



Quantifiable Aspects of Language Acquisition

Grigory Kazakov

MGIMO University
76, Prospekt Vernadskogo, Moscow, 119454, Russia

Abstract. This paper investigates the quantifiable aspects of language acquisition such as the amount of input (perceived speech material) and vocabulary size. In light of the disillusionment of language education methodology in the 20th century with the search for an ideal method of language learning, the need for a new theoretical framework that will account for the available research data is emphasized. Stephen Krashen's theory of second language acquisition is described as the most significant contribution to this end. Krashen's hypotheses of acquisition and learning, natural order, monitor, comprehensible input and emotional filter are discussed. It is theorized that statistical processing of language input by the brain lies at the core of language acquisition. Reference is made to evidence from first language acquisition and polyglots' study of foreign languages confirming this statement. The requirements for optimal input are examined, leading to the question of sufficient input quantity (measured by the number of words). Relevant data from first language acquisition, corpus linguistics and polyglottery is presented. Then, the question of vocabulary size corresponding to different levels of language competence is addressed by reference to estimates found in polyglot literature, exam-based research and online sources. In conclusion, it is argued that reaching an intermediate level in foreign language skills (B1–B2 in the European framework), requires the input of about one million words while mother tongue acquisition in childhood requires several dozen times more. This result can serve as a reference point for language education curriculum and materials development. On this basis, some practical recommendations for the improvement of language education are proposed, namely the need for talking with children more, calculating input in the learning process, using audio and video materials more widely, stimulating extensive reading, encouraging students to perceive information on topics of interest in a foreign language, and working with their motivation. Finally, the importance of further quantitative studies on language acquisition is underscored.

Keywords: quantitative aspects of language acquisition, second language acquisition, input, vocabulary size, child speech, corpus linguistics, polyglottery

For citation: Kazakov, G.A. (2025). Quantifiable Aspects of Language Acquisition. *Linguistics & Polyglot Studies*, 11(3), pp. 10–23. <https://doi.org/10.24833/2410-2423-2025-3-44-10-23>

Измеряемые аспекты овладения языком

Г.А. Казаков

Московский государственный институт международных отношений (университет), МИД России
119454, Россия, Москва, пр. Вернадского, 76

Аннотация. В данной статье исследуются количественно измеряемые аспекты овладения языком, такие как объём инпута (воспринимаемого речевого материала) и размер словарно-го запаса. В свете того, что методика языкового образования в конце XX века разочаровалась в поиске идеального метода изучения языка, подчёркивается необходимость новой теоретической основы, которая будет учитывать имеющиеся исследовательские данные. Теория усвоения неродного языка С. Крашена описывается как наиболее значительный вклад в достижение этой цели. Обсуждаются гипотезы Крашена об усвоении и изучении языка, естественном порядке усвоения языковых структур, редакторе, понятном инпуте и эмоциональном фильтре. Выдвигается теория, что в основе усвоения языка лежит статистическая обработка инпута мозгом. Приводятся ссылки на данные о процессе усвоения родного языка и изучении иностранных языков полиглотирами, подтверждающие этот тезис. Рассматриваются требования к оптимальному инпуту, что приводит к вопросу о достаточном объёме инпута (измеряемого количеством слов). Анализируются релевантные данные из области усвоения родного языка, корпусной лингвистики и полиглотии. Затем изучается вопрос о размере словарного запаса, соответствующего разным уровням языковой компетенции, с учётом оценок, опубликованных в полиглотической литературе, исследованиях на основе экзаменов и интернет-источниках. В заключение утверждается, что для достижения среднего уровня владения иностранным языком (B1–B2 по европейской системе) требуется инпут в количестве около 1 млн. слов, тогда как для усвоения родного языка в детстве необходимо в несколько десятков раз больше. Этот результат может служить ориентиром для разработки учебных программ и материалов по иностранным языкам. На этой основе предлагаются некоторые практические рекомендации для усовершенствования языкового образования, а именно необходимость большего количества речевого общения с детьми, учёта инпута в учебном процессе, более широкого использования аудио- и видеоматериалов, стимулирования экстенсивного чтения, поощрения работы студентов с интересующим их тематическим материалом и внимания к их мотивации. Наконец, подчёркивается важность проведения дальнейших количественных исследований в области овладения языком.

Ключевые слова: количественные аспекты овладения языком, теория усвоения неродного языка, инпут, словарный запас, детская речь, корпусная лингвистика, полиглотия

Для цитирования: Казаков, Г.А. (2025). Измеряемые аспекты овладения языком. *Филологические науки в МГИМО*. 11(3), С. 10–23. <https://doi.org/10.24833/2410-2423-2025-3-44-10-23>

1. Введение

С конца XIX века методическая наука вращалась вокруг поисков идеального метода обучения иностранным языкам. Однако в конце XX века методический поиск разочаровался в понятии метода, то есть совокупности определённых принципов, которые вместе должны дать оптимальный результат обучения иностранным языкам. Идеальный метод не был найден, поскольку каждый метод имеет свои ограничения и применим для той или иной категории людей, для тех или иных целей изучения иностранного языка. Соответственно, стал вопрос о поиске новых понятий, которые позволили бы дать ориентиры в сфере изучения и преподавания иностранных языков [24, р. 14–16].

Если нет только одного определённого пути, тогда как подходить к обучению иностранным языкам? Что может служить его опорой? В свете данных о фундаментальных факторах процесса овладения языком, нейробиологии речи и природе человеческого языка, перед нами стоит задача создать такую модель, выстроить такую систему понятий, которая бы учитывала эти объективные данные, а не противоречила им, и позволяла предсказывать степень продуктивности того или иного подхода к обучению иностранным языкам.

2. Теория усвоения неродного языка С. Крашена

Наиболее значительный шаг в направлении поиска такой системы понятий, которая позволила бы в целом описать причины успеха изучения иностранного языка, как представляется, был предпринят американским лингвистом и методистом С. Крашеном. Его теория была сформулирована в 1980-е годы и не потеряла своей актуальности и сегодня. Из доступных автору данных она видится наиболее удачным описанием тех факторов, от которых зависит успешность изучения и преподавания иностранных языков. Название сферы её применения – теория усвоения неродного, или второго, языка (*second language acquisition*) – это наименование той дисциплины, которая более или менее соответствует тому, что в русскоязычной традиции называется методикой (возможно, с более прочной опорой на психолингвистические исследования). Теория Крашена формулируется в пяти гипотезах [20, р. 9–33].

1. Первая гипотеза заключается в различении усвоения и изучения языка (*acquisition and learning*). Крашен пишет о том, что эти два процесса не идентичны, что есть два источника, из которых мы черпаем знание языка. Один источник – это «усвоение» языка (*acquisition*). Это непровольный и не контролируемый сознанием процесс. Ученик не концентрируется на изучении, но произвольно усваивает структуру языка, понимает смысл новых слов и т.д. В то же время «изучение» языка (*learning*) – сознательный процесс, когда целенаправленно изучаются правила языка: грамматика, лексика и т.д. Крашен говорит о том, что это два отдельных процесса, которые играют разную роль в овладении новым языком.

2. Вторая гипотеза называется гипотезой о естественном порядке, согласно которой язык усваивается не в какой-то непредсказуемой последовательности, то есть каждый раз как получится, но в определённом естественном порядке формирования его структур. В своей книге Крашен на основе ряда исследований приводит примеры из английского языка, которые показывают, какие формы усваиваются раньше или позже. Например, среди первых, которые усваиваются быстрее всего, суффикс *-ing* (длительная форма глагола), суффикс множественного числа *-s*, глагол-связка *to be* и затем другие формы. Таким образом, утверждается, что в усвоении языка всегда прослеживается определённая закономерность.

3. Третья гипотеза называется по-английски *The Monitor Hypothesis*, гипотеза о редакторе. Эта гипотеза проливает свет на роль изучения языка, то есть сознательного достижения языковых знаний в результате систематических занятий и запоминания. Функция редактора заключается в контроле применения наличествующих знаний. Когда мы хотим сказать что-то, наша

способность порождать речь спонтанно не зависит от сознательного изучения («редактора»). Роль этих сознательно выученных языковых фактов лишь в том, чтобы редактировать высказывания, которые мы способны порождать спонтанно, но источником их порождения является только непроизвольное усвоение (*acquisition*). Иными словами, благодаря тому, что мы усваиваем непроизвольно, мы способны порождать речь, не задумываясь над формами и грамматическими нюансами. Это неосознанный процесс. А при целенаправленном использовании знаний включается редактор, который реагирует на готовое высказывание и анализирует его с точки зрения грамматической, лексической, стилистической и т.д. правильности, то есть мы оцениваем то, что уже непроизвольно родилось в нашей голове. В этом заключается различие между процессами усвоения и изучения языка.

Каким же образом в таком случае усваивается язык и откуда появляется способность порождать речь, то есть именно то, что называют свободным владением языком, когда мы не задумываемся над каждым суффиксом и окончанием, а просто свободно говорим и высказываем собственные мысли, суждения и т.д.? На эти вопросы отвечают гипотеза о понятном инпуте и гипотеза об эмоциональном фильтре.

4. Главная из них – это гипотеза о понятном инпуте: она представляет собой центр всей теории. Согласно данной гипотезе, источником приращения знания языка в процессе непроизвольного усвоения является только понимание доступного языкового материала. Когда мы встречаемся с образцами речи в письменной или устной форме, то есть слышим речь других людей или аудиозапись, смотрим видеofilm, читаем тексты, наш мозг получает инпут (*input*) – воспринимаемые нами устные и письменные тексты. И только когда мы понимаем этот языковой материал, происходит приращение знания языка. Крашен описывает это формулой $i + 1$, где i – это имеющийся языковой уровень, наше знание, а $+1$ означает материал немного выше текущего уровня знаний, прибавление нового к уже имеющемуся багажу.

5. Эмоциональный фильтр (*affective filter*), в свою очередь, – это психологический барьер, своего рода заслонка на пути инпута. Психологическая неготовность или закрытость, вызванная негативными эмоциями, страхом, тревожностью и т.д., блокирует восприятие языкового материала. Зрительно это можно изобразить следующим образом (см. рис. 1 и 2).

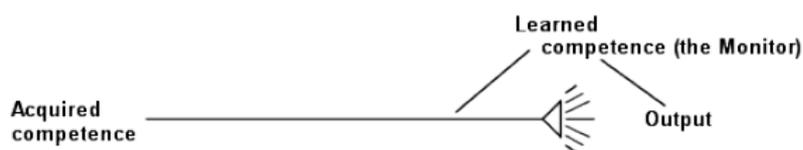


Рис. 1. Роль «редактора» [20, p. 16]

Мы, благодаря уже ранее усвоенному языковому материалу, приобретаем способность к производству речи (*acquired competence*) и порождаем некое высказывание (*output*), которое впоследствии контролируем на предмет его правильности: с точки зрения использования падежей, расстановки окончаний и т.д.

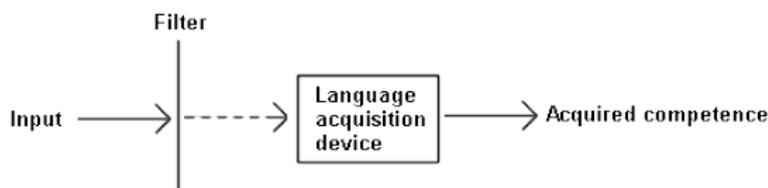


Рис. 2. Роль инпута и эмоционального фильтра [20, p. 32]

Даже в родном языке мы иногда задумываемся постфактум, правильно ли, например, мы всё согласовали в предложении. Это и есть иллюстрация роли редактора, то есть сознательно изученных языковых сведений. Инпут же работает как некий поток необходимой языковой информации, поступающей в наш мозг. Крашен называет это *language acquisition device* (термин Н. Хомского), буквально «устройство по усвоению языка» – некий предполагаемый участок мозга, некая функция, которая ответственна за усвоение языкового материала. И если инпут доставляется в этот участок благополучно, это приводит к приращению непроизвольно усвоенных языковых знаний, то есть повышается наша речевая способность, способность спонтанно производить речь. Но на пути инпута может стать фильтр как заслонка, которая не даст ему пройти в наш мозг, и тогда процесс не приносит результатов.

Главным понятием здесь выступает именно инпут, поэтому представляется необходимым остановиться на этом подробнее. Это можно наблюдать на примере усвоения родного языка ребёнком: без направленного на него достаточного количества инпута оно невозможно [15, р. 235], [17, р. 118]. При этом ребёнок должен получать инпут от родителей и других окружающих людей – технические средства (телевизор, радио и другие подобные устройства) не достигают должного эффекта, поскольку отсутствует компонент непосредственного взаимодействия и эмоционального общения [7, с. 51], [13, с. 265], [20, р. 63].

Таким образом, инпут – это действительно главная переменная в теории усвоения неродного языка и центральное понятие, потому что даже эмоциональный фильтр, вторая главная переменная, служит как бы поправкой к работе по усвоению языка в связи с получением инпута, что можно представить в виде дроби, где инпут – это числитель, а эмоциональный фильтр – знаменатель. Чем больше инпут, тем выше результат усвоения языка; чем выше эмоциональный фильтр (уровень страха, тревожности, напряжения), тем ниже результат, тем меньше эффект от инпута. Соответственно, чем больше инпута и чем меньше негативных эмоций, то есть чем ниже эмоциональный фильтр, тем лучше итоговый результат (рис. 3). Таким образом, процесс усвоения речи можно представить в виде следующей формулы:

$$\frac{\text{инпут}}{\text{эмоциональный фильтр}} = \text{усвоение языка}$$

Рис. 3. Формула усвоения языка

Другие источники, в частности эмпирические наблюдения полиглотов, также подтверждают значение инпута, ключевую роль достаточного количества языкового материала [8], [19, р. 115], [21]. В мозге работает некий статистический механизм, который прозаически можно сравнить с водопроводным счётчиком. Чем больше воды протекает по трубе, тем выше значение показаний на счётчике. Чем больше через мозг проходит языкового материала, который можно измерить количеством слов, содержащихся в поступающих устных и письменных текстах, тем лучше результат. Таким образом, измерение инпута позволяет предсказывать эффективность овладения иностранным языком в каждом конкретном случае.

3. Требования к инпуту

Однако для того, чтобы инпут был усвоен наилучшим образом, он должен соответствовать нескольким требованиям [20, р. 62–73].

1. Прежде всего, инпут должен быть понятным. Если включить аудио- или видеозапись на совершенно неизвестном языке и несколько часов её слушать либо же взять текст, пусть даже записанный, скажем, латинским алфавитом, который можно произнести, но где не понятно ни одно слово, то никакого эффекта от этого не будет. Инпут присутствует, но не соблюдается условие о его понятности. Приращение знания языка происходит, когда инпут в целом понятен и содержит небольшое количество новых элементов, о значении которых можно догадаться по контексту.

Но в этом случае возникает вопрос: если мы начинаем изучать новый язык, как сделать инпут понятным? Ведь любое сообщение на новом языке для нас будет совершенно непонятно. Здесь потребуются вспомогательные инструменты, такие как параллельный перевод, когда на одной или соседних страницах представлены и целевой текст, и его перевод на уже знакомый читателю язык [9]. Кроме того, можно использовать субтитры (если речь идёт о видеоматериалах), комментарии к устным или письменным текстам, или просто читать или смотреть материалы, которые уже знакомы на родном языке. В последнем случае мы уже имеем представление об истории, и поэтому, встретив иноязычные формы, многое можем вывести из нашего фонового знания.

2. Инпут также должен быть интересным и соответствовать личным потребностям человека, изучающего язык. Иными словами, тематика текстов должна вызывать какой-то отклик. Если ученик интересуется историей, а читает технические инструкции, это будет вызывать повышение эмоционального фильтра, поскольку будет восприниматься как неактуальная информация. Хотя языковой материал есть, у изучающего нет на него запроса. Таким образом, речь идёт о тематическом подборе. При этом сам характер темы не важен: язык можно изучать по текстам любой тематики. Важно только, чтобы она соответствовала нашим интересам и личным задачам.

3. Инпут должен быть представлен в достаточном количестве. Если прочитать один абзац текста или послушать пять минут аудиозаписи, этого будет недостаточно для того, чтобы усвоить необходимые структуры языка. Становится ясно, что инпута должно быть большое количество. Но сколько именно? Если это главная переменная при усвоении неродного языка (и родного тоже), то совершенно необходимо понимать, какой объём текстов нужно закладывать в учебники, сколько часов видеозаписи требуется для того, чтобы достичь того или иного уровня владения языком, и т.д. Таким образом, это становится центральным вопросом методических разработок, на который Крашен не даёт ответа. Он говорит, что на момент написания его книги (начало 1980-х годов) у него для этого было недостаточно данных, но в целом инпута должно быть значительно больше, чем сейчас содержится в наших программах [20, р. 73]. Однако за последние десятилетия, как представляется, определённые эмпирические данные были накоплены, и теперь можно в некоторой степени приблизиться к ответу на вопрос, сколько же именно инпута необходимо для усвоения языка [10].

4. Вопрос о количестве инпута

4.1. Данные детской речи

Первый источник информации на этот счёт – это усвоение родного языка, детская речь. В конце XX века в США было проведено исследование, связанное с усвоением английского языка как родного. Б. Харт и Т. Рисли в течение 2,5 лет наблюдали за речевым развитием детей в нескольких десятках американских семей. Эти семьи были разного социального уровня, разного материального достатка. В результате наблюдений исследователи пришли к выводу, что за час времени в семьях малого достатка ребёнок получает инпут в размере в среднем 616 слов; в семьях среднего достатка практически в два раза больше – 1251 слово; а в очень обеспеченных семьях и того больше – более 2000 слов за час (см. табл. 1). Исследователи записывали на аудиоустройство всю речь, которую дети слышали за это время в своей семье, то есть которая была направлена в сторону ребёнка, и затем посчитали количество слов. Далее, исходя из представлений о 14 часах бодрствования ребёнка в сутки, они посчитали, что за четыре года, отталкиваясь от приведённых выше показателей, ребёнок в малообеспеченной семье в США получит инпут в размере 13 млн. слов; в семьях среднего достатка – 26 млн.; и в богатых семьях – 45 млн. слов. Таким образом, налицо очень существенный разрыв. Статья, в которой были опубликованы эти данные, так и называлась: “30 million word gap” (разрыв в 30 млн. слов) [18].

Таблица 1

Результаты исследования Б. Харт и Т. Рисли

<i>Семьи</i>	<i>Инпут</i>	<i>Количество слов в час</i>	<i>Количество слов за 4 года</i>
Малообеспеченные		616	13 млн.
Среднего достатка		1251	26 млн.
Богатые		2153	45 млн.

Возникает вопрос, почему имеет место такой разрыв между семьями разного достатка? По всей видимости, это связано с тем, что родители в более обеспеченных семьях имеют возможность проводить с ребёнком больше времени и, соответственно, больше с ним общаются. Аналогичную тенденцию можно проследить в процессе обучения чтению на родном языке. Речь идёт о так называемом «эффекте Матфея», когда в более обеспеченных семьях ребёнок получает больше общения со стороны родителей: с ним больше разговаривают, ему больше читают вслух, благодаря чему он становится более подготовлен к чтению и в дальнейшем лучше читает. Это, в свою очередь, помогает ему наращивать запас слов и увеличивает систему понятий у него в голове, что способствует его дальнейшему когнитивному развитию. Вследствие этого он читает ещё лучше и больше и таким образом процесс ускоряется. А ребёнок, который недополучает подобной подготовки к чтению из устной речи, в дальнейшем, когда начинаются школьные занятия и нужно постоянно читать, сталкивается с трудностями. Он начинает отставать, и этот разрыв, который можно измерить словарным запасом, становится всё больше [5, с. 185]. Харт и Рисли также устанавливают некоторые корреляции между выявленным ими лексическим разрывом и разницей в учебных успехах и в целом интеллектуальном развитии детей из разных семей [18, р. 4, 9]. Их социальные выводы вызвали большой резонанс и споры в США.

Но в данном случае нас интересует именно числовая сторона, а именно что это исследование предоставляет нам данные о том, сколько слов усваивается ребёнком за определённый период. Что из этого можно вывести? Если посчитать среднее арифметическое между указанными величинами (13, 26 и 45 млн.), мы получим 28 млн. слов за четыре года. Данное число можно соотнести с тем, что словарный запас ребёнка в этом возрасте в родном языке варьирует в размере от 1500 до 2000 слов. Таким образом, можно видеть, какая колоссальная величина, около 30 млн. слов инпута из устной речи окружающих, прежде всего родителей, нужна, чтобы из них выкристаллизовались только около 2000 слов собственного лексического запаса ребёнка.

4.2. Данные корпусной лингвистики

Другое исследование, которое может служить источником данных о количестве инпута, было выполнено в русле корпусной лингвистики новозеландским лингвистом П. Нейшном на материале английского языка [23]. В любом языке одни слова используются в речи чаще, а другие – реже, в связи с чем возможно составление списков слов по их убывающей частотности. Нейшн поставил перед собой задачу выяснить, какой объём текста необходим, чтобы наиболее частотные английские слова (от второй до девятой тысячи) естественным образом встретились при чтении достаточное для запоминания количество раз. В данном случае он исходил из того, что для запоминания необходимо встретить новое слово в разных контекстах порядка 12 раз. Выполнив статистический поиск в корпусах английского языка, Нейшн пришёл к следующим выводам. Для второй тысячи самых частотных лексических гнезд (групп однокоренных слов) нужен текст объёмом в 200 тыс. слов, для третьей тысячи – 300 тыс., для четвёртой – 500 тыс., для пятой тысячи – 1 млн. слов, для шестой – 1,5 млн., для седьмой – 2 млн., для восьмой – 2,5 млн., и для девятой тысячи – 3 млн. слов (см. табл. 2). Таким образом, чем реже слова, тем больший объём текста необходим, чтобы произвольным образом усвоить их при чтении.

Таблица 2

Результаты исследования П. Нейшна

Самые частотные лексические гнезда английского языка	Количество инпута, необходимое, чтобы встретить в тексте порядка 12 раз
2-я тысяча	200 тыс.
3-я тысяча	300 тыс.
4-я тысяча	500 тыс.
5-я тысяча	1 млн.
6-я тысяча	1,5 млн.
7-я тысяча	2 млн.
8-я тысяча	2,5 млн.
9-я тысяча	3 млн.

4.3. Данные полиглотии

И, наконец, из данных полиглотии, то есть информации об овладении многими языками полиглотики, также можно извлечь некоторые, как представляется, интересные сведения. На Съезде полиглотов 2017 года в Братиславе, Р. Дирман выступил с докладом, в котором озвучил эмпирические наблюдения некоторых полиглотов, пользователей определённого интернет-форума, о том, сколько часов просмотра видеозаписей им потребовалось для того, чтобы перейти от одного уровня владения языком к другому (пользуясь европейской шкалой языковых уровней от А1 до С2). Из этих усреднённых данных следовало, что для достижения уровня В1 нужно было порядка 150 часов видеоматериала, для уровня В2 – 366 часов, для уровня С1 – 590 часов, и для С2 – более 2000 часов [16].

Чтобы составить представление об инпуте, который стоит за этим объёмом видеоматериалов, можно воспользоваться подсчётами П. Нейшна, который в своём ранее упомянутом исследовании отмечает, что один час чтения даёт инпут порядка 10 тыс. слов, тогда как час просмотра видеоматериалов даёт в два раза меньше, примерно 5 тыс. слов [23, р. 8]¹. Исходя из этих данных, несложными математическими вычислениями мы получаем такие величины, как 750 тыс. слов для уровня В1, примерно 2 млн. слов для уровня В2, около 3 млн. слов для уровня С1 и 10 млн. для уровня С2 (см. табл. 3).

Таблица 3

Данные о соотношении объёма видеоматериалов, уровня владения языком и лексического инпута

Уровень	Видео в часах	Инпут в словах
В1	150	750 тыс.
В2	366	1 млн. 830 тыс.
С1	590	2 млн. 950 тыс.
С2	более 2000	10 млн.

5. Словарный запас и уровень владения языком

Что же можно из этого всего вывести? Предположим, у нас есть представление об объёме инпута, но какое количество слов собственного словарного запаса будет ему соответствовать? И как соотносятся уровни владения языком с разным размером словарного запаса? Парадоксально, но по такому важному для методики изучения и преподавания языков вопросу практически

¹ Эта разница объясняется тем, что в фильмах часть лексических описаний заменяется зрительным рядом, и, кроме того, словарный запас устной речи значительно более ограничен, по сравнению с письменной. По всем этим причинам чтение представляет собой гораздо более богатый источник инпута, чем видео и аудио, хотя преимуществом последних является невербальный контекст.

нет исследований. Автору удалось найти только одно специальное исследование, посвящённое этой теме (даже не вопросу об инпуте, а вопросу о том, какие показатели лексического запаса соответствуют уровням владения языком по европейской шкале). Кроме того, очень разные оценки встречаются в интернет-источниках.

Э. Гуннемарк в своей книге «Искусство изучать языки» и, ссылаясь на него, Д. Б. Никуличева в книге «Как найти свой путь к иностранным языкам» приводят цифры, которые можно свести в следующую систему. Уровню владения языком А соответствует словарный запас в размере от 400 до 500 слов, уровню В1 – примерно в два раза больше, уровню В2 – от 1500 до 2000 слов, и уровню С – 8000 слов [6, с. 50–51], [12, с. 239]. Основой для подсчётов стали, во-видимому, наблюдения Гуннемарка о степени беглости речи и понимания текста в зависимости от количества выученных слов, которые Д. Б. Никуличева соотносит с уровнями европейской шкалы. Можно видеть, что в приведённых вычислениях между уровнями предполагается примерно двукратное увеличение. Поскольку Гуннемарк пользуется собственной шкалой языковых уровней, а Д. Б. Никуличева соотносит их с уровнями А, В1, В2 и С (не подразделяя А и С), следуя логике двукратного увеличения, для уровня А1 можно вывести величину от 200 до 250 слов, а для уровня С1 – словарный запас от 3000 до 4000 слов, который Гуннемарк рассматривает как ступень, непосредственно следующую за лексическим уровнем 1500–2000 слов.

Единственное известное специальное исследование, посвящённое вопросу о том, как объём словарного запаса соотносится с уровнем владения языком, основано на материалах кембриджских экзаменов [22]. Дж. Милтон и Т. Алексиу исследовали размер словарного запаса у людей, сдававших кембриджские экзамены для того или иного уровня владения языком, и в дальнейшем соотнесли сведения о словарном запасе этих людей с тем уровнем, которого они достигли. В результате были получены другие, отличные от Гуннемарка и Никуличевой, данные, хотя и совпадающие в некоторых местах, в частности, в уровне С1. Но отправная точка в любом случае другая (более высокая). Согласно этому исследованию, уровень А1, должно быть, соответствует словарному запасу менее 1500 слов, А2 – в среднем 2000, В1 – 3000, В2 – 3500, С1 – около 4000, и С2 приближается к 5000 (из 5000 наиболее частотных слов).

В качестве ещё одного источника информации по данному вопросу можно рассмотреть опубликованные на интернет-форумах практические наблюдения людей, изучающих языки². Среди них встречаются весьма разные оценки (в табл. 4 они приводятся как диапазон значений для каждого уровня). Как можно видеть, различие очень большое вплоть до того, что на уровне С2 наблюдается варьирование от 10 до 16 тыс. слов.

Таблица 4

Данные о соотношении уровня владения языком и словарного запаса

Уровень владения	Объём словарного запаса		
	Гуннемарк, Никуличева	Milton, Alexiou (max. 5000)	Форумы
A1	200–250	<1500	300–500
A2	400–500	1500–2500	600–1000
B1	800–1000	2750–3250	1200–2000
B2	1500–2000	3250–3750	2500–4000
C1	3000–4000	3750–4500	5000–8000
C2	8000	4500–5000	10 000–16 000

Как же оценить все эти разрозненные данные? Если подходить к этому вопросу с точки зрения частотности (чем выше наш уровень владения языком, тем более редкими словами мы владеем) и руководствоваться законом Ципфа (частота использования слов в тексте уменьшается

² What are estimates of vocabulary size for each CEFR level? URL: <https://languagelearning.stackexchange.com/questions/3061/what-are-estimates-of-vocabulary-size-for-each-cefr-level/3063> (дата обращения: 10.09.2025).

пропорционально их рангу: второе по употребительности слово встречается вдвое реже, чем первое, и т.д.), то нужно, по всей видимости, предполагать более значительное увеличение словарного запаса при переходе от одного уровня к другому. Таким образом, те подсчёты, в которых мы будем видеть большой разрыв между уровнями, будут ближе к истине. Однако необходимо оговориться, что, возможно, объём словарного запаса как математически исчисляемая величина в целом не соотносим с уровнями владения языком по европейской системе, которые представляют собой совокупность описательных характеристик. Открытым остаётся и вопрос о количестве инпута, соответствующего тому или иному размеру словарного запаса. Иными словами, сколько нужно услышать на иностранном языке устной речи, прочитав текстов (что можно измерить количеством содержащихся в них слов), чтобы достичь определённого уровня в обучении языку и того объёма словарного запаса, который ему соответствует? Этот вопрос остаётся предметом дальнейших исследований.

6. Выводы

Исходя из всех приведённых выше данных, представляется возможным сделать следующие выводы.

1. Необходимое количество инпута нужно рассчитывать применительно к каждому целевому уровню владения языком. Если мы просто называем некий объём инпута, неясно, какому уровню он соответствует. Однако одну определённую величину, соответствующую пороговому уровню, можно вывести.

2. Уровню B1–B2 по европейской системе соответствует инпут в размере 1 млн. слов. Если исходить из точки зрения, что пороговый уровень достаточен для большинства целей изучения иностранного языка, то на эту величину можно ориентироваться при составлении учебных программ и разработке учебных материалов³.

Как представить себе объём в 1 млн. слов? Посчитано, например, что семь книг о Гарри Поттере содержат на английском немногим более 1 млн. слов [25, р. 16]. Эта величина представляется вполне достижимой, если мы зададимся целью и распределим этот объём, притом в него включается не только чтение письменных текстов, но вся речь, которую воспринимает человек: речь преподавателя, обращённая к студентам, речь, воспринимаемая в разговоре с носителем языка, и т.д. Всё это учитывается статистическим механизмом, если инпут (возвращаясь к вопросу о необходимых требованиях) понятен, соответствует интересам, потребностям изучающего и предоставляется в достаточном количестве. Соответственно, исходя из этого числа, мы можем планировать, сколько текстов на иностранном языке нужно включить в пособия, сколько заложить в программы часов аудио- и видеозаписи и т.д. Таким образом, появляется научно обоснованная отправная точка для расчётов, связанных с учебными программами по иностранным языкам.

3. Из приведённых данных можно видеть, что ребёнку при усвоении родного языка нужно значительно (множественно, в десятки раз) больше инпута, чем взрослому человеку при изучении иностранного языка. Здесь можно видеть подтверждение закономерности, заключающейся в том, что родной язык усваивается медленнее, чем иностранный. Когда мы изучаем иностранный язык, у нас уже есть когнитивная система на основе родного языка, и новые данные ложатся на эту систему, поэтому навыки иностранного языка можно развивать значительно быстрее (хотя они и зависят от соответствующих навыков родного языка).

³ Разумеется, здесь речь не идёт о более высоких уровнях владения языком (профессиональное лингвистическое образование и т.п.). В этом случае нужно предполагать значительно большее количество инпута.

7. Практические рекомендации

Из теории усвоения неродного языка, из гипотезы о понятном инпуте, подтверждаемой многочисленными данными, и из тех сведений, которые у нас есть для того, чтобы ответить на вопрос о необходимом количестве инпута, можно вывести следующие практические рекомендации.

1. Если речь идёт об усвоении родного языка, нужно больше разговаривать с детьми для их полноценного речевого, лексического и интеллектуального развития. Это подтверждается и исследованиями, связанными с освоением навыка чтения. Чем больше с ребёнком говорят, тем лучше он готов к дальнейшему чтению, тем лучше развивается его словарный запас и в дальнейшем когнитивные способности в целом.

2. Мы можем учитывать инпут, когда мы разрабатываем учебные пособия и программы. В этом случае преподаватель, который захочет выстроить все учебники в систему, исходя из теории усвоения неродного языка, сможет посчитать предполагаемое количество инпута, который будет дан студентам в программах за счёт занятий, чтения, аудиоматериалов и т.д.⁴ Эти данные можно без труда получить с помощью функции статистики в текстовом редакторе. Если речь идёт об аудио- и видеозаписях, объём инпута можно вычислить исходя из того, что один час записи содержит примерно 5 тыс. слов.

3. Необходимо включать больше аудио- и видеоматериалов в занятия по иностранным языкам. Почему? Не с точки зрения их богатства инпута – выше было показано, что чтение даёт существенно больше лексического материала. Но аудио- и видеоматериалы хороши на более ранних стадиях изучения нового языка, поскольку содержат более широкий контекст. Когда мы читаем письменный текст, мы не видим в письменных знаках ни выражения лица, ни жесты, ни зрительного ряда и т.д. Чем шире контекст, тем больше у нас возможностей для догадки о значении новых лексических единиц. Таким образом, повышается понятность инпута. К тому же аудио- и видеоматериалы задействуют более широкий спектр эмоций и поэтому в большей степени способствуют снижению эмоционального фильтра (обычно ученики с большим расположением готовы смотреть видео и слушать аудио, чем читать письменный текст, если у них недостаточная мотивация). Кроме того, известно, что устная речь предшествует развитию письменной речи [3, с. 100, 104]. Поэтому, если в начале изучения нового языка не дать достаточного количества аудио- и видеоматериалов, то переход к письменным текстам будет в дальнейшем объективно затруднён.

4. Следует стимулировать экстенсивное чтение текстов по выбору учащихся. «Экстенсивное» означает направленное на максимальное количество, а не на глубокий анализ каждой конкретной языковой формы, которая встречается на странице. Интенсивное чтение имеет место тогда, когда мы берём небольшой отрывок текста и детально его анализируем. В этом случае происходит работа «редактора», пользуясь технологией Крашена, то есть мы пользуемся сознательно изученными языковыми сведениями. Таким образом происходит изучение (*learning*), но не усвоение (*acquisition*). Усвоение же будет происходить в том случае, когда мы предоставим большое количество текстов, в которых человек, изучающий язык, естественным образом столкнётся с новым для него материалом, значение которого он сможет вывести из контекста и таким образом неосознанно усвоить.

5. Следует также поощрять работу студентов с информацией по интересующим их темам на иностранном языке. Человек интересуется, может быть, какими-то научными дисциплинами, общественной жизнью, международными отношениями и т.д. Ему интересно получать информацию, читать новости, связанные с этой тематикой. Всё то же самое можно делать на иностранном языке. Когда нам интересен текст не только ради того, чтобы он помог нам овладеть языком, но и просто ради своего содержания, тогда происходит наилучшее усвоение языка.

⁴ Пример учёта количества инпута см. в пособиях МГИМО, подготовленных при участии автора [1], [2], [11].

Крашен называет это «принципом забвения» (*forgetting principle*) [20, р. 66]. Предположим, мы читаем или слушаем текст на иностранном языке, и мы настолько заинтересованы в его содержании, что даже забываем, на каком языке оно выражено. Именно в этот момент, утверждает Крашен, усвоение языка происходит самым лучшим образом, то есть можно констатировать наиболее низкий эмоциональный фильтр, при котором инпут свободно поступает в наш мозг и усваивается, повышая языковую компетентность.

6. Наконец, поскольку инпут всегда работает в паре с эмоциональным фильтром (даже если у нас в наличии самый лучший инпут, психологический барьер, чувство тревожности, страха будет препятствовать усвоению языкового материала), необходимо обращать внимание на психологическое состояние и установки учащихся: как они чувствуют себя во время занятий, с какими эмоциями ассоциируется у них процесс изучения языка и т.д. Естественно, нужно максимально поддерживать те формы работы, которые связаны с позитивными эмоциями, и постараться устранить факторы, вызывающие негативные эмоции. Следует всячески подчёркивать значение, полезность иностранного языка, то есть способствовать развитию мотивации, осознанного отношения студентов к обучению, потому что это будет также снижать эмоциональный фильтр и повышать интерес к доступному материалу, вследствие чего инпут будет усваиваться лучше.

8. Перспектива исследования

Поиск ответов на вопросы, связанные с измеряемыми аспектами овладения языком, требует получения большего количества данных, в том числе с помощью экспериментов по изучению языков, в ходе которых будет регистрироваться объём поступающего инпута и рост словарного запаса⁵. Сейчас мы имеем представление о тех ключевых факторах, от которых зависит успех этого процесса, и можем предсказывать его эффективность. Если наличествует необходимый речевой материал, отвечающий известным требованиям, если атмосфера способствует низкому эмоциональному фильтру, мы можем предсказывать эффективное усвоение языка. Если нет, мы будем предполагать, что развитие речевых способностей будет протекать значительно медленнее. Однако задачи образовательного процесса требуют более конкретного выстраивания учебных программ и пособий. И именно данные об измеряемых аспектах овладения языком, в том числе о количестве инпута и размере словарного запаса, позволят, как представляется, разработать такие методики, которые будут отличаться значительно большей эффективностью, в сравнении с тем, что история методической науки знала до этого.

© Г.А. Казаков, 2025

Список литературы

1. Английский язык. *Discovering British History* / Н. С. Измайлова, С. Д. Камалова, А. Е. Локонова и др.; под ред. М. М. Конколь, Г. А. Казакова. М.: МГИМО-Университет, 2022. 144 с.
2. Английский язык. Социология: новые перспективы / А. О. Козьмин, Н. В. Панич, Д. И. Милюхина и др.; под ред. Г. А. Казакова. М.: МГИМО-Университет, 2024. 232 с.
3. Беляев Б.В. Очерки по психологии обучения иностранным языкам. Изд. 2-е. М.: Просвещение, 1965. 227 с.
4. Бигулов А.К. Эксперимент по применению методики гиперинтенсива по вхождению в разговорную практику немецкого языка только во взаимодействии с ИИ-тьюторами // Теория речевой деятельности – новая парадигма в науке о языке. Воронеж: Научная книга, 2025. С. 60–61.
5. Вулф М. Пруст и кальмар: нейробиология чтения. М.: КоЛибри, Азбука-Аттикус, 2020. 384 с.
6. Гуннемарк Э. Искусство изучать языки. Изд. 2-е. СПб.: Тесса, 2002. 208 с.
7. Иванов Вяч. Вс. Чет и нечет: асимметрия мозга и знаковых систем. М.: Сов. радио, 1978. 184 с.

⁵ В качестве новейшего примера такого рода можно привести эксперимент А. К. Бигулова [4]. Фактор времени также относится к числу измеряемых аспектов овладения языком, но в силу своего значения заслуживает отдельного рассмотрения. Пример количественного исследования по этой теме см. в работе Я. В. Алешкевича-Суслова [14].

8. Казаков Г.А. Съезд полиглотов в Братиславе: Polyglot Gathering 2017 // Вопросы психолингвистики, 2017, №3 (33). С. 258–260.
9. Казаков Г. Параллельный перевод как инструмент овладения иностранным языком // Translatoryka i glottodydaktyka: od teorii do praktyki. Gdansk: Wydawnictwo Uniwersytetu Gdańskiego, 2019. С. 189–199.
10. Казаков Г.А. Сколько нужно инпута для овладения языком: данные детской речи, корпусной лингвистики и полиглоти // Теория речевой деятельности: вызовы современности. М.: Канцлер, 2019. С. 245–246.
11. Казаков Г.А. Английский язык. Discovering Shakespeare / Г. А. Казаков, С. С. Шеменкова; под ред. Т. А. Ивушкиной. М.: МГИМО-Университет, 2024. 162 с.
12. Никуличева Д.Б. Как найти свой путь к иностранным языкам: лингвистические и психологические стратегии полиглотов. М.: Флинта: Наука, 2009. 304 с.
13. Пинкер С. Язык как инстинкт. М.: УРСС, 2004. 456 с.
14. Aleshkevich-Suslov Ya. Assessing successive learning acceleration in Germanic languages: a research paradigm proposal // Филологические науки в МГИМО. 2022. Vol. 8, No. 3. P. 35–46. DOI: 10.24833/2410-2423-2022-3-32-35-46.
15. Crystal D. The Cambridge Encyclopedia of Language. Cambridge University Press, 1987. 472 p.
16. Dearman R. Using TV, Anki, and subtitles to learn a language. Paper presented at Polyglot Gathering 2017. Bratislava, 01.06.2017.
17. Fry D. Homo Loquens. Man as a Talking Animal. Cambridge University Press, 1978. 177 p.
18. Hart B. The early catastrophe: The 30 million word gap by age 3 / B. Hart, T. Risley // American Educator, 2003 (spring). P. 4–9.
19. Kaufmann S. The Linguist: A Personal Guide to Language Learning. Canada, 2003. 146 p.
20. Krashen S. Principles and Practice in Second Language Acquisition. New York: Phoenix ELT, 1995. 202 p.
21. Krashen S. Polyglots and the comprehension hypothesis // Turkish Online Journal of English Language Teaching, 2017, vol. 2, no. 3. P. 113–119.
22. Milton J. Vocabulary size and the Common European Framework of Reference for Languages / J. Milton, T. Alexiou // Vocabulary Studies in First and Second Language Acquisition. London: Palgrave Macmillan, 2009. P. 194–211.
23. Nation P. How much input do you need to learn the most frequent 9,000 words? // Reading in a Foreign Language, 2014, vol. 26, no. 2. P. 1–16.
24. Richards J. Approaches and Methods in Language Teaching / J. Richards, T. Rodgers. 2nd ed. Cambridge University Press, 2007. 270 p.
25. Seikei Freshers' English Extensive Reading Teacher's Manual. 3rd ed. Tokyo: Seikei Institute for International Studies (Seikei University), 2015. 18 p.

References

1. Izmailova, N., et al. *Angliyskiy Iazyk* [English]. *Discovering British History* (M. Konkol & G. Kazakov eds.). Moscow: MGIMO University, 2022. 144 p.
2. Kozmin, A., et al. *Angliyskiy Iazyk* [English]. *Sociology: A Novel Perspective* (G. Kazakov ed.). Moscow: MGIMO University, 2024. 232 p.
3. Belyaev, B. *Ocherki po Psikhologii Obucheniia Inostrannym Iazykam* [Essays on the Psychology of Language Learning]. 2nd ed. Moscow: Prosveshchenie, 1965. 227 p.
4. Bigulov, A. Eksperiment po primeneniiu metodiki giperintensiva po vhozhdениu v razgovornuiu praktiku nemetskogo iazyka tolko vo vzaimodeystvii s II-tutorami [Experiment on applying the hyperintensive method for entering conversational practice in German using only AI tutors]. *Teoriia Rechevoy Deiatelnosti – Novaia Paradigma v Nauke o Iazyke* [Theory of Speech Activity – New Paradigm in the Science of Language]. Voronezh: Nauchnaia Kniga, 2025. P. 60–61.
5. Wolf, M. *Prust i Kalmar: Neyrobiologіia Chteniia* [Proust and the Squid: Neorobiology of Reading]. Moscow: KoLibri, Azbuka-Attikus, 2020. 384 p.
6. Gunneemark, E. *Iskusstvo Izhuchat' Iazyki* [The Art of Learning Languages]. 2nd ed. Saint Petersburg: Tessa, 2002. 208 p.
7. Ivanov, V. *Chet i Nechet: Assimetriia Mozga i Znakovykh Sistem* [Even and Odd: Asymmetry of the Brain and Sign Systems]. Moscow: Sovetskoe Radio, 1978. 184 p.
8. Kazakov, G. Polyglot Gathering in Bratislava 2017. *Journal of Psycholinguistics*, 2017, no. 3 (33). P. 258–260.
9. Kazakov, G. Parallelny perevod kak instrument ovladeniia inostrannym iazykom [Parallel translation as an instrument of language learning]. *Translatoryka i Glottodydaktyka: Od Teorii do Praktyki*. Gdansk: Wydawnictwo Uniwersytetu Gdańskiego, 2019. P. 189–199.
10. Kazakov, G. Skol'ko nuzhno inputa dlia ovladeniia iazykom: dannye detskoй rechi, korpusnoy lingvistiki i poliglotti [How much input do we need to learn a language: data from child speech, corpus linguistics and polyglottery]. *Teoriia Rechevoy Deiatelnosti: Vyzovy Sovremennosti* [Theory of Speech Activity: Modern Challenges]. Moscow: Kanzler, 2019. P. 245–246.
11. Kazakov, G., & Shemenkova, S. *Angliyskiy Iazyk* [English]. *Discovering Shakespeare* (T. Ivushkina ed.). Moscow: MGIMO University, 2024. 162 p.
12. Nikulicheva, D. *Kak Nayti Svoй Put' k Inostrannym Iazykam: Lingvisticheskie i Psikhologicheskie Strategii Poliglotov* [How to Find Your Way to Foreign Languages: Linguistic and Psychological Strategies of Polyglots]. Moscow: Flinta; Nauka, 2009. 304 p.
13. Pinker, S. *Iazyk kak Instinkt* [The Language Instinct]. Moscow: URSS, 2004. 456 p.
14. Aleshkevich-Suslov, Ya. Assessing successive learning acceleration in Germanic languages: a research paradigm proposal. *Linguistics & Polyglot Studies*, 2022, vol. 8, no. 3. P. 35–46. DOI: 10.24833/2410-2423-2022-3-32-35-46.
15. Crystal, D. *The Cambridge Encyclopedia of Language*. Cambridge University Press, 1987. 472 p.

16. Dearman, R. Using TV, Anki, and subtitles to learn a language. Paper presented at Polyglot Gathering 2017. Bratislava, 1 June 2017.
17. Fry, D. *Homo Loquens. Man as a Talking Animal*. Cambridge University Press, 1978. 177 p.
18. Hart, B., & Risley, T. The early catastrophe: The 30 million word gap by age 3. *American Educator*, 2003 (spring). P. 4–9.
19. Kaufmann, S. *The Linguist: A Personal Guide to Language Learning*. Canada, 2003. 146 p.
20. Krashen, S. *Principles and Practice in Second Language Acquisition*. New York: Phoenix ELT, 1995. 202 p.
21. Krashen, S. Polyglots and the comprehension hypothesis. *Turkish Online Journal of English Language Teaching*, 2017, vol. 2, no. 3. P. 113–119.
22. Milton, J., & Alexiou, T. Vocabulary size and the Common European Framework of Reference for Languages. *Vocabulary Studies in First and Second Language Acquisition*. London: Palgrave Macmillan, 2009. P. 194–211.
23. Nation, P. How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 2014, vol. 26, no. 2. P. 1–16.
24. Richards, J., & Rodgers, T. *Approaches and Methods in Language Teaching*. 2nd ed. Cambridge University Press, 2007. 270 p.
25. *Seikei Freshers' English Extensive Reading Teacher's Manual*. 3rd ed. Tokyo: Seikei Institute for International Studies (Seikei University), 2015. 18 p.

Сведения об авторе:

Григорий Александрович Казаков – кандидат филологических наук, PhD, ответственный секретарь журнала «Linguistics & Polyglot Studies» («Филологические науки в МГИМО»), доцент кафедры английского языка №3 МГИМО МИД России. Сфера научных интересов: языковое образование, полиглотия, психолингвистика, психология мастерства, язык и религия, перевод священных текстов, внешкольная педагогика. E-mail: g.kazakov@my.mgimo.ru. ORCID: 0000-0002-6481-2564.

About the author:

Grigory Kazakov, PhD, is Executive Secretary of *Linguistics & Polyglot Studies* and Associate Professor at Department of English No.3, MGIMO University. Research interests: language education, polyglottery, psycholinguistics, psychology of expertise, language and religion, translation of sacred texts, Scouting pedagogy. E-mail: g.kazakov@my.mgimo.ru. ORCID: 0000-0002-6481-2564.

* * *