

Vocabulary Pedagogy:
A Wittgensteinian Approach

Emily Comer

March 28, 2014

Table of Contents

1. Introduction	3
2. Wittgenstein's Theory of Vocabulary Pedagogy	5
2.1 Introduction	5
2.2 The argument against definitions	6
2.3 The argument for examples	12
2.4 Methods for using examples	14
2.5 Philosophical conclusion	16
3. Empirical Support for a Wittgensteinian Vocabulary Pedagogy	17
3.1 Introduction	17
3.2 Empirical support for family resemblances	17
3.3 Empirical support for blurred edges and contextual influence	20
3.4 Meta-analysis of vocabulary instruction	24
3.5 Empirical support for incidental word learning	28
3.6 Empirical conclusion	29
4. A Wittgensteinian Analysis of Words, Words, Words	30
4.1 Introduction	30
4.2 Integration, repetition, and meaningful use	31
4.3 Incidental acquisition: extensive reading	32
4.4 Explicit instruction: tools and techniques	34
4.5 Assessment: a holistic approach	37
4.6 Conclusion to the book analysis	38
5. Conclusion	38
5.1 Lessons for philosophy	39
5.2 Lessons for education	40

1. Introduction

Vocabulary is vital: the more words one knows, the better. The number of words a person can understand and use has far-reaching influence in many areas, such as his or her abilities to understand complex ideas, to express himself or herself, and to learn almost anything about almost any subject. More specifically, research indicates that “vocabulary knowledge is related to and affects [reading] comprehension. The relationship between word knowledge and comprehension is unequivocal” (Davis 1944, 1968, quoted in Allen 1999: 5). With respect to students, “[c]hildren who know more words understand text better” (Nagy & Herman 1987: 27).¹ The ability to comprehend what one reads is universally regarded as necessary for success in school from about third grade onwards (see the National Reading Panel (2000)). Therefore, we should strive for vocabulary pedagogy² that increases students’ abilities to understand and use words. We should incorporate useful insights from as many relevant perspectives and areas of study as possible. Accordingly, this thesis takes an interdisciplinary approach to vocabulary pedagogy, combining philosophical reasoning with the findings of empirical research in psychology and education.

Despite the importance of increasing students’ word knowledge, classroom-based vocabulary pedagogy often falls short of this goal. Janet Allen, whose book *Words, Words, Words: Teaching Vocabulary in Grades 4-12* I will discuss in Section Four, reflects on experiences common to many teachers:

¹ For a discussion of the underlying causality, see Nagy and Herman 1987: 28.

² In this thesis I use ‘instruction’ to refer to specific, explicit vocabulary-teaching practices such as copying dictionary definitions and making graphic organizers of word meanings. ‘Pedagogy’ refers to a cumulative approach to imparting word knowledge. Thus, it includes both explicit techniques that fall under ‘instruction’ and, when they are used, implicit strategies such as extensive reading and discussions about words and concepts.

Dictionaries and programmed vocabulary books have been the mainstay of vocabulary instruction in language arts classrooms for many years. I spent most of my career telling students, “Look it up in the dictionary,” when they asked me what a word meant. I handed out lists of words and had students copy definitions and write the words in sentences. Still they didn’t know the words. They asked me which definition to copy from the dictionary. I told them to copy the one that made sense, the one that fit the context. They looked at me as if I were an alien and asked, “Can we copy the shortest one?” None of the definitions made sense to them. Often they didn’t even understand the words used in the definitions... They had not internalized a meaning during our reading. At best, they knew only the meaning they copied. Often they didn’t even know that meaning. (Allen 1999: 33)

I myself have seen the prevalence of the definition-copying approach firsthand, both in the United States and in the United Kingdom, and across multiple grades. In the schools where I volunteered in Rockbridge County, Virginia, and Oxford, England, I have watched students bent over dictionaries, laboriously copying every definition of the assigned word – even when one or more of those definitions clearly did not fit the original context. This exercise is sufficiently boring to bring tears to the eyes of students and teacher alike.

My claim in this thesis is that we do not need to rely solely on studying dictionary definitions as the exclusive means of learning new words. Rather, there are far more powerful ways to increase students’ vocabularies – including contextual examples, conversations about words, wide reading, and both dictionary and student-written definitions. To argue for this more holistic vocabulary pedagogy, I first consider the word-learning theory of prominent twentieth-century philosopher Ludwig Wittgenstein and draw several implications for vocabulary pedagogy from his work *Philosophical Investigations*. Then, using those implications as a framework, I examine several empirical studies that tested word knowledge and word learning. Next, I review Allen’s well-regarded teacher resource book, *Words, Words, Words*, as an example of powerful, holistic

vocabulary pedagogy. In accordance with the interdisciplinary nature of this thesis, I conclude with lessons for philosophy and lessons for education.

2. Wittgenstein's Theory of Vocabulary Pedagogy

2.1 Introduction

My goal in this section is to apply the ideas of the later Wittgenstein in his *Philosophical Investigations* (1953/2009) to classroom-based vocabulary pedagogy. Of course, Wittgenstein himself was concerned less with actual vocabulary pedagogy than with how language learning in general occurs and with what language learning reveals about language itself. For this reason, I sometimes apply his ideas to contexts he did not originally envision. In fact, because his concerns do differ from mine, the distance between the new context in which I am using his ideas and his original context might at times be large enough to raise the question of whether I am still being faithful to the ideas themselves.

However, let me offer two defenses of this methodology. First, one reason for Wittgenstein's continuing influence as a philosopher is that his ideas *can* be applied to contexts that he himself did not imagine. In this, Wittgenstein is no different from other great philosophers. Second, and perhaps more importantly, this thesis is not directly about Wittgenstein's ideas. Instead, it is about a philosophical and practical argument regarding vocabulary pedagogy, an argument *based* on his ideas (again, sometimes applied to a new context). And, though I try to remain faithful to Wittgenstein's intent, my goal is to *use* his ideas, which necessarily means providing my own interpretation of them. I do not enter into exegetical debates concerning the varied potential interpretations of Wittgenstein's intent itself.

In this section, I first argue against overusing dictionary definitions in vocabulary pedagogy. I then argue in favor of examples as pedagogical tools. Finally, in the spirit of giving examples, I discuss several ways that Wittgenstein illustrates using examples to teach word meanings.

2.2 The argument against definitions

The outline of this argument runs thus:

- (a) Definitions delineate necessary and sufficient conditions for word meanings.
- (b) Ordinary language use does not adhere to those necessary and sufficient conditions.
- (c) Therefore, from (a) and (b), studying only definitions is not an effective way to prepare students for ordinary language use.

I will briefly explain (a) before examining (b), which is the main contention of the *Philosophical Investigations*.

The purpose of a formal definition is to specify all and only those features that constitute a word's meaning. Instead of listing every situation where the word could potentially be used (which would be impossible), the definition presents rules for its use.³ If the definition contains every rule *necessary* and only those rules *sufficient* for capturing a word's meaning, then the reader is able to apply those rules to determine whether the word may appropriately be used in a given situation. Glock, a leading commentator on Wittgenstein's philosophy, identifies this as language essentialism: "the view that... the only adequate or legitimate explanation of a word is an analytic definition which lays down necessary and sufficient conditions for its application, entailing that, for example, explanations by reference to examples are inadequate" (1996: 120). Throughout the

³ Wittgenstein argues that a word's use *is* its meaning. Although I am sympathetic to this position, the important point for vocabulary pedagogy is that a word's use at least *expresses* its meaning.

Philosophical Investigations, Wittgenstein implicitly argues that definitions — composed of necessary and sufficient conditions for using a word appropriately — present a false image of language, one that does not correspond to ordinary language use⁴ (that is, the ways in which most people use language in most situations, including many academic situations⁵). I now discuss his reasons for claiming this; after giving each reason, I explain its implications for vocabulary pedagogy.

First, formal definitions distort ordinary language by drawing sharp, rigid distinctions (*Philosophical Investigations* §76) where there are instead “blurred edges” (*PI* §71). Think of cases that *are* ‘games’ (or ‘languages,’ or ‘revolutions,’ or ‘experiments’) and cases that *are not* ‘games’ (etc.). A particular activity might be labeled a ‘game’ or an ‘experiment’ under one set of circumstances and not under another. Similarly, two mature, native English speakers might disagree about whether the same activity in the same circumstances should be called ‘game,’ ‘experiment,’ etc. These and other words are entirely usable with such blurred edges, though we can and sometimes do make their meanings more distinct. “To repeat, we can draw a boundary — for a special purpose. Does it take this to make the concept usable? Not at all!” (*PI* §69) When rigid boundaries do occur, they are usually imposed by some authority (such as a dictionary or a teacher) to prevent confusion in a specialized context (e.g., when a chemistry teacher says that ‘water’ refers to H₂O). However, in ordinary language use, word meanings do not have such rigid

⁴ The controversy surrounding ordinary language philosophy, and the relationship between “metaphysical” and “everyday” uses of words, are not the topic of this thesis. For discussion of these points see Cook 1999.

⁵ It is possible that definitions based on necessary and sufficient conditions are appropriate in certain contexts, including mathematics and some areas of philosophy. However, even if they are, the point for vocabulary pedagogy still stands.

boundaries. When students study dictionary definitions alone, they do not acquire words with blurred edges; so they do not acquire words as they are actually used in ordinary language.

Because definitions analyze word meanings in terms of necessary and sufficient conditions, a dictionary definition is often *more* difficult for students to understand than the word as it is used in context. (I will refer to the latter as the ‘word-in-use.’) Wittgenstein compares using a definition instead of the word it defines to asking for “the broomstick and the brush which is fitted on to it” instead of the broom. “Is [the listener] going to understand the further analysed sentence better? — This sentence [i.e., one asking for the parts of the broom instead of the broom itself], one might say, comes to the same thing as the ordinary one, but in a more roundabout way” (*PI* §60).

Similarly, definitions typically use words that are even more unfamiliar to students than the defined word. For example, Merriam-Webster’s *Word Central*, a student dictionary available online, defines ‘experiment’ as follows: “trial; especially: a procedure or operation carried out under controlled conditions in order to discover something, to test a hypothesis, or to serve as an example.” This definition is intended to use simple vocabulary (such as “serve as an example”), yet it remains unusable if students do not understand what is meant by (at a minimum) the words and phrases ‘trial,’ ‘procedure,’ ‘operation,’ ‘controlled conditions,’ and ‘hypothesis.’ Because definitions are further analyzed forms of word meanings, teachers should only use them when students will understand them better than they will understand the word-in-use.

Wittgenstein’s second reason for contending that definitions misrepresent ordinary language is that definitions present words as independent units, discrete and unconnected bundles of such necessary and sufficient conditions. But words in ordinary language are actually related to

each other by what Wittgenstein calls ‘family resemblances’ (*PI* §67).⁶ The two most common types of family resemblances occur when the same word is used in different contexts and when different words are used in the same context – though other types occur as well.⁷ Wittgenstein primarily addresses the first type, so I begin with it.

Wittgenstein frequently uses the ‘game’ example to illustrate the impossibility of devising accurate necessary and sufficient conditions for words in ordinary language. He observes that the things we call ‘games’ share many features or attributes: rules, players, pieces, winning and losing, entertainment, physical movement, mental skill, luck, etc. (*PI* §66). These shared features result in “a complicated network of similarities overlapping and criss-crossing: similarities in the large and in the small” (*PI* §66), which he compares to “the various resemblances between members of a family” (*PI* §67). Yet despite these overlapping similarities, there is no single feature or set of features that everything called a ‘game’ possesses; thus, Wittgenstein is unable to find “something in common” by which we call them ‘games.’ Of course, as Glock notes,

Wittgenstein does not maintain that games have nothing in common – he refers to them as ‘procedures’, and it is manifest that they are all activities. But this falls short of a definition, since there are many activities which are not games. The claim is that there is no set of conditions which all and only games satisfy, and hence no analytic definition of ‘game’ in terms of necessary and sufficient conditions. (1996: 121)

But the goal of a definition is precisely to set forth “something in common” to every instance of a word, by defining necessary and sufficient conditions for using the word appropriately *in any context*.

⁶ Glock explicitly connects Wittgenstein’s family resemblance concept to his critique of essentialism (1996: 120).

⁷ For instance, Allen (1999) refers to morphological variants (e.g., experiment, experiments, experimenting, experimented) as ‘word families;’ see Section Four of this thesis.

Of course, dictionary entries for words that occur in sufficiently diverse contexts usually give distinct definitions for each meaning. But these separate definitions are only useful for readers who are already able to determine which meaning is contextually appropriate — that is, readers who at least partially understand the word. For word learners, conversely, multiple definitions are especially confusing. Consider *Word Central's* definitions of ‘revolution’:

- 1a: the action by a heavenly body of going round in an orbit; b: the time taken to complete one orbit
- 2: completion of a course (as of years): **CYCLE**
- 3a: the action or motion of revolving; a turning round a center or axis: **ROTATION**; b: a single complete turn (as of a wheel or a phonograph record)
- 4a: a sudden, extreme, or complete change; b: a basic change in government; especially: the overthrow of one government and the substitution of another by the governed

An experienced user of the word ‘revolution’ can easily see that definition 1 comes from natural science, definition 2 from either natural science or social studies, definition 3 from natural science or mathematics, and definition 4 from social studies (although 4a could also come from natural science). But students who are not already familiar with ‘revolution’ and are instructed merely to study the definition have no way of determining which definition they will need in a particular context. (Indeed, the fact that at least two of these definitions could be used in more than one subject illustrates the difficulty of understanding the definitions without substantial context.)

Moreover, definitions cannot effectively capture a second kind of word families: different words that frequently occur together (what Wittgenstein calls “cousins” (*PI* §224)). *Some* cousins may appear in the definition, but even then it will not necessarily be clear to students *how* language users typically combine the words in ordinary language. And the definition alone does not reveal

the word's relationships with words that do not appear in the definition. For instance, *Word*

Central includes the following definitions of 'democracy':

- 1a: government by the people; especially: rule of the majority; b: government in which the supreme power is held by the people and used by them directly or indirectly through representation
- 2: a political unit (as a nation) that has a democratic government
- 3: belief in or practice of the idea that all people are socially equal

Just like the definitions of 'revolution,' these definitions are also subject to the problem of deciphering multiple definitions, discussed above. In addition, they fail to teach students how to use 'democracy' in conjunction with other, related words. They use, but do not explain, the semi-technical meanings of 'people,' 'majority,' and 'equal;' and the concepts of 'vote' and 'elect' are missing entirely.

However, the solution is *not* to add more words to definitions, so that they contain all cousins of the defined word. For one thing, many of a vocabulary word's cousins are likely also to be unfamiliar to students, so adding them to the word's definition would only make it that much more difficult to comprehend. Moreover, needing to know words in order to learn other words is not the same as learning words together. In fact, Wittgenstein observes of such "cousins" that "[i]f I teach anyone the use of the one word, he learns the use of the other with it" (*PI* §224). Students cannot fully acquire 'democracy' without also acquiring 'vote,' 'people,' 'election,' 'majority,' 'equal,' etc. Thus, instead of studying unconnected definitions, they can learn sets of related words as they arise naturally in examples taken from ordinary language.

Finally, Wittgenstein contends that vocabulary pedagogy based solely on definitions misrepresents ordinary language by stripping the word from its natural context — what Wittgenstein

calls “the language in which it is at home” (*PI* §116).⁸ Knowing how to use a word is *not* the same as knowing how to define it: “an aspect of the matter is lost to the latter no less than to the former” (*PI* §63). The different forms – word-in-use and definition – are not interchangeable; we use them for different purposes. Although knowing a word’s definition may be a helpful bridge to using the word, students must be able to do more than just define words if they are to continue using them after the vocabulary test: they must also be able to use vocabulary words in a variety of situations, such as research projects, lab work, one-to-one and whole group discussions, as well as many other non-academic contexts.

For all these reasons, knowledge of definitions should be thought of as a bridge to or preparation for knowledge of ordinary language – not as the final goal of vocabulary pedagogy. In cases where students understand definitions more easily than they understand vocabulary words, they can use definitions to help them acquire new words, gradually abandoning rigid boundaries and unconnected meanings as they become more comfortable with the blurred edges and family resemblances that the words themselves have. Yet each newly-learned definition must be connected to the language as a whole that students are acquiring. The next section suggests that this can best be done using examples of the word-in-use.

2.3 The argument for examples

So far, my discussion has concentrated on Wittgenstein’s *negative* approach to language learning. But he also makes positive arguments regarding vocabulary acquisition and proposes an alternative

⁸ For Wittgenstein, the idea that word meanings can be understood only in the contexts in which they are used is closely connected with the concept of the ‘language-game;’ however, I have chosen not to attempt to do justice to language-games in this thesis. For more on this topic, including the relationship between language-games and rules, see Kenny 1973: 159-177.

word-teaching tool. Wittgenstein poses the following question: “How would we explain to someone what a game is?” (Substitute ‘revolution,’ ‘experiment,’ ‘democracy,’ or any other word students are expected to learn.) His answer is simple yet profound: “I think that we’d describe *games* to him, and we might add to the description: ‘This *and similar things* are called “games”.’ ” (*PI* §69, emphasis original) Teachers can help students understand vocabulary words by describing real-world instances of those words. For example, the American Revolution, the French Revolution, and the Industrial Revolution are all instances of revolutions, so discussing their similarities and differences can help students learn the meaning of ‘revolution.’ Additionally, students can learn word meanings through exposure to examples of the word-in-use.

Describing examples might appear to be little different from simply defining — both involve explaining one word in terms of others. Nevertheless, examples (see also *PI* §135) are effective in ways that definitions are not. First, in contrast to definitions, examples do not require students to comprehend an *additional* layer of abstraction: the rules for the word’s meaning. By providing and discussing many examples of the word, teachers enable students to discover its meaning through direct exposure to and practice with the word itself, without the intermediary of a definition.

Moreover, a vocabulary pedagogy that incorporates examples in addition to definitions does not exclusively rely on the futile attempt to devise necessary and sufficient conditions that exhaustively cover the word’s meaning. Definitions brush over the idiosyncrasies of specific instances of a word, claiming to give students “something in common” (*PI* §65) that purportedly can be captured in a set of necessary and sufficient conditions and employed for determining where and how the word may be used appropriately. Examples, on the other hand, reveal that instances of a word’s use may differ in many features; thus, they are better able to show students

the family resemblances among the word's different uses. In addition, acknowledging that "[t]his *and similar things* are called [by the word]" (*PI* §69, emphasis original) alerts students that the word encompasses cases beyond the given examples, including ones the teacher chooses not to mention or does not know about. Students can even extend the word's usage themselves. In contrast, a definition purports to cover all possible applications of the word; in effect, a definition discourages students from using the word creatively and experimentally and thereby making it their own.

Finally, examples are taken directly from "the language in which [the word] is at home" (*PI* §116), so they present the word as it is actually used in ordinary language. Through examining the word as it appears in contexts taken from ordinary language, students learn not only *that* word, but also many others that commonly occur with that word (see *PI* §224). For these reasons, examples are often more effective tools than formal definitions for teaching students how words are used in ordinary language.

2.4 Methods for using examples

Wittgenstein not only presents arguments in favor of examples as ways of understanding meaning, which can be extended into arguments for using examples as tools in vocabulary pedagogy; he also illustrates several ways of using examples for this purpose. He considers how he might try to get someone to understand what is called 'yellow ochre':

Suppose I show someone various multicoloured pictures, and say: 'The colour you see in all these is called "yellow ochre".' — This is an explanation that another person will come to understand by looking for, and seeing, what is common to the pictures. Then he can look at, can point to, the common feature.

Compare with this a case in which I show him figures of different shapes, all painted the same colour, and say: 'What these have in common is called 'yellow ochre'.'

And compare this case: I show him samples of different shades of blue, and say: ‘The colour that is common to all these is what I call “blue”.’ (PI §72)

Each of the three methods mentioned by Wittgenstein is useful for a different piece of vocabulary pedagogy. The first enables students to learn one meaning of a word. Unlike a dictionary entry, this exercise depends on many, varied examples: sentences and paragraphs that students can understand sufficiently without knowing the target word. The emphasis, though, is on *variety* (“multicoloured pictures”). Students compare and understand the different contexts in which the word appears – but, as Baker and Hacker point out (1980: 360-362), they may fully understand those contexts without “seeing what is common” to them all. Through the combination of these processes, in the context of a discussion with each other and their teacher, students discover the contribution of the unknown word.

However, if examples for the first method are too similar, that method gives way to the second; without sufficiently diverse input, students cannot see the whole meaning of the word. Yet the second method too can be useful, particularly for discussing morphology – that is, grammatical changes in a word’s form such as plural or past tense markings. Students examine different variants of a word (e.g., ‘experiment,’ ‘experiments,’ ‘experimenting,’ etc.), noting that the core meaning remains the same even as its form varies. They also consider which forms are used in which contexts, and for which purposes.

The teacher or students might use the third method by pointing out that ‘revolution’ is used both in social studies (e.g., the French Revolution, the Industrial Revolution) and in science (e.g., the revolutions of a wheel, the revolution of the earth around the sun). They could then analyze the commonalities (or “affinities”) among uses of the word, such as turning and change.

They could also try to determine what is different about each situation (e.g., what is turning or changing, how fast, how often, etc.). This conversation, containing numerous, varied examples and back-and-forth between teacher and students, would increase the students' understanding of each use of the word, both strengthening the connections and clarifying the distinctions.

As this passage from the *Philosophical Investigations* illustrates, vocabulary pedagogy can use examples in numerous ways. In Section Four, when I analyze Allen's book *Words, Words, Words*, I will discuss other uses of examples, as well as ways that examples and definitions can be combined. To help students learn new words, teachers must skillfully choose the most effective tools for each word, situation, and group of learners.

2.5 Philosophical conclusion

Through analyzing Wittgenstein's *Philosophical Investigations*, I have shown several reasons why studying definitions alone is a less effective vocabulary pedagogy. Definitions present discrete sets of necessary and sufficient conditions for word meanings. Moreover, definitions are essentially independent of one another and any supporting context. In contrast, word meanings in ordinary language have "blurred edges" (*PI* §71). They are related to the contexts in which they are used (*PI* §49) and to one another through "family resemblances" (*PI* §67).

Based on Wittgenstein's theory of language, I argue that definitional knowledge itself should not be the goal of vocabulary pedagogy, but should be one of many tools used to teach new words. I also argue that an effective way to acquire words is through conversations about examples. When students have opportunities to notice how words are used in ordinary language and to practice using those words themselves, they learn to use new words "in the language[s] in which [they] are at home" (*PI* §116).

3. Empirical Support for a Wittgensteinian Vocabulary Pedagogy

3.1 Introduction

In this section I show that Wittgenstein's theory of vocabulary pedagogy is supported by empirical evidence from disciplines such as psycholinguistics and education. Wittgenstein's view of language is opposed to what psycholinguist Labov calls the "categorical" view, which assumes that word meanings are defined by necessary and sufficient conditions or essential properties:

[According to the categorical view] there is a set of properties associated with the unit [i.e., the word] which are in some way criterial or necessary, essential as opposed to other properties which are unnecessary, accidental, or redundant, and... all of these essential properties must be present for the category to be recognized. (Labov 1973: 342ff)

Numerous empirical studies have discredited this view and supported a Wittgensteinian theory of word meanings based on family resemblances and blurred edges. Accordingly, I first consider the seminal experiments conducted by prominent cognitive psychologists Rosch and Mervis (1975) on the structure of concepts and words. I then consider the experiments of psycholinguist Labov (1973) on blurred edges in word use. Finally, I consider research in education by examining a meta-analysis of studies on effective vocabulary instruction and a study of incidental vocabulary growth.

3.2 Empirical support for family resemblances

Rosch and Mervis (1975) conducted a series of experiments to test whether, in contrast to the categorical view described above, concepts do in fact exhibit family resemblance structures; of these experiments, the first two are most relevant for supporting the argument of this thesis.

Experiments 1 and 2 focus on items taken from six superordinate semantic categories:⁹

FURNITURE, VEHICLE, FRUIT, WEAPON, VEGETABLE, and CLOTHING. The level of abstraction for these superordinate semantic categories is most similar to that of the concepts and vocabulary words that are commonly taught in classrooms, so I will focus on these two experiments.

Experiments 3 through 6 followed a slightly modified procedure for two other types of categories, basic (i.e., CHAIR and CAR) and artificial (nonsense strings of letters created by the researchers).

Although also derived from family resemblance theory, these other experiments tested related issues in cognitive psychology (such as level of categorization) that are not directly significant for my purposes.

The researchers explicitly trace the theoretical foundation for their experiments to the later Wittgenstein's theory of family resemblances:

Wittgenstein (1953) argued that the referents of a word need not have common elements in order for the word to be understood and used in the normal functioning of language. He suggested that, rather, a family resemblance might be what linked the various referents of a word. A family resemblance relationship consists of a set of items of the form AB, BC, CD, DE. That is, each item has at least one, and probably several, elements in common with one or more other items, but no, or few, elements are common to all items. (1975: 574-575)

From this view of family resemblances Rosch and Mervis derive two hypotheses: Hypothesis 1 (tested in Experiment 1), "that a measure of the degree to which an item bore a family resemblance to other members of the category would prove significantly correlated with previously obtained prototypicality ratings of the members of the category" (1975: 577), and Hypothesis 2 (tested in Experiment 2), "that the more prototypical a member of a superordinate category, the

⁹ Semantic categories may be divided into three levels: basic (e.g. CHAIR, CAR), containing familiar, everyday objects; subordinate (e.g. ARMCHAIR, SPORTS CAR), containing more specific instances of basic categories; and superordinate (e.g., FURNITURE, VEHICLE), containing groups of related basic categories.

less dominant its membership would prove to be in categories other than the superordinate in question” (1975: 585). In other words, they measured how subjects categorized the same concepts across multiple tasks to gain converging evidence for family resemblances.

Before performing the experiments, Rosch and Mervis gathered data for how subjects rated (on a 7-point scale) the prototypicality or ‘goodness-of-example’ of fifty to sixty items taken from the six common superordinate semantic categories listed above (Rosch 1975a, cited in Rosch & Mervis 1975). For instance, in the category WEAPON, ‘gun’ was viewed as a very good example of the category, and so received a high prototypicality rating. ‘Bow and arrow’ is still clearly a WEAPON, but was considered not quite as prototypical as ‘gun.’ ‘Poison’ was on the border between WEAPON and non-WEAPON, so received a rather low prototypicality rating. And ‘screwdriver’ was not seen as a WEAPON at all (though the researchers included it for contrast).

In Experiment 1, 400 undergraduate subjects listed attributes for six of these same items, one from each superordinate category; each item was rated by twenty subjects. These attributes were then scored according to how many items they were attributed to, and the items were scored according to how common their attributes were. Based on these scores, the researchers developed a picture of the conceptual structure of the category. The prevalence of common but not universal attributes indicated that the category had a family resemblance structure. In addition, items that had many attributes in common with other items were also those with high prototypicality ratings (correlations ranged from 0.84 to 0.94, $p < .001$), as expected (Hypothesis 1).

In Experiment 2, 400 undergraduates who had not participated in Experiment 1 were shown the same items as in Experiment 1, but were asked to list categories that those items belonged to. Based on these categories, the researchers determined the category dominance of

each item, which they used as an indirect measure for salient membership in another category – i.e., if an item had high category dominance in one particular category, they concluded that it had low dominance in other, similar categories. High category dominance correlated strongly with high prototypicality ratings (correlations ranged from 0.67 to 0.83, $p < .001$), as predicted (Hypothesis 2).

Rosch and Mervis found that “the majority of attributes listed for items in the six [superordinate semantic] categories demonstrated a family resemblance relationship; that is, they were common to only some of the category members” (1975: 581). Moreover, both the hypotheses were strongly supported: more prototypical category members had greater commonality (family resemblance) with other members (Hypothesis 1), and were less strongly located in other categories (Hypothesis 2). This research laid the foundation for further research on prototypicality. Although the picture has grown more nuanced and complex since Rosch and Mervis’ original experiments, the basic assumptions of their research continue to be supported.¹⁰ As Rosch and Mervis claim (1975: 603), their experiments provide empirical support in favor of Wittgenstein’s family resemblance theory and against the categorical view of words and concepts.

3.3 Empirical support for blurred edges and contextual influence

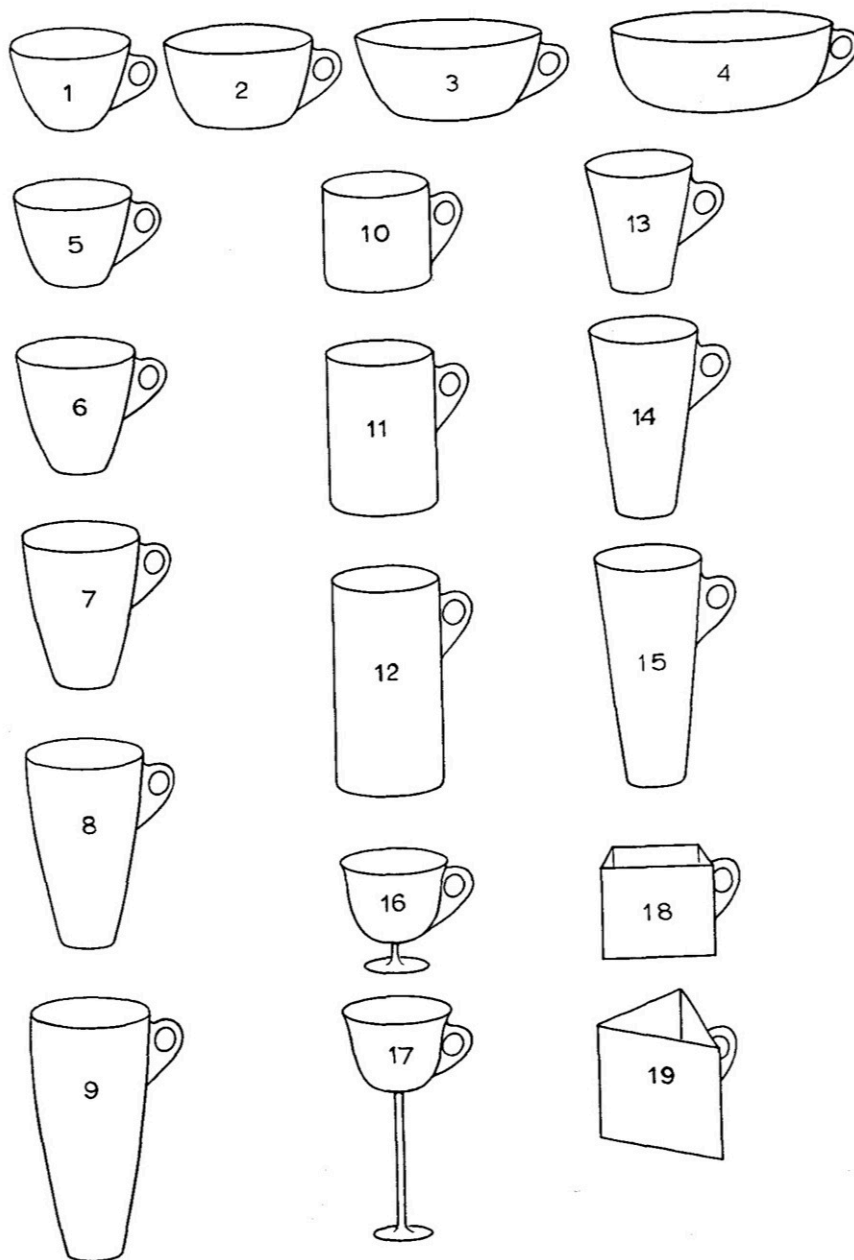
Two other aspects of Wittgenstein’s language theory, blurred edges and contextual influence, are supported by Labov’s 1973 study of word boundaries (also known as the famous “cup” experiments).¹¹ To examine the use of words such as ‘cup,’ ‘bowl,’ ‘vase,’ etc., Labov created drawings of the prototypical cup (one handle, 1:1 ratio of width to height) along with various

¹⁰ For further citations, see Reisberg 2006.

¹¹ Although Labov does not explicitly trace the theoretical foundation of his experiments to Wittgenstein in the way that Rosch and Mervis do, he was opposed to the categorical view of word meanings just as Wittgenstein was.

alterations: wider, deeper, long stem, short stem, non-circular cross-section, etc. (see Labov 1973: 354 [Figure 5, reproduced below]). He presented his subjects with these drawings in a randomized order and asked them to name the object they saw. In some versions of the experiment, he also told them that the object had particular contents (e.g., coffee, potatoes, flowers) or was made of particular materials (e.g., glass). Using only the main (head) noun of the subjects' descriptive phrases (e.g., "a long cup," "a funny cup with a stem," and "a kind of a cup" all become simply 'cup'), he plotted the frequency of each name for each object. Finally, he arranged these frequencies in charts showing the effect of variables such as width-to-depth ratio and contents (Labov 1973: 356-365; Figures 6-7 are reproduced below).

FIGURE 5. Series of cup-like objects.



(Labov 1973: 354)

FIGURE 6. Consistency profiles for cup and bowl in Neutral and Food contexts, Group A, N = 11.

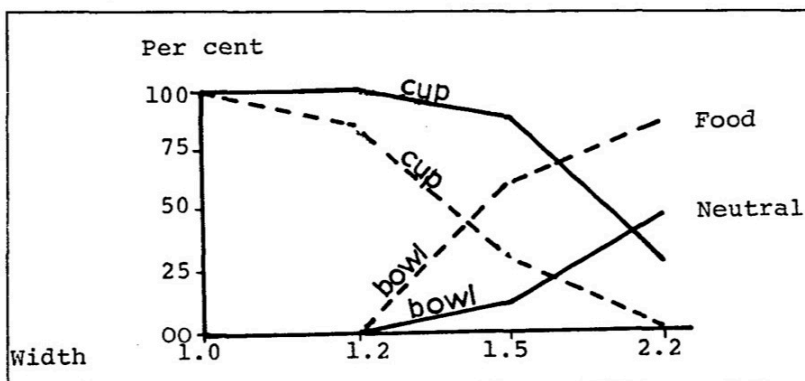
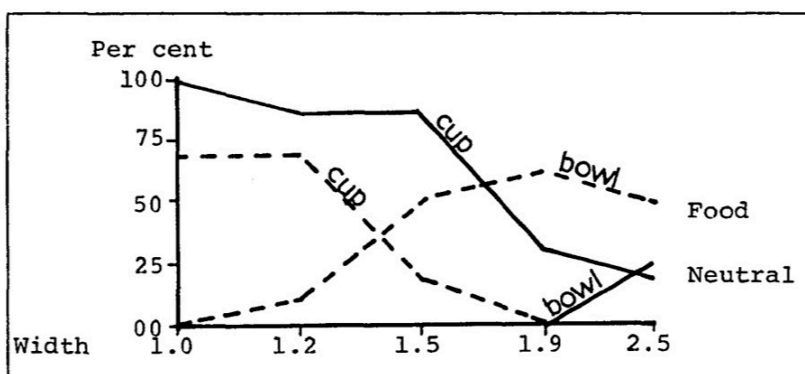


FIGURE 7. Consistency profiles denoted as cup and bowl in Neutral and Food contexts, Group B, N = 11.



(Labov 1973: 356)

As these graphs indicate, there is no clear-cut distinction between a 'cup' and a 'bowl.' Rather, as the width grew larger relative to the depth, the terms 'cup' and 'bowl' *gradually* became less common and more common, respectively. (Labov found corresponding results for 'cup' and 'vase' as the depth grew larger relative to the width [Labov 1973: 358-359].) Notice especially that there is always a point at which the 'cup' and 'bowl' lines cross; that is, there exists an object with a width-to-height ratio such that an equal number of people called it a 'cup' and a 'bowl.' If there

were a rigid boundary between the meaning of ‘cup’ and the meaning of ‘bowl,’ we would expect to see that reflected in graphs such as those above. Therefore, Labov’s results uphold Wittgenstein’s position that word meanings have blurred edges.

Labov’s experiments also showed quite clearly that word choice depends not only on the physical qualities of the object in question but also on that object’s use. When he asked his subjects to imagine that the objects held coffee, potatoes, flowers, or other things, he found that the subjects used ‘cup’ less frequently in the food context than in the neutral context. As the width-to-depth ratio increased, they switched from ‘cup’ to ‘bowl’ much sooner in the food context than in the neutral context (see Figures 6-7, above). In addition, they used ‘cup’ in the drink context slightly more often than in the neutral context, and used ‘cup’ slightly less often in the flowers context than in the food context (Labov 1973: 358). These results corroborate Wittgenstein’s emphasis on the linguistic and environmental context(s) in which a word is used.

To summarize, Labov’s experimental research supports the Wittgensteinian theory of language developed in the previous section in two ways: words have blurred edges instead of rigid boundaries (Labov’s “categories” [1973: 342ff]), and speakers’ choice of which word to use is heavily influenced by context.

3.4 Meta-analysis of vocabulary instruction

The experiments just described offer empirical evidence for the three aspects of Wittgenstein’s language theory that I have identified as most salient for vocabulary pedagogy: family resemblances, blurred edges, and contextual influence. However, for the most part this psycholinguistic research only examines the *present* state of subjects’ word intuitions; it does not address how subjects acquired their knowledge of those words and their appropriate uses. Unlike

Wittgenstein, these researchers do not consider the role of development in word knowledge or the ways in which we learn words, much less whether and how words should be taught. Therefore, we must be cautious in applying these psycholinguistic findings directly to the primary concern of this thesis: vocabulary *pedagogy*.

A substantial number of studies in education research have focused on the strength of various methods for vocabulary instruction. For my purposes, it is most useful to consider the results of Stahl and Fairbanks' comprehensive meta-analysis (1986). Stahl and Fairbanks created a series of predictions about using different methods of vocabulary instruction as independent variables. Then they classified "[a]ll vocabulary instruction studies suitable for meta-analysis available in April 1985" (1986: 78) according to the following variables: content (i.e., definitions and context), number of exposures to the target words, type of processing (drill-and-practice vs. depth-of-processing), group size, and instructional time. Finally, to measure the effectiveness of each independent variable, they used a statistical measure called "effect size" to determine which variables in vocabulary pedagogy were most powerful on multiple assessments of word knowledge (contextual vocabulary, definitional vocabulary, and passage comprehension). Although all of the results of this influential meta-analysis are instructive for vocabulary instruction, I will discuss only those three that relate to Wittgenstein's pedagogical implications:¹² definitional versus contextual content, number of exposures, and type of processing.

With respect to content, Stahl and Fairbank divided vocabulary instruction methods into five categories:

¹² For instance, Stahl and Fairbanks considered the size of the instruction group, but Wittgenstein does not address this issue and it is beyond the scope of this thesis.

1. **Definitional only:** The only information provided is a definition, synonym, and so forth. There are no examples of the word used in context.
2. **Definitional emphasis:** Some exposure is given to the word in context, but the emphasis is on the child learning the definition.
3. **Balanced:** A balance or near balance between definitional and contextual information is given.
4. **Contextual emphasis:** Although a definition is given, the major emphasis is on learning the word in context.
5. **Context only:** The child is exposed only to each word in context, with no attempt to have the child derive a definition. (1986: 75)

They called types 2, 3, and 4 “mixed” methods. These produced greater effect sizes than methods categorized as type 1¹³ (1986: 95) – a finding that supports Wittgenstein’s theory of vocabulary pedagogy. Of the three mixed methods, type 4 methods had the greatest effect sizes on word-specific vocabulary measures, while type 3 methods had the greatest effect sizes on passage comprehension measures. The passage comprehension effect sizes of type 3 methods were significantly greater even than those of type 2 methods (1986: 95-96). While these results support Wittgenstein’s recommendation to de-emphasize definitions in vocabulary instruction, they also show that there is an appropriate, though limited, use for definitions. When used in conjunction with rich contextual exposure, definitions become simply one more piece of information about the word and its uses – rather than the exclusive source of such information.

Moreover, Stahl and Fairbanks found that methods with multiple exposures or repetitions clearly produced greater effects than those involving only one or two exposures. This was true whether the multiples were repetition of the same information or exposure to various contexts

¹³ Type 5 methods, which gave the context alone with no definition, had the lowest mean effect sizes of all. But only four studies examined Type 5 methods, so this result does not tell us much about the effectiveness of context-only vocabulary pedagogy methods.

(1986: 96). This finding is consistent with the pedagogical implications I have drawn from Wittgenstein.

Finally, Stahl and Fairbanks compared drill-and-practice methods such as memorizing definitions with those that required deeper, generative processing involving breadth of knowledge. Unexpectedly, drill-and-practice methods outperformed depth-of-processing methods on most assessments, but Stahl and Fairbanks hypothesize that this may be due to the large number of exposures involved (compared to generative processing methods, which devote more time to a relatively small number of exposures), rather than to the success of the method itself. They note:

On the passage comprehension measures, only the breadth of knowledge methods produced a mean effect reliably different from zero. However, the number of effect sizes on passage comprehension measures for the drill-and-practice methods were so small [that is, few relevant studies were included in the meta-analysis] that no firm conclusion should be drawn. (1986: 97)

In other words, depth-of-processing methods are better able to achieve the goal of increasing reading comprehension, discussed in the introduction. If further studies that control for number of exposures are able to show that breadth of knowledge methods are in fact more effective than drill-and-practice methods, then this also supports Wittgenstein's emphasis on examples and family resemblances rather than definitions. But Stahl and Fairbanks were unable to confirm this.

Stahl and Fairbanks summarize the relevant portion of their results as follows:

Methods that provided only definitional information about each to-be-learned word did not produce a reliable effect on comprehension, nor did methods that gave only one or two exposures of meaningful information about each word. Also, drill-and-practice methods, which involve multiple repetitions of the same type of information about a target word using only associative processing, did not appear to have reliable effects on comprehension. Interpretation of these findings should be tempered by the small number of effect sizes [that is, studies with relevant data] involved. (1986: 101)

One conclusion to draw from this meta-analysis is that we need more research — and research that uses consistent categories, terminology, and measures — on specific methods for vocabulary instruction, such as drill-and-practice and depth of processing. However, we *can* conclude that a mixed method, involving both definitional and contextual information, is more powerful than one involving only definitions. We can also conclude that multiple exposures are more powerful than only one or two. Both findings support the implications for vocabulary pedagogy derived from Wittgenstein’s language theory.

3.5 Empirical support for incidental word learning

Even the most powerful vocabulary instruction — as determined by both philosophical theory and empirical research — will necessarily fall short of covering all the words that students must and do learn. Researchers Nagy and Herman (1987) calculate that the average student acquires approximately 3,000 words *each year* from third to twelfth grade.¹⁴ Of course, some students gain substantially more, and others substantially less. But even setting these differences aside, all students encounter and learn a large number of unknown words outside of explicit vocabulary instruction. Note that these are not words students *should* learn, to meet some pre-determined goal or standard; they are words students *do* learn. This enormous rate of incidental vocabulary growth is a fact, not an ideal.

Such a high rate of incidental word learning supports Wittgenstein’s language theory and its pedagogical implications. For it is reasonable to assume that students are not learning the definition of every unknown or uncertain word they encounter in their reading and conversations. Even if

¹⁴ They derive this estimate from five different studies that measured the average student’s vocabulary size in both Grade 3 and Grade 12. The difference between the two, divided by the number of years that have elapsed (nine), yields an estimated average annual growth. For details and citations, see Nagy and Herman 1987: 22.

they do look up or ask for a definition, they meet the definition in a social setting (i.e., a discussion with a teacher, parent, or more knowledgeable peer) and a natural linguistic context (i.e., the text they are reading or conversation they are engaging in). They certainly do not copy, drill, and memorize the definition, as they would in some vocabulary instruction methods. And yet they manage to learn several thousand words a year through context or context reinforced by definitions. This research suggests that word learners need not rely exclusively on studying dictionary definitions to acquire a large number of words.

If students do not acquire these words through explicit vocabulary instruction, then how *do* they acquire them? Nagy and Herman suggest a number of possibilities: “the speech of parents and peers, classroom lectures and discussions, school reading, free reading, and television” (1987: 24). They then consider whether oral or written language produces a context more conducive to word learning, and which is actually responsible for the bulk of word learning. The results of this debate are not relevant to this thesis;¹⁵ but we can infer that large-scale incidental vocabulary growth probably relies on a combination of extensive reading and word-rich conversations.

3.6 Empirical conclusion

Empirical evidence supports Wittgenstein’s theory of language as it is relevant to vocabulary pedagogy. The set of experiments performed by Rosch and Mervis confirm that word meanings and concepts have family resemblance relationships. In addition, Labov’s experiments suggest that word meanings have blurred edges and are influenced by context. Moreover, research on effective

¹⁵ Nagy and Herman (1987) themselves claim that, while oral context accounts for the bulk of word learning throughout the preschool years, once students can read they acquire most new words through written contexts. They also suggest that word learning from context is so difficult to measure because it occurs incrementally; word knowledge develops in tiny amounts across multiple exposures to a word in context. This difficulty may partially account for the exceptionally small number of context-only studies available for Stahl and Fairbank’s meta-analysis.

vocabulary instruction itself also supports Wittgenstein's pedagogical implications as I have interpreted them. Vocabulary instruction methods based exclusively on drilling definitions are measurably less effective than methods that combine definitions and context; in addition, methods that expose students to target words multiple times and in multiple contexts are particularly effective. Finally, most words are never explicitly taught through specific vocabulary assignments (though they may still be learned in a school context) but are acquired incidentally and contextually, through extensive interaction with oral and written language. On the basis of this empirical support, it is reasonable to conclude that vocabulary pedagogy based on Wittgenstein's language theory will be effective.

4. A Wittgensteinian Analysis of *Words, Words, Words*

4.1 Introduction

So far I have developed Wittgenstein's theory of vocabulary pedagogy and supported it philosophically and empirically. But I have yet to consider what this vocabulary pedagogy might actually look like in a classroom, or what resources exist for teachers to implement it. To answer the second question, and the first by means of the second, I will analyze one teacher resource book, *Words, Words, Words: Teaching Vocabulary in Grades 4-12* by Janet Allen (1999). I chose this particular text because Allen bases her recommendations for teaching vocabulary on empirical and theoretical research, in addition to conversations with current teachers and her own observations over many years of teaching. Moreover, she translates research into specific practices, guidelines, and tools that teachers can apply directly to their own classrooms. So teachers can use

her approach to implement effective vocabulary pedagogy in accordance with Wittgenstein's theory.

From the very first pages, Allen's pedagogical approach is holistic, grounded in ordinary language use. She observes that her students "used and played with the language we created together — not the language I assigned" (1999: 3). Accordingly, she bases not only her explicit vocabulary instruction but also her overall use of language in the classroom on the goal of developing students' linguistic abilities: "I saw my role as one of demonstrating a more advanced level of language. I tried not to take my language to their level but rather to bring their language to mine" (1999: 4). With this perspective as a promising start, she goes on to present a thorough vocabulary pedagogy that corresponds well with both Wittgenstein's theory and empirical research.

4.2 Integration, repetition, and meaningful use

Allen focuses on three properties of effective vocabulary instruction that she considers critical: integration, repetition, and meaningful use (1999: 12, 35, 69; see also Nagy 1988, cited in Allen 1999). Clearly, these properties are in line with the philosophical and empirical stance developed in this paper. Integration and meaningful use require vocabulary words to be learned *not* as isolated lists and definitions, but rather as components of a rich and relevant linguistic context, replete with family resemblances and blurred edges. Repetition, overwhelmingly shown to be effective in Stahl and Fairbank's meta-analysis, enables students to incorporate new words into their existing linguistic knowledge.

Allen applies integration, repetition, and meaningful use to her vocabulary pedagogy in three ways: extensive reading, explicit instruction, and holistic assessment. Accordingly, I consider each of these pedagogical methods in turn.

4.3 Incidental acquisition: extensive reading

According to Allen, increasing the time students spend reading is not simply a matter of replacing an over-emphasis on definitions with an equally dangerous over-emphasis on context. In fact, she mentions numerous studies such as those discussed in the previous section to demonstrate that a word's meaning can be almost impossible to derive from a single context (such as one example sentence). However, she also cites research that “the amount of time spent reading [is] the best predictor of vocabulary growth.” For this reason, she continues, “I’m not willing to abandon the use of context; rather, I suggest we expand our teaching of what it means to use context and increase the amount of time students spend reading” (1999: 21). She even places extensive reading – to, with, and by students – above all other vocabulary-teaching methods:

None of the strategies in this book, nor all of them combined, will take the place of the wealth of words learned in a strong reading program that includes time for you to read to your students, time for them to read with you and other students, and time for them to read self-selected books independently. This reading forms the larger context for any word study a teacher may choose to do. (1999: 31)

This emphasis on reading grants students access to natural language about topics that interest and delight them.

To achieve the goal of exposing students to as much natural language as possible, Allen urges teachers to fill their classrooms with “books that many students... find truly engaging, that... build specialized vocabulary knowledge, and that... help them find answers to their questions” (1999: 83). Accordingly, she then lists ten pages’ worth of suggestions (1999: 84-94), emphasizing non-fiction, informational books “because of the rich, diverse vocabulary found there” (1999: 84). She draws these book suggestions from a wide range of topics that students may find interesting and relevant: art and artists, health and physical education, language arts, math, science, and social

studies, as well as magazines and alphabet books. The inclusion of books about science and especially math is particularly important, because it might be easy to suppose that, although more “general” vocabulary can be learned in context, the specialized vocabulary of these technical subjects must be taught through definitions. But Allen observes that books such as *Anno’s Mysterious Multiplying Jar* and *Sir Cumference and the First Round Table* “convey mathematical information and specialized language that help students transfer word and concept knowledge to both textbook and real-life mathematical challenges” (1999: 89). These kinds of books can be used to introduce concepts, to reinforce definitional knowledge, and to increase interest in learning about math and science.

Two more of Allen’s observations from this list of book recommendations are worthy of note. First, she encourages the use of sports books, because “[f]or many readers, understanding words like *spectacular*, *well-deserved*, and *affiliate* in the context of sports will improve their ability to transfer those words to other areas” (1999: 86). In other words, exposure to new vocabulary in a natural and engaging context facilitates students’ understandings of multi-contextual word families. Second, Allen mentions two books of short stories written by teen writers, giving the following reason: “It is important that students hear their teachers using such language, of course, but encountering peers who use content-rich language is even more significant. It helps student writers give voice to their thoughts in language they previously might not have considered” (1999: 89). Students will mimic the language they read and hear, especially when it comes from those they

respect or identify with.¹⁶ Both of these specific comments, in addition to the booklist as a whole, indicate that Allen is in line with the vocabulary pedagogy developed in this thesis.

4.4 Explicit instruction: tools and techniques

To supplement – but by no means supplant – this emphasis on extensive reading, Allen recommends “varying levels of direct instruction” in vocabulary words (1999: 6). All of her tools and techniques go well beyond “look it up in the dictionary,” though they may include definitions as one of many ways to give students “a world of knowledge” about words and the world (1999: 35). Although many of her tools and techniques are praiseworthy, I will focus on just three that illustrate the correspondence between her approach and Wittgenstein’s.

All of Allen’s methods promote blurred edges in students’ understandings of words, because they present those words in natural contexts and emphasize meaningful use over memorization of definitions. But one tool in particular, called the “linear array,” focuses directly on blurred edges in word meanings. Allen describes linear arrays as “visual representations of degree... graphic organizer[s] for depicting gradations between two related words” (1999: 52). Students place two contrasting words (e.g., “freezing/boiling,” “minute/immense,” “always/never”) at opposite ends of a piece of paper. Then they fill in the space between with an ordered list of words that connect the two extremes. Allen observes that “[a]n activity like this helps students examine subtle distinctions in words” (1999: 53) – distinctions that are important for understanding both blurred edges and family resemblances.

¹⁶ This particularly agrees with Wittgenstein’s views on learning language-games, which I have not addressed in this thesis.

Allen also suggests tools that are specifically designed to bring out connections and relationships among words. An excellent example is the chart labeled “Multiple Meanings” (1999: 61). In the center of the worksheet, students copy the context where they originally found the target word. On the left, they write generalized meanings of the word, and on the right, more specific instances where the word would be appropriately applied. At the bottom they put what Allen calls “family words:” morphological variants such as past tenses and plurals, adjective or adverb forms, etc. Allen does not use the term “family words” in the same way that Wittgenstein uses “family resemblances,” so the two concepts should not be confused;¹⁷ however, the entire chart enables students to discuss and learn different sorts of family resemblances, such as a word’s various meanings and contexts.

Finally, many of Allen’s tools ask students to list both examples and non-examples (1999: 43, 50, 57, 58, 59). This one technique alone encompasses nearly all the theoretical principles I have derived from Wittgenstein. At the most basic level, it requires students to consider word meanings in terms of examples, rather than abstract, acontextual definitions. When asked to think of things that are and are not ‘preposterous,’ one student listed “platypus” as an example and “dog” as a non-example (1999: 57). Both are concrete and easily memorable, and will probably help the student internalize the word’s meaning. Moreover, the technique emphasizes family resemblances by indicating different uses of a word, as well as the word’s opposite(s) (see the Analysis Map of ‘immigration,’ reproduced on the next page).

¹⁷ It appears that there is a family resemblance structure to the concept ‘family resemblances.’

E.11 Analysis Map

Word or name
immigration

Define or rename
when people move to another country
looking for a "better life"

Compare to:

Irish move to escape potato
Famine

Island movement to U.S.
looking for work

Mexican migrant
workers

Jews escape from Hitler

Blacks flee to Canada to
escape slavery

Contrast with:

moving from state to state
vacation

moving from Puerto Rico
to U.S.

visiting family in another
country

business trip

Examples

Underground Railroad

Nazis

Hitler

Copyright © 1999 Janet Allen, Words, Words, Words. Stenhouse Publishers.

(Allen 1999: 50)

The examples and non-examples technique also encourages blurred edges: two words could have the same example but different non-examples, or vice versa. For example, "bad behavior" is given as a characteristic of both 'preposterous' and 'reprehensible,' but something 'preposterous' is unexpected while something 'reprehensible' might be predictable. (Notice that Allen is fairly relaxed about what counts as an example or non-example; adjectives and descriptions are

acceptable in addition to things.) In all these ways, the tool of examples and non-examples corresponds to Wittgenstein's theory of word learning. It is also in line with the findings of empirical research on mixed-method vocabulary instruction: Allen frequently includes a definition either before or after using this technique. But nowhere does she rely *solely* on definitions, or examples and non-examples, or context, or any other single way of giving information about words.

4.5 Assessment: a holistic approach

Just as Allen advocates explicit vocabulary instruction that incorporates a variety of methods, so too she argues that “[t]he assessment of vocabulary instruction should be varied and meaningful” (1999: 104). In fact, the connection between teaching and assessment in her approach is quite clear: “If vocabulary instruction changes... the tests must reflect a different way of thinking about language” (1999: 96). Instruction based upon integration and meaningful use must be complemented by tests formulated on the same principles.

Allen gives a number of sample questions for assessing students' word knowledge, all of which “[r]equire students to think and write about the word, not just match definitions.” She notes that “[w]hile these questions may take a bit more time to design and grade... these tasks help students realize that there is no single definition for a word” (1999: 98). Indeed, the implicit message that a particular testing method sends to students may be at least as important as the explicit information that teachers gain about their students' learning (1999: 99). Thus, it is especially important that vocabulary assessments properly embody the view of language we want students to embrace. As Allen remarks,

Whatever assessment you choose, moving students into roles that require them to recognize words, think about ways the words could be used in multiple contexts,

and write about their personal connections to the words will be a step forward in helping them become independent word learners. (1999: 104)

Because the ultimate goal of Allen's vocabulary pedagogy is meaningful use, not memorization of definitions, her approach to both instruction and assessment is in accordance with Wittgenstein's theory of language.

4.6 Conclusion to the book analysis

To summarize, Allen presents a philosophically and empirically sound approach to vocabulary teaching based on the key principles of integration, repetition, and meaningful use. She emphasizes extensive reading, thereby encouraging students to develop their word knowledge in context, with family resemblances and blurred edges. The explicit instructional tools she suggests also promote family resemblances and blurred edges, and employ a mixed method (i.e., a combination of both definitions and context) that is supported by empirical research on effective vocabulary instruction. Teachers who wish to implement Wittgenstein's vocabulary pedagogy in their own classrooms need have no fear about using Allen's methods and techniques.

5. Conclusion

After pointing out the relationship between vocabulary and reading comprehension in Section One, in Section Two I used the later Wittgenstein's theory of language, as expressed in the *Philosophical Investigations*, to develop implications for vocabulary pedagogy. In Section Three, I then compared those implications with the findings of empirical studies on concept structures, word meanings, effective vocabulary instruction, and incidental word learning. Lastly, in Section Four, I reviewed a teacher resource book, *Words, Words, Words*, that puts holistic vocabulary pedagogy into practice.

Due to the interdisciplinary nature of my project, I have divided my conclusion into two sets of lessons: lessons for philosophy and lessons for education. Although the fields are so closely interwoven in my context that the division may at times seem unnatural or even arbitrary, I hope the reader will find this organization helpful.

5.1 Lessons for philosophy

I draw from this thesis four lessons for the field of philosophy. First, I take Wittgenstein to be correct. Word meanings have blurred edges, not rigid boundaries. They exist in word families of various kinds. And they must be considered in the contexts in which they are used. Any adequate philosophy of language must recognize these facts. Moreover, on the basis of this theory, I have suggested that vocabulary pedagogy that uses examples and discussions in addition to definitions will be better at increasing students' word knowledge than vocabulary pedagogy that relies solely on memorizing definitions. Any adequate philosophy of language must recognize *this* as well.

Second, the research of this thesis could be expanded by looking more closely than I have done at Wittgenstein's views on word meanings and language acquisition, including his theory of language-games and his other works besides the *Philosophical Investigations*. A closer examination of his writing might bolster or nuance the interpretation I have given. Alternatively, it might give rise to a new interpretation. That interpretation would in turn yield different implications for vocabulary pedagogy, which could then be tested against empirical research, including (but not limited to) the studies I have discussed.

Third, the relevance of philosophical ideas and arguments stretches far beyond philosophy itself. In particular, philosophy has important implications for empirical research, pedagogy, and education policy. Philosophy of education professors and books often discuss the relationship

between philosophy and education only in terms of ethics, epistemology, and philosophy of mind. Yet I have shown how Wittgenstein's theories about language and word learning can be applied directly to vocabulary pedagogy and related to research in cognitive psychology, psycholinguistics, and education. Thus, philosophy of language also connects to education in significant ways.

A fourth and final lesson is that future philosophical research might trace the pedagogical implications of other theories in philosophy of language, ones that contrast or even conflict with that of the later Wittgenstein. A more nuanced philosophy of language might yield an even more powerful vocabulary pedagogy; on the other hand, if the pedagogical implications derived from a particular philosophy of language are noticeably ineffective, perhaps it should be considered suspect even as a philosophy of language. In sum, I hope that this thesis will engender a greater degree of cross-disciplinary communication between philosophers and researchers in other fields, especially education.

5.2 Lessons for education

I also draw from this thesis four lessons for the field of education. First, one area related to vocabulary pedagogy in particular needs further investigation. Stahl and Fairbanks (1986) demonstrated that a "mixed method" (one that combined definitional and contextual information) was more effective than either a definition-only method or a context-only method. However, they were not able to determine which of the three types of methods within the "mixed" category — contextual emphasis, definitional emphasis, or balanced between the two — was most effective:

The methods that did appear to produce the highest effects on comprehension and vocabulary measures were methods that included both definitional and contextual information about each to-be-learned word (or "mixed" methods). It was difficult to draw conclusions as to the effects of relative emphases of definitional or contextual

information from these data. These effects need to be examined further. (1986: 101)

It would therefore be useful to conduct research on this specific issue: Should one type of information about words receive greater emphasis than the other in vocabulary instruction – and if so, which one? Or should the two be used roughly equally?

Stahl and Fairbanks were also unable to determine the effectiveness of depth-of-processing methods as compared to drill-and-practice methods, because the number of repetitions is usually so much greater in the latter than in the former. So the second lesson is that we need to find out what happens when number of repetitions is held constant between the two types of methods: Are drill-and-practice methods still more effective than depth-of-processing methods even when both have the same number of repetitions? Studies specifically addressing these questions and those above could also confirm Stahl and Fairbanks' findings more directly, without having to resort to a meta-analysis.

Another problem encountered by Stahl and Fairbanks was the difficulty of knowing exactly what happened in the classroom in terms of vocabulary instruction. They based their categorization of the methods in the various studies on the descriptions found in those studies; yet they caution that “[t]he degree to which the actual methods differed from the descriptions reported cannot be determined, but we can assume there were some differences.” In fact, “[o]ne of us (Stahl, 1983) reported that a method that was intended to include only definitional information actually included a great deal of contextual information. This contextual information came from the students and from the dictionaries” (1986: 103). Clearly, then, the third lesson is that more information is needed about actual classroom practices for teaching vocabulary. I propose that the

best way to gain this information is through classroom observations conducted by the fields of sociology and anthropology of education. Such ethnographic information would reveal what teachers and students are actually doing in schools when they claim to be studying vocabulary.

Fourth and finally, to whatever group or groups will make the final decisions about implementing my findings in the classroom (test or textbook writers, policy-makers at the district, state, or national level, school boards, principals, or even teachers themselves), I offer this concluding summary of my findings. Although studying definitions can be helpful for learning some words in some contexts, knowledge of definitions should be a tool of vocabulary pedagogy — not its goal. Moreover, vocabulary pedagogy that emphasizes copying and memorizing dictionary definitions is relatively unproductive. A more helpful way to encourage word learning is to have students read widely and well, and to supplement that extensive reading with more explicit vocabulary instruction that includes both definitional and contextual information. Assessments, too, should address holistic word knowledge and multiple meanings rather than the ability to repeat single definitions. Such a vocabulary pedagogy is philosophically logical and empirically sound, and so can be trusted to be successful in the classroom. It communicates to students that words are relevant to their lives and the world and therefore worth learning.

References

- Allen, J. (1999). *Words, words, words: Teaching vocabulary in grades 4-12*. Portland, Me: Stenhouse Publishers.
- Baker, G. P., & Hacker, P. M. S. (1980). *Wittgenstein, understanding and meaning*. Chicago: University of Chicago Press.
- Cook, J. W. (2000). *Wittgenstein, empiricism, and language*. New York: Oxford University Press.
- Glock, H.-J. (1996). *A Wittgenstein dictionary*. Oxford, UK: Blackwell Reference.
- Kenny, A. (1973). *Wittgenstein*. Cambridge, Mass: Harvard University Press.
- Labov, William. (1973). The boundaries of words and their meanings. In Bailey, C. J. N., & Shuy, R. W. *New ways of analyzing variation in English*. Washington: Georgetown University Press.
- Merriam-Webster's Word Central Dictionary*. (2007). Retrieved from <http://www.wordcentral.com/home.html>.
- Nagy, W. and Herman. P. (1987). Breadth and Depth of Vocabulary Knowledge: Implications for Acquisition and Instruction. In McKeown, M. G., & Curtis, M. E. (2009). *The Nature of vocabulary acquisition*. New York: Psychology Press.
- National Reading Panel (U.S.), & National Institute of Child Health and Human Development (U.S.). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: reports of the subgroups*. Washington, D.C: National Institute of Child Health and Human Development, National Institutes of Health.
- Reisberg, D. (2006). *Cognition: Exploring the science of the mind*. New York: W. W. Norton.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 4, 573-605.
- Stahl, S. A., & Fairbanks, M. M. (1986). The Effects of Vocabulary Instruction: A Model-Based Meta-Analysis. *Review of Educational Research*, 56, 1, 72-110.
- Wittgenstein, L., & Anscombe, G. E. M. (2001). *Philosophical investigations: The German text, with a revised English translation*. Oxford: Blackwell.