

The Order of the Acquisition  
of the English Article System by Polish Learners  
in Different Proficiency Groups



The Order of the Acquisition  
of the English Article System by Polish Learners  
in Different Proficiency Groups

By

Artur Świątek

**CAMBRIDGE  
SCHOLARS**

---

P U B L I S H I N G

The Order of the Acquisition of the English Article System  
by Polish Learners in Different Proficiency Groups  
By Artur Świątek

This book first published 2014

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data  
A catalogue record for this book is available from the British Library

Copyright © 2014 by Artur Świątek

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-5629-0, ISBN (13): 978-1-4438-5629-4

To my wife and sons



# TABLE OF CONTENTS

<b>Chapter One: Acquisition of Articles in L1</b> .....	1
1. Introduction .....	1
1.1. L1 Acquisition of Articles .....	2
1.2. The Age of Acquisition .....	4
1.3. Evaluating Competence from the Naturalistic Data .....	5
1.4. Specificity and Non-Specificity .....	6
1.5. Competence with Discrepancy of Viewpoints .....	7
1.6. Ensuring Specificity for the Addressee .....	7
1.7. Specificity Introduced via Spoken Communication .....	8
1.8. Specificity by Entailment .....	9
1.9. The Semantics of Articles .....	9
1.10. Naming .....	10
1.11. The Conceptual Basis for Definite and Indefinite Reference .....	11
1.12. Understanding Classes and Class Membership .....	12
1.13. Summary .....	14
<b>Chapter Two: Acquisition of Articles in L2</b> .....	15
2. Introduction .....	15
2.1. L2 Acquisition of Articles .....	15
2.2. Comparing L1 and L2 Acquisition Data .....	19
2.3. Definiteness and Indefiniteness in Polish .....	21
2.4. Foreign and Polish Learners' Problems with the Acquisition of English Article System: Current Research Analysis .....	23
2.4.1. Discussion of Tryzna's Study .....	30
2.5. Conclusion .....	32
<b>Chapter Three: The Use of English Article System</b> .....	35
3. Introduction .....	35
3.1. Types of Articles .....	35
3.2. Classes of Nouns .....	38
3.3. Usage Contexts .....	42
3.3.1. Quirk et al.'s Classification .....	42
3.3.2. Langacker's Classification .....	46
3.4. Classification of the Article Environments .....	48
3.5. Function and Meaning of the Articles Classified by Berezowski .....	51
3.6. Summary .....	58

<b>Chapter Four: The Empirical Study</b> .....	60
4. Explanation of Study Design.....	60
4.1. Aim of the Study .....	60
4.2. Schema of the Study.....	60
4.3. Accepted Methodology Procedures.....	61
4.4. Criteria of Sample Selection.....	61
4.5. Sample Description.....	63
4.6. Tools for Measurement.....	64
4.6.1. Evaluation Criteria.....	64
4.7. Course of the Study .....	65
4.8. Presentation of Study Results .....	65
4.9. Schema of Data Presentation.....	65
4.10. Data Analysis Interpretation.....	65
4.10.1. Type 1 [-SR, +HK]: Generic Nouns (a, the, 0).....	66
4.10.2 Type 2 [+SR, +HK]: Referential Definites (the).....	66
4.10.3. Type 3 [+SR, -HK]: Referential Indefinites (a, 0).....	67
4.10.4. Type 4 [-SR, -HK]: Nonreferentials (a, 0).....	68
4.10.5. Type 5: Idioms and Other Conventional Uses (a, the, 0).....	69
4.11. Global Juxtaposition among Groups.....	70
4.12. The Impact of Each Type of Article on Particular Groups of Polish Learners.....	72
4.12.1. Elementary.....	72
4.13. The Detailed Tabular Analysis of Sentences According to Each Type of Article.....	73
4.14. Intermediate.....	96
4.15. Advanced.....	120
4.16. Research Conclusions And Pedagogical Implications for Teaching the Use of English Articles for Polish L2 Learners.....	143
4.16.1 Elementary Group.....	143
4.16.2. Intermediate Group.....	144
4.16.3. Advanced Group.....	145
<b>4.17. Final Conclusions</b> .....	145
4.17.1. Elementary Group.....	147
4.17.2. Intermediate group.....	147
4.17.3. Advanced group.....	148
<b>5. Streszczenie</b> .....	148
Kolejność Przyswajania Przedimka.....	150



<b>6. References</b> .....	152
<b>7. Appendices</b> .....	163
7.1. Appendix 1: Typology of article use .....	163
7.2. Appendix 2: Test Instrument (Version with inserted articles) .....	169
7.3. Appendix 3: Test Instrument (Version for subjects/students) .....	172



# CHAPTER ONE

## ACQUISITION OF ARTICLES IN L1

### 1. Introduction

The English system of articles is intricate, and various aspects of the language bear on the distribution of articles (Thomas, 1989:335). Both L1 and L2 research traditionally commence by identifying contexts for the appearance of articles, then investigate learners' production of articles in those environments. One common classification divides environments for articles according to whether the noun on which the article is dependent is used referentially or non-referentially, and whether or not that noun is (or can be treated as) identifiable by the recipient. These two features vary independently, yielding four cross-classified environments, each associated with one or more possible articles. The lexical properties of the noun (singular or plural, mass or count) determine article choice from among the possibilities available in a given environment (Thomas, 1989:336). Some of the four environments comprise several subclasses of contexts sharing the same features.

This classificatory system is based on Huebner (1983). It classifies nouns as plus or minus *specific referent* ([+/- SR]) and plus or minus *assumed known to the hearer* ([+/- HK]). Nouns classified as [-SR, +HK] are *generics* and are marked with *a*, *the* or *0*. Nouns classified as [-SR, -HK] are *non-referentials*. This comprises nouns that name a class to which another noun is asserted to belong to or that refer to an unspecified member of a class; *a* and *0* are the relevant articles. The category [+SR, -HK] includes *first mention nouns*, whose referent is identifiable to the speaker but not the listener, that is, nouns that the speaker is entering into the discourse for the first time. These are marked with *a* or *0*. Once a noun has entered the discourse, it normally becomes a *referential definite* [+SR, +HK] and is marked with *the*.

## 1.1. L1 Acquisition of Articles

Cziko (1986) proposes a four-stage sequence in the L1 acquisition of articles and seeks support for his proposal in independent work by seven researchers: Bresson (1974), Brown (1973), Emslie and Stevenson (1981), Garton (1983), Karmiloff-Smith (1979), Maratsos (1976), and Warden (1976).

In Stage 1 (two years and three months), children mark all referential nouns, both [+HK] and [-HK], with either *a* or *the*, but do not use articles with non-referential nouns. At Stage 2 (two years and ten months), children employ *the* in [+SR] contexts and *a* in [-SR] contexts. It is not until Stage 3 (three years and one month) that the child begins to acquire sensitivity to feature [+/- HK], possibly resulting in the reintroduction of *a* into both [+SR] environments. At Stage 4 (three years and eight months) the child has acquired the adult system of classification of nouns, which assigns articles according to both the features [+/- SR] and [+/- HK]. All stages are presented in Table 1 below, adapted from Cziko (1986, 881).

Cziko's four stages are in part a projection of Bickerton's (1981, 1984) Language Bioprogram Hypothesis, which holds that learners have an innate sensitivity to specificity and nonspecificity of reference. Bickerton observes that creole languages produce zero-form articles in [-SR] contexts (namely with generics and nonreferential nouns) and overt articles in [+SR] contexts. He sees in this fact a parallel to data from L1 acquisition studies, citing some of the same research already mentioned. Bickerton argues that children would be *'highly unlikely to derive [a sensitivity to specificity] from analysis of purely linguistic context* (1981, 151), and that this therefore represents an aspect of the child's 'language bioprogram'.

Referring back to Cziko's proposed developmental sequence in referential indefinite ([+SR, -HK]) environments one may state that children from Stage 2 onwards use *a* appropriately in nonreferential ([-SR, -HK]) environments and *the* appropriately in referential definite ([+SR, +HK]) environments. But these early successes contrast with their mistakes in [+SR, -HK] environments: children frequently use definite rather than indefinite articles with first-mention nouns.

In Brown's (1973: 353) naturalistic longitudinal study of three 2-4-year-olds (Adam, Eve and Sarah, the ones that will be discussed later), this was the largest single class of errors with articles. Experimental work by Warden (1976:109) demonstrated that 3-year-olds (n=16) used definite articles to introduce first-mention nouns in a story-telling task at an overall

frequency of 54%. This figure declined gradually to 18% for his 9-year-olds.

**Table 1. Proposed four stages in the L1 acquisition of English articles**

	Stage 1	Stage 2	Stage 3	Stage 4
[-SR, +HK] Generics	* $\emptyset$	<i>a</i>	<i>a</i>	$\emptyset$ , <i>a</i> , <i>the</i>
[-SR, -HK] Nonreferentials Attributive indefinites Nonreferential indefinites	* $\emptyset$	<i>a</i>	<i>a</i>	<i>a</i>
[+SR, -HK] Referential indefinites First-mention nouns	<i>a</i> , * <i>the</i>	* <i>the</i>	<i>a</i> , * <i>the</i>	<i>a</i>
[+SR, +HK] Referential definites Previous- mention nouns Specification by entailment Specification by definition Unique in all contexts	* <i>a</i> , <i>the</i>	<i>the</i>	(* <i>a</i> ), <i>the</i>	<i>the</i>

\*Predicted errors in article use.

Power and Dal Martello (1986:150) replicated Warden's results with Italian-speaking children (n=50), although they found that use of the definite article declined more sharply with age.

Karmiloff-Smith's (1979:144) French-speaking children aged 3-11 (n=68) consistently used indefinite articles in [+SR, -HK] contexts at lower frequencies than they used definite articles in [+SR, +HK] contexts in a story-completion task. As age increased, overgeneralization of definite articles decreased, and the appropriate use of indefinite articles increased.

There is various evidence that children employ the definite article when introducing a noun for the first time, in contrast to adult use of the

indefinite article. This result has surfaced in several languages, and in both naturalistic and experimental data. Brown, Maratsos, Warden, and Power and Dal Martello speculate that this finding constitutes evidence of the child's 'egocentricity' in the Piagetian sense. Warden (1976:110) writes that a child '*is unable to adopt his audience's point of view. From his own egocentric view point, a referent is specified as soon as he (the speaker) is familiar with it; he fails to realize that his audience will only become familiar with his referent after he has identified it for them verbally*'. Therefore, young children mark first-mention nouns with **the** because they assume that whatever is known to them is also known to their listener. Cziko (1986:881) seems at first to accept that egocentricity shapes the child's use of definite and indefinite reference. However, he later (1986:896) proposes that overgeneralization of the definite article is due instead to the association of **the** with [+SR] environments and **a** with [-SR] environments.

Butler (2002:453) replicates the opinion that children seem to acquire the article system at a relatively early age in L1 acquisition (somewhere between 2,8 and 3,8 years old), and they typically exhibit a low frequency of overall errors. It has been reported that although L1 child acquirers overuse the definite article **the** on occasions in which listeners do not have any knowledge of the reference (i.e., the [+SR, -HK] case), they do not make errors when the referents are nonspecific for both the speakers and hearers (i.e. the [-SR, -HK] case; Brown, 1973; Maratsos, 1971, 1976). That is, although L1 child acquirers do not seem to be able to sufficiently detect a given listener's presumed knowledge, they do appear to be able to easily distinguish specificity from nonspecificity (as coded by articles) from a very young age. This result is particularly interesting because '*specific and nonspecific references are connected in no clear way with external physical attributes or relations of perceived objects*'

(Maratsos, 1976: 94). Such results led in part to Bickerton's (1981, 1984) Language Bioprogram Hypothesis, mentioned above.

## 1.2. The Age of Acquisition

The articles abound in examples of forms that may profitably be observed naturalistically (Maratsos, 1976:15). Particularly in the study of the development of word meanings, this is not always so. Unlike basic syntactic constructions, individual words may occur only infrequently in speech samples. Even rather basic lexical items such as *front* and *back* or *know* or *more* and *less* may occur with extremely low frequency in spontaneous speech. The first uses discovered in samples may fall well

after the child has begun to understand and use the items, for the same reason that an adult's understanding of words such as *astronomy* is represented only poorly in speech. Some forms, however, occur quite commonly in speech, even obligatorily so. Brown (1973) has with a group of associated investigators (Bellugi, 1967; Brown and Bellugi, 1964; Brown, Cazden, and Bellugi, 1968; Brown and Hanlon, 1970; Cromer, 1968) gathered longitudinal, naturalistic recordings of the speech of three children who have become well-known in the literature of language acquisition as Adam, Eve and Sarah. Among the problems studied (Brown, 1973) has been the children's acquisition of fourteen very commonly used grammatical morphemes such as *in*, *on*, the progressive ending *-ing*, the past tense *-ed* and conveniently, *the articles a and the*.

For each of these morphemes Brown (also Cazden, 1968) has defined a point of stable acquisition. By various criteria it can be judged for various morphemes whether or not the morpheme should appear in a context. Some cases are simple. The definite article *the* must appear before some terms, e.g. in a frame such as '*This is \_\_\_\_\_ middle one*', or before the word *same*, e.g. '*I saw \_\_\_\_\_ same one.*' In other cases a combination of grammatical context and situational context determines an obligation to use one of the articles. Acquisition of morpheme is said to take place when the morpheme has appeared in its obligatory contexts at least 90 per cent of the time in three consecutive speech samples. In chronological terms, using these criteria, stable usage of the articles appeared for the three children studied between an estimated thirty-two months (for Eve, the generally fastest child) and forty-one months of age (Sarah, generally the slowest child).

### 1.3. Evaluating Competence from the Naturalistic Data

It would be remarkable for children to command the full semantics of definite and indefinite articles at around three years of age, but information about the stability of use cannot by itself provide the information necessary to draw this conclusion. Brown (1973) has also attempted an assessment of the semantic appropriateness of the children's usage, drawing on samples of the children's speech around the time of stable acquisition, and comparing the use of the article with the verbal and non-verbal context of its use. With such data he studied the appropriateness of the child's use where possible, although there were a good many doubtful cases.

### 1.4. Specificity and Non-Specificity

For those instances in which the child's point of view was compatible with his addressee's, successful references very largely outnumbered the unsuccessful. A large number of references specific for neither speaker or listener were successfully made. The examples included 'Put *a band-aid* on it' (Eve), 'This don't have *a wheel* on it', and 'Make *a B*' (Sarah). No example was found of a child erroneously referring definitely to non-unique or non-existent class members. Despite the abstractness of this category, the three children seemed to have good knowledge of its use by the time they were using the articles stably. They also displayed many correct uses of references specific for both the speaker and the listener alike. References included those **unique for all** (*the sky, the ground*), **unique in a given setting** (*the floor, the couch, the ceiling*), **salient for a social group** (*the mailman, the TV, the subway*) and **definite reference to other conspicuous unique objects**. Definite references were also found which were **specified by entailment** (*the driver's wheel, the motor* (of a train), *the nose, the nurse* (at a doctor's office), parts of a family (*the grandma*), and of a band-aid (*the sticky of the bandage*); **by definition** ('That's *the middle*', *the next page*); and **specified by prior utterance** ('That's *a jeep*. I put some in *the jeep*'). There were also errors, especially in the **categories of entailment** (e.g. 'Where there's *a heel*?' said of a particular sock; *a chin*, in naming features of a face) and in **references specified by prior utterance** ('I never drop *a watch*', said of an already specified watch; '*a jeep* is coming', of an already mentioned jeep). Brown (1973:35) suggests, however, that

'there are far too many correct unlimited instances of both categories of reference to suppose that the children did not know that when a whole entailed one of a certain part of the article should be **the** and that when a reference is repeated the article should be **the**'.

The occurring errors are ascribed to children's possible lack of knowledge of part-whole assemblages in the case of entailment, and occasional failure to keep track of previously specified references rather than general ignorance in the failure of definite reference to previously referred objects. The children's early usage apparently displays considerable knowledge of the semantic factor of specificity vs. non-specificity.



## 1.5. Competence with Discrepancy of Viewpoints

The case is different for instances in which the child and his listener's knowledge did not converge (Maratsos, 1976: 17). The examples were infrequent when the child made indefinite reference to objects for which reference was specific for the child but not his interlocutor. Many of the examples in this category appear to be types of naming statements, such as 'He's *a witch*' (Adam) and 'It's *a gun*.' In contrast, the transcripts indicated many apparently erroneous definite references made by the child when the reference was specific only for himself. A convincing sign of the erroneousness of the reference in many categories was the puzzled or interrogative response of the mother, presented in two exchanges below:

Sarah:     *The cat's* dead.  
 Mother:    What cat?  
 Adam:     Put it up, *the man*, says.  
 Mother:    Who's the man?

Such replies demonstrate the failure of the reference to be uniquely identifiable for the listener quite directly. In other cases the response may display the problem in a more subtle fashion. Eve, for instance, once asked 'Where's *the stool*?', apparently having one in mind, and her mother replied 'There's one over here', demonstrating that the reference to *stool* did not elicit knowledge in her mother of a particular stool Eve intended to refer to. In this category, children's erroneous definite references were far more frequent than correct indefinite ones. Children's expected failure to take into account the point of view of the other finds substantiation in these examples.

## 1.6. Ensuring Specificity for the Addressee

How can the child be convinced that his addressee understands a reference with the same specificity as he does? There exist a number of ways to justify it. Some references are specific for all even without further specification because of their general uniqueness. Such references include, for instance *the sun* or *the moon* or *the ground*. In various social groups references will be specific for all because of shared knowledge of the members of the group (Maratsos, 1976: 3). In a house with a dog for a pet, a reference to *the dog* will be easily comprehended, as will a reference to *the car* in a family with one car. Some references are specific because of the conspicuousness of their referent in the immediate environment. In a living room, *the couch* generally refers unambiguously. Or pushing a

table, it is clear what is meant by *the table* in asking ‘Where should we put *the table*?’ Sometimes a speaker may be able to take precautions to make sure that his reference is comprehended. If he has a particular table in mind out of a number in the room, he may point to the table or act on it in some other way to make it conspicuous. Or he may elaborate on the class description so that only one member is present, by saying, for example ‘Let’s put this on *the table over by the piano*’. Where *table* may not make specific reference, *the table over by the piano* might. Specificity of reference does not inhere in the object referred to, but in the relation between the object and the class membership description given by the linguistic expression. Apparently, the same speaker may participate in various contexts, and the meaning of his references changes accordingly. In the United States generally reference to *the president* without further description reliably means the current chief executive officer of the country; on the contrary to a group of academicians discussing university affairs, a reference to *the president* could easily mean the university executive.

## 1.7. Specificity Introduced via Spoken Communication

The problem frequently arises that a reference specific for the speaker cannot be made specific for his interlocutor by any of the above means, not by socially shared knowledge or by induced physical conspicuousness. The problem becomes more evident in speaking of a referent neither already known to the listener nor physically present. The speaker must then use purely conversational means to lend specificity to the reference for his listener.

What the speaker must do in such a case is introduce the referent with what can be called a *specific indefinite* expression. Although the intended reference is specific for himself, the speaker nevertheless defers to his listener’s lack of knowledge of the particular referent intended by initial indefinite reference. Let us consider the instance of the boy who had been bitten by a strange dog in the street. It was inappropriate for him to tell his mother immediately ‘*The dog* bit me’. What would be appropriate is an indefinite reference to the dog: ‘*A dog* bit me’, or ‘There was *a dog*, and *The dog* bit me.’ The use of the indefinite reference indicates to his mother that he is referring to a member of the class of dogs not already known to her. When the speaker’s reference is specific for himself but not for his addressee, such an introductory indefinite reference becomes appropriate.

Once a referent in a discourse has become established as a unique member of its class for both the speaker and listener in the discourse,

subsequent references to it should be definite ones, such as *the dog*. Such references are referred to by Karttunen (1968a,b) as *discourse referents*: a referent that is to be referred to specifically in the discourse for both the speaker and the listener.

### 1.8. Specificity by Entailment

The kind of specificity involved in a conversation about absent referents differs sharply from the perceptual specificity provided in a conversation about physically present referents, especially given the different means of establishing discourse referents in each. But the two kinds of situations are closely linked. The case of *entailment*, discussed by Karttunen (1968b), provides clear conceptual bridge between the two types of context. The workings of entailment derive from the fact that simply mentioning some referents or situations necessarily entails the existence of the other, immediately specified referents, which can themselves become discourse referents. Karttunen exemplifies the workings of entailment in a hypothetical discourse:

‘I was driving on the freeway some other day when suddenly *the engine* began to make a funny noise. I stopped *the car* and when I opened *the hood*, I saw that *the radiator* was boiling’. (Karttunen, 1968b:10)

The speaker did not have to introduce the car, hood and radiator as discourse referents by means of introductory specific indefinite expressions, e.g. ‘I was driving *a car*. It has *a hood* and *a radiator*.’ All of the italicized expressions are properly definite without such introduction because driving on a freeway entails the existence of a particular car that was driven. In turn the existence of the car entails a hood, radiator, and engine belonging to the car. Speakers use words when conversing about absent referents to construct situations. Discourse referents can be created or prepared in conversations without the use of overt verbal introduction and treated referentially much like those in physically present contexts. The rules of definite and indefinite reference apply similarly through different kinds of discourse, with the addition that specific indefinite expressions may be necessary to introduce a referent to the listener when no other means suffices (Maratsos, 1976: 5).

### 1.9. The Semantics of Articles

A child learning the use of definite and indefinite articles must formulate a semantic system both abstract and sensitive to discourse

variables such as his listener's likely knowledge of particular referents. The categories of references, mentioned above, are summarized in Table 2, which has been adapted from Brown (1973) with different examples.

**Table 2. The relation between definite and non-definite forms and specific and non-specific reference in speaker and listener**

	Speaker specific	Speaker non-specific
Listener specific	Definite: <i>the</i> Where should we put <i>the table</i> ?  <i>The engine</i> began to make a funny noise.	Null?
Listener non-specific	<i>A dog</i> bit me.  There's <i>a table</i> over here.	Draw <i>a horse</i> .  I haven't got <i>a car</i> .

The upper left-hand quadrant corresponds to instances in which the speaker has in mind a particular member of the class and is confident that his listener will be able to understand the expression he uses as referring to just the same unique member or the class.

A reference such as *the dog* is appropriate only when specific in this way both for the speaker and the listener. The lower left-hand quadrant corresponds to the case in which the reference is specific for the speaker but not for his listener. In this divergence of viewpoints, the speaker must defer to his listener's lack of knowledge and refer with an indefinite expression. The lower right-hand quadrant exemplifies the case in which reference is specific for neither speaker nor listener: any member of the class may be intended, as in 'Give me *a short dress*', or the referent may be non-existent, as in 'I haven't got *a car*'.

## 1.10. Naming

Children must learn the definite-indefinite referential system. One of the problematic instances is the case of *naming* or *nomination*, one of the essential operations of reference. Brown (1973) depicts the problem clearly:

‘When pointing and naming something new, a thing both parents and children often do, one says **That’s a train** or **That’s a bear** and then goes on to use the definite forms: **it** or **the train** or **the bear**. Why does the introductory sentence use a non-definite form? Nominatives of this sort are used in situations in which both speaker and listener are attending to the same referent, and in addition, the speaker is likely to be pointing at it. I.e., it seems as though reference should be in the definite (p. 347)’

Brown believes that such instances fall into the category of references specific for the speaker but not the listener. The speaker pointing out the name knows that it applies to the conspicuous train or bear, but the listener presumably does not. Hence the reference of *the train* or *the bear* would not be clear to the listener even though the object is apparent.

This argument has some interest, but there is difficulty for it with formally similar statements in which both the speaker and the listener are aware of the class membership and the reference is still indefinite: such as the sentence ‘It is, after all, only *a bear*’, or ‘The fact that it is *a bear* should not affect us.’ In such cases it is no longer possible to argue that the speaker and listener share differing knowledge about the appropriateness of the reference for the nominated object; nevertheless the reference remains indefinite. (Maratsos, 1976: 7)

If nominal indefinite expressions fit into the categories of Table 2, they probably do so as non-specific references. In the act of naming or attributing further characteristics to an object, a speaker is concerned only with placing the named object in its relationship to the rest of the members of the named class. In the example above of a parent naming a bear for his child, he is not concerned with it as a particular bear, but only as a non-particular member of the class of bears. It is contrastively possible for nominal expressions to acquire specific status as members of their class. Some gain this unique status by virtue of definition, e.g. ‘*This is the biggest bear in the world*’ (there is only one such bear). In other cases the situation may provide the context, for example ‘One of these is a bear and the other one is a raccoon. That one is *the bear*.’ The possibility should not be overlooked, however, that nomination and attribution statements do not fall properly into the categories discussed above at all, though there is much overlap of meaning.

### 1.11. The Conceptual Basis for Definite and Indefinite Reference

The general problem facing the child in making semantic acquisition is a complex and challenging one. His task is to discover the proper

situational uses of various phonological segments. According to Maratsos (1976), his data for the task are the grammatical and semantic knowledge he may have of other sound segments used around the target segments and the non-linguistic context in which he hears all these segments used. In the case of articles the segments are phonologically slight ones, *a* (or *an* in front of words beginning with vowels, as in *an elephant*) and *the*. Their meanings are abstruse. They refer to no particular object, class of objects, or class of actions, as do, e.g. *mommy*, *dog*, or *push*, or even a consistent internal feeling such as is nominated by *want*. Their meaning inheres in the semantically abstract notions of specificity of reference and the specificity of a reference for their listener. Each of these presents what abstractly seem like severe problems of conceptualization. It is worth considering here each of the conceptual bases of definite-indefinite reference to attempt greater insight into the child's problem.

### 1.12. Understanding Classes and Class Membership

The system of specificity is the abstract system of classes, class membership, the relation of class members to other class members as well as simply the notion of any class member. The basis for specificity cannot be found in particular objects or external physical attributes.

A specific reference rests on the cognitive notion of unique member of a class. A specific reference is a reference to some member of the class nominated which has all the attributes required to be a member of that class plus others which make it distinctive. In situations where the referent is physically present such distinctive properties may be perceptual ones. Each object has its particular individual physical characteristics, and a certain individuality attained by the occupation of a unique spatio-temporal segment. Referents introduced to the speaker only verbally have as distinctive properties only the propositional context in which they were introduced. When a speaker hears someone say 'I was walking down the street and a dog growled at me', the speaker has access to *dog* as a specific one only through the information 'growled at X when X was walking down the street.' Such discourse referents may be even more abstract in nature, both short-lived and highly hypothetical. In the cases discussed so far, expressions such as *the dog* have referred to a real dog unique in the class of dogs. But discourse referents may be introduced only hypothetically, in statements such as 'I wish we had *a dog* and *a ball* (non-specific references). We could throw *the ball* to *the dog* (specific references).' Referents hypothetically introduced in this manner exist only

in the hypothetical mode of the discourse. It would be peculiar to continue afterwards ‘Let’s see how big the dog is’, or ‘I wonder where the dog is’.

Non-specific reference contrastively rests on the idea of any member of a class. The class member may be presumed to be existent, as when a speaker says ‘I’d like to take out a book’ at the library. Its existence may be only variably perspective, as in ‘Let’s buy a car’, or ‘Let’s have a baby’. Or only the bare notion of any class member at all may remain, as in references to no member at all, e.g. ‘I can’t drive a car’.

The child must not have only developed awareness of such differences in types of class members and class membership. For use in the linguistic system he must operate recursively on his own awareness and monitor it for use in the verbal system. This is evidently true for any instance of verbal expression. But the cognitive dimension operated on seems peculiarly abstract, defined neither on particular perceptions or classes of perceptions, either external or internal.

The hopes for children’s early acquisition of the ability to operate consistently with such a semantic system should be dim. Piaget (1962) concluded that children between the ages of two and four have great difficulties with formulating the relations between individual class members and the generic notion of the class:

‘We find one constant characteristic of the **‘preconcepts’** of this age which seems to be decisive. The child at this state achieves neither true generality nor true individuality, the notions he uses fluctuating incessantly between the two extremes (Piaget, 1962: 224)’

Piaget produced a famous observation of his daughter Jacqueline, aged two and a half (2;6) at the time to support his claim:

But also at 2;6 she used the term **‘the slug’** for the slugs we went to see every morning along a certain road. At 2;7 she cried: ‘There it is’ on seeing one, and when she saw another ten yards further on she said: ‘There’s the slug again’. I answered: ‘But isn’t it another one?’ Jacqueline went back to see the first one. ‘Is it the same one? – **Yes** – Another or the same? – ...?’ The question obviously had no meaning for Jacqueline (p. 225).

He quotes a similar observation of Jacqueline at 3;3:

Jacqueline was playing with a red insect, which disappeared. A quarter of an hour later when we were out for a walk we tried to look at a lizard, which darted away. Ten minutes later afterwards we found another red insect. ‘It’s the red animal again’. (Piaget, 1962:225)

These and other observations convinced Piaget that the young child does not differentiate individual members of a class clearly from one another, nor from the general class to which they belong.

Other work also suggests that the cognitive dimensions presupposed by specific and non-specific reference may give problems to the young child. Bruner *et al.* (1966) hold that the cognitive competence of the preschool child is limited in a serious way by the child's greater dependence on iconic representation, the use of the perceptual imagery to represent the world. Imagic representation seems peculiarly ill-suited for representation of the difference between particular and non-particular, individual and general. Should a child translate expressions like *a dog* as used in 'I don't want *a dog*' or 'Let's get *a dog*' into imagic terms alone, he would have nothing but problems in distinguishing this non-particular dog from individual, particular dogs. Mature referential ability cannot depend heavily on iconic representation. If the young child's representations are heavily laden with imagery and only lightly based on amodal, abstract representations, the conceptual basis underlying specific and non-specific reference can only cause difficulty.

### 1.13. Summary

According to Brown (1973) young children frequently fail in their initial use of articles to distinguish between those occasions on which their listener's knowledge converges with their own and those which it does not. Adam, Eve and Sarah's, the children described above, definite references frequently failed to elicit recognition of the intended referent in their listeners, and instances of introductory, specific indefinite references were few and limited.

In Piaget's terminology, such children fail to 'decenter' from their own viewpoint in their use of articles and so are egocentric. Even children who have used articles for some time have continued varied difficulties in their non-egocentric use. Although skill in employing articles non-egocentrically improves, we shall find that unlike the case for specificity of reference, it is more difficult to mark the acquisition of a single, apparently more unified non-egocentric competence.





	<b>Niejaki</b> <b>Pewien</b>
3) prepositions, e.g. two pence <b>a</b> kilo	dwa pency <b>za</b> kilo

However, the definite article has at least five counterparts in Polish:

1) zero article, e.g. <b>The</b> Polish that live... <b>The</b> Alps	Polacy, którzy mieszkają... Alpy
2) pronouns, e.g. <b>The</b> man there <b>The</b> fellows	<b>Ten</b> człowiek tam <b>Ci</b> faceci
3) adjectives, e.g. <b>The</b> Nowak	<b>Słynny</b> Nowak
4) prepositions, e.g. 80 pence <b>the</b> yard	80 pensów <b>za</b> jard
5) word order, e.g. <b>The</b> woman looked at him. <b>A</b> woman looked at him.	Kobieta spojrziała na niego. Spojrziała na niego kobieta.

Each article is rendered by many different types of structures and this diversity is responsible for the difficulty in acquiring the article as a grammatical structure.

As Arabski (1990) claims (after doing research with Polish first year university students at the Faculty of Philology), that in the process of learning or acquisition of English article system by Polish learners, it is not associated with one structure as in the case of, e.g. **book – książka**, which is associated with a Polish noun. Syntactically both stay within one class. The problem gets more serious when a given English structure is associated with two or more Polish items of the same class.

Article errors are apparently the most common because they appear in English more often than other structures. Their misuse is more apparent and more striking than in the case of other structures which are equally unlearned. In a free conversation or a composition, article errors among Polish learners strike us as very common ones. They are common proportionally to their frequency of occurrence in a normal native speaker's production.

English articles are very often redundant structures and like any other redundant elements are omitted in the process of acquisition at least at the beginning and intermediate levels.

The difficulty of acquiring English articles by Polish learners has at least three aspects:

- 1) the differences between L1 and L2,
- 2) the inherent L2 difficulty,
- 3) statistical status of the structure.

Following the opinions expressed above, Szwedek (1976) claims that indefinite and demonstrative pronouns are the most likely candidates to function in the way parallel to the English articles.

Pisarek (1968) wrote that *'in Polish, where there are no articles nor explicit definiteness or the lack of it realized by morphological features, there are no pronouns which can be called definite. The opposition of definiteness to indefiniteness is expressed in a specific way: for example, by the opposition of a given pronoun to the lack of the pronoun. This is how I understand the sense of grammatical definiteness in Polish...'*

There is a history of contrastive studies (English vs. Persian in Jafarpur (1979); English vs. Slavic in Kaluza (1963)) and of pedagogically-oriented analyses of the article system (Grannis, 1972; Hok, 1970; McEldowney, 1977). There exist descriptions of the use of English articles by learners sharing a given L1 (Agnihotri, Khanna, & Mukherjee (1984) for Hindi and Punjabi; Kharma (1981) for Arabic; Yamada & Matsuura (1982) for Japanese)). These studies are often based on data from cloze tests, where subjects insert articles into a written text. Unfortunately, there are risks related to the interpretation of cloze tests, and in any case, data gathered in this way give an inaccurate view of how learners actually use articles (Thomas, 1988).

There are relatively few studies of the L2 acquisition of articles based on naturalistic data. Hakuta's (1975, 1976) longitudinal study of a single 5-year-old Japanese speaker learning English includes data on the development of *a* and *the*, although he does not analyze *zero* article.

Huebner's (1983, 1985) analysis of the acquisition of English by one adult Hmong speaker provides the most in-depth longitudinal study of L2 acquisition of the definite article. However, Huebner's concern lies not so much in tracing the development of the native-speaker system of article usage than in describing the systematicity inherent in his subject's interlanguage. This makes his results difficult to compare with more conventional work.

Parrish (1987) follows Huebner in looking at the acquisition of articles by a single L2 learner, in this instance a native speaker of Japanese, over a period of 4 months.

Master (1987) performed a pseudo-longitudinal study of the use of articles in spontaneous speech. His subjects were 20 L2 learners comprising one subject at each of four developmental levels, across five L1 groups, two of which have formal equivalents of English articles

(German and Spanish), and three of which do not (Japanese, Chinese, and Russian).

From these sources, some tentative generalizations emerge about the development of articles in the speech of L2 learners.

Master reports that *zero* article dominates in all environments for articles in the early stages of L2 acquisition, at least for learners whose native languages lack articles (which is highly consistent with Arabski's research, described above, and his identical conclusions that articles are omitted in the process of acquisition at least at the beginning and intermediate levels). The first context in which overt articles are used with consistent appropriateness is [+SR +HK]. This seems to be true of all Master's subjects, whether their L1 includes articles or not.

The first sample of Huebner's (1983) subject's speech shows the definite article in 64% of [+SR +HK] contexts, with this figure rising to around 88% by the 18<sup>th</sup> week of observation.

Hakuta's (1976) data are less exact, but he reports that his subject performs better with *the* than with *a*, measuring the number of correct uses over total usage.

The emergence of *a* in its appropriate environments is later and more gradual, according to Master. Parrish also finds that her subject's accurate use of *a* is delayed relative to *the*. Huebner (1985) observes that during his original year-long study, '*a* appeared to be a phonological variant of *zero* or a hesitation phenomenon'. But by the time of a follow-up study performed 20 weeks after the end of the first study (in the equivalent of the 71<sup>st</sup> week of observation), his subject was using *a* appropriately in both [+SR -HK] and [-SR -HK] contexts.

Assuming that this course of development differs somewhat from that proposed for L1 learners, it is significant that Huebner and Master both find L2 learners overgeneralizing *the*. Huebner (1983) calls this *the*-*'flooding'* and reports that by the 6<sup>th</sup> week of observation his subject marked 90% of *all* nouns with the definite article. The flood gradually receded, with the definite article disappearing first from [-SR -HK] contexts, but persevering longer in [+SR -HK] contexts. His subject continued to use the definite article virtually exclusively in [+SR +HK] and [-SR +HK] contexts at the end of the first year. Master observes surges in the use of *the* among subjects in the first two of his four proficiency levels, but only for speakers whose languages lack formal equivalents of articles. Except for a single subject who produced *the* in front of all nouns, there was no evidence of *the* flooding into [-SR -HK] contexts. Master claims that *the* floods mostly into [+HK] environments. But paradoxically, in first-mention [+SR -HK] contexts, there are higher

levels of *the* than of *a* in the production of half of his level 3 and 4 subjects whose L1s lack articles. One problem in noticing what is and is not a flood is that neither Huebner nor Master defines the term, except loosely as ‘a dramatic rise in usage’.

Summing up generally the report of the results of the research concerning L2 acquisition of English articles by foreign learners, we may conclude that *the* emerges early and *a* later in L2 acquisition. Huebner (1983) speculates that after an early flooding of the definite article across all environments, his subject associated *the* with the feature [+HK], thus accounting for its sustained use with first-mention and non-referential nouns. Because Master finds *the* dominating in [+SR +HK] and [-SR +HK] contexts among learners whose L1 lacks articles, he agrees that they may associate *the* with the feature [+HK]. However, this leaves unexplained the evidence that L1 learners overproduce *the* in [+SR -HK] environments. Parrish, for instance, reports that her subject’s only instances of inappropriate use of *the* occurred in [+SR -HK] contexts.

## 2.2. Comparing L1 and L2 Acquisition Data

We have observed parallel claims that L1 and L2 learners overgeneralize the definite article. Several studies concur that L1 learners do so in [+SR -HK] first-mention environments. We must remember that this has been invariably attributed to children’s egocentricity (see Chapter One) or to association of the definite article with the feature [+SR], since *the* also appears early in [+SR +HK] contexts. The facts of L2 acquisition are less clear. Parrish’s (1987) results are consistent with the facts of L1 acquisition, as she finds *the* being overproduced only in referential indefinite contexts. Huebner’s subject seems first to produce *the* everywhere, then to cut back on its appearance in [-HK] contexts. Master suggests that some L2 learners first associate *the* with [+HK], leading them to flood [+SR +HK] and [-SR +HK] contexts as a result.

Huebner and Master conclude that L2 learners associate *the* with the feature [+HK], rather than with [+SR]. Their evidence is based on relatively high rates of *the* in [-SR +HK] (generic) contexts, in addition to the high rates of *the* in [+SR +HK] contexts. However, this observation needs examination. The first problem is the deficit of data concerning article use with generic nouns. Huebner and Master both indicate that generics are rare in the production of L2 learners. Huebner (1983) tallies only 27 [-SR +HK] environments in a full year of observation (cf. 377, 441 and 1,613 instances of the other three environments). Master (1987) finds generic *a* and *the* contexts so infrequent that he excludes them from

his analysis, and 8 out of 20 subjects produce fewer than five instances of generic *zero* contexts. Unfortunately, he does not specify how many generic *zero* contexts do appear, but the total number may be so insignificant as to make generalizations about article use in this context unreliable. A second problem is the difficulty of defining [-SR +HK] contexts. Master's examples include *he saw his brother going to Ø school* which might be analysed as an idiom rather than a generic, and *to call a virus a live thing*, potentially as [-SR -HK] non-referential context. Huebner neither defines what makes a noun 'generic' nor gives examples. Without consistent standards, claims about how L2 learners use articles in generic contexts are hard to evaluate. A third problem is that data from Parrish are inconsistent with the idea that L2 learners associate *the* with the feature [+HK]. Her subject produced only four tokens of *the* in 42 [-SR +HK] contexts, all of which were appropriate uses (we must remember that generic contexts license *a*, *the* or  $\emptyset$ , depending on the lexical features of the noun). Moreover, we lack data about how children mark generic contexts, so we cannot compare L1 learners with L2 learners.

That is why it is advisable to put off the data on generic nouns and compare how L1 and L2 learners use articles in other contexts (see Chapter Four – the empirical study). We see from the early stages of acquisition that both groups use *the* appropriately in [+SR +HK] environments. In addition, both L1 learners and L2 learners whose native languages lack articles produce more instances of *the* in [+SR -HK] contexts than in [-SR -HK] contexts. The data on L2 acquisition are insufficient, but if these generalizations are supportive, it would mean that:

1. L2 as well as L1 learners associate the definite article with contexts sharing the feature [+SR], rather than with [+HK] and therefore,
2. the claim that children's use of *the* in [+SR -HK] contexts is due to egocentrism (see Chapter One) becomes less tenable. Adult L2 learners are unlikely to be influenced by egocentrism, and yet they seem to overuse *the* in ways similar to L1 learners. On the other hand, Bickerton's suggestion that language learners are sensitive to the specificity of nouns may be extended to L2 learners.

Concluding the above considerations we may state that these hypotheses are based on rather slim evidence that L2 learners do overgeneralize *the* into [+SR -HK] contexts, but do not do so in [-SR -HK] contexts. To gather more data on the use of articles in generic environments as well as in other contexts, the author conducted the empirical research in three different proficiency groups of Polish L2