factors in order to detect the unique contribution of frequency. Year and Gordon (2009) listed some of these factors as “the nature of a construction, implicit/explicit learning mode adopted during learning, the number of stimuli taught, exposure duration, the order in which stimuli are presented, and construction complexity.” Furthermore, word frequency distributions are highly skewed by nature, which makes even the mega corpora insufficient. Still, as Baayen (2001) argues, a bigger corpus is always better than a smaller one. The number of the studies conducted with mega corpora is rare in the area of frequency effects. Future studies may focus on this shortness. Furthermore, the analysis of frequency studies shows that the frequency effects in language learning are not satisfactorily investigated in the Turkish context. So, more studies should also be conducted with Turkish EFL learners. Although, the studies on the frequency effects are multidimensional and complicated, there are some implications for the stake holders of English Language Teaching (ELT) contexts. For instance, the curriculum designers and ELT coursebook writers should consider utilizing corpus findings to make the language teaching easier.
and effective for learners. Similarly, language teachers should be aware of the fact that frequency has an important role in language acquisition and thus they should make use of more frequent constructions in language classrooms. Last but not least, for learners to function well, they should keep in mind that some words have more frequent use than the others. Diependaele et al. (2013) found a negative relationship between language proficiency and frequency effects. In other words, less proficient learners were found to be more sensitive to frequency effects. So, utilizing frequency effects with less proficient learners might be more useful. One way of utilizing frequency is using corpus tools.

As exposure to the authentic linguistic items in non-native speakers’ communities is very limited outside of the classroom, use of corpora has gained more significance recently. Johns (1994) states that authentic items are more helpful to teach the target language than the fabricated non-authentic ones. Similarly, Wiebe (2000) argues that using corpora for obtaining linguistic knowledge is popular in natural language processing. At this point, a crucial question comes into mind: How can language teachers help EFL learners to reach authentic usage of linguistic items both inside and outside of the classroom? Recent research findings came up with a solution to this problem regarding the concordancing programs with the help of technological advancements in the electronic compilation of authentic texts, namely the corpora. Particularly, the combination of the corpora and the concordance has encouraged language teachers and learners to reach natural occurring authentic language.

There are numerous corpora that are available to language learners. Two of them are considered as mega corpora as they include millions of words. One of them is the Corpus of Contemporary American English (COCA). It is a contemporary and genre-based corpus. It is free to access and includes over 450 million words, which provides comprehensive and representative data. There are five main sub-corpora of COCA: spoken, fiction, magazine, newspaper, and academic. The academic sub-corpora has about 83 million words and the data comes from 148 journals. The divisions in the academic subcorpus of COCA are history, education, geography/social science, law/political science, humanities, philosophy/religion, science/technology, medicine, and miscellaneous. The other corpus is the British National Corpus (BNC). It is also a large-scale corpus with over 100 million words. It is free to access these corpora, and they have a very comprehensive and representative data set. Thus, both of the corpora can be used appropriately by teachers of English for educational aims especially for vocabulary teaching in foreign language classes. Furthermore, it will be a good idea giving some clues to the students about the usage of them and teaching some strategies for learning vocabulary.

Note: The former version of this paper was presented at the International Conference on the Changing World and Social Research I, held in Vienna-Austria on August 25-28, 2015.

References


Copyrights
Copyright for this article is retained by the author(s), with first publication rights granted to the journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.