



**LACUS
FORUM
XXVII**

*Speaking and
Comprehending*



RICE

© 2009 The Linguistic Association of Canada and the United States (LACUS).
The content of this article is from *LACUS Forum* 27 (published 2001). This article and others
from this volume may be found on the Internet at <http://www.lacus.org/volumes/27>.

YOUR RIGHTS

This electronic copy is provided free of charge with no implied warranty. It is made available
to you under the terms of the Creative Commons Attribution-NonCommercial license
version 3.0

(<http://creativecommons.org/licenses/by-nc/3.0/>)

Under this license you are free:

- **to Share** — to copy, distribute and transmit the work
- **to Remix** — to adapt the work

Under the following conditions:

- **Attribution** — You must attribute the work in the manner specified by the author
or licensor (but not in any way that suggests that they endorse you or your use of the
work).
- **Noncommercial** — You may not use this work for commercial purposes.

With the understanding that:

- **Waiver** — Any of the above conditions can be waived if you get permission from the
copyright holder.
- **Other Rights** — In no way are any of the following rights affected by the license:
 - Your fair dealing or fair use rights;
 - The author's moral rights;
 - Rights other persons may have either in the work itself or in how the work is
used, such as publicity or privacy rights.

Notice: For any reuse or distribution, you must make clear to others the license terms of
this work. The best way to do this is with a link to the web page cited above.

For inquiries concerning commercial use of this work, please visit
<http://www.lacus.org/volumes/republication>

Cover: The front cover of this document is licensed under the Creative Commons
Attribution-No Derivative Works 3.0 license (<http://creativecommons.org/licenses/by-nd/3.0/>) and may not be altered in any fashion. The LACUS “lakes” logo and Rice University
logo on the cover are trademarks of LACUS and Rice University respectively. The Rice
University logo is used here with permission from the trademark holder. No license for use
of these trademarks outside of redistribution of this exact file is granted. These trademarks
may not be included in any adaptation of this work.

LEXICALIZATION AND OPACITY

AYA KATZ

Providence University

INTRODUCTION. Lexicalization is a process whereby new lexemes are formed. The process is usually seen as historical, and the resultant lexeme becomes a synchronic unit. A certain degree of componential opacity is associated with a fully lexicalized item. (Heine et al. 1991:95) To the extent that speakers still view a lexeme as composed of its parts, the lexicalization process may be considered less than complete within a functional theoretical framework. Such an analysis is consonant with certain typologies, but flies in the face of synchronic evidence in languages where componential transparency is the norm rather than the exception.

For example, it is well known that Chinese has compound lexemes (Li & Thompson 1998). However, there is evidence from brain damaged patients that a component's status as verbal or nominal can affect access to same within a fully lexicalized compound. So Mandarin 吃飯 *chi fan* 'eat', while enjoying full lexeme status, is more likely to be retrieved as just 飯 *fan* 'rice' by patients with damage to Broca's area, and as 吃 *chi* 'eat' alone by those suffering from damage to Wernicke's area (Bates & Chen 1991; Zhang ms.). On the other hand, speakers of English often have difficulty accessing the componential makeup of words whose root morphemes have independent lexemic status: e.g., *rooster* [*< roost + er*]. In Hebrew most morphemes are bound, because consonants carry mostly lexico-semantic information, while vowels code mostly grammatical and categorization data. Despite the highly dependent nature of Hebrew morphology, speakers find it extremely transparent.

The purpose of this paper is to determine the implications for lexicalization from the relative opacity or transparency of morphemes in typologically distinct languages. Lexicality will be defined independently from componential opacity and morphological dependence. Opacity will be subcategorized as circumstantial (objective) versus psychological (subjective). Correlations between and among opacity, dependence and lexicality will then be tracked.

This paper will use examples from Hebrew, English and Chinese to determine how overall systematicity in lexical derivation affects speakers' strategies in processing lexemes. The actual transparency/opacity of a particular item may be seen to be less important than whether a speaker expects lexemes to be componentially transparent. We will consider whether systematicity in derivation primes a subject to use componential analysis even where it is less productive. We will explore whether lack of systematic derivation overall deters a speaker from using a strategy of componential analysis even on lexemes whose components are well known to the speaker. The

psychological underpinnings of opacity will be explored within a framework of typological comparison.

1. THE THEORETICAL BACKGROUND. *Lexicalization*, as a linguistic label, emerges from the functional linguistic school of thought, and it is often paired with the term *grammaticalization*. (Heine et al. 1991; Hopper & Traugott 1993; Traugott & Heine 1991). As historical processes, both lexicalization and grammaticalization often result from reanalysis. However, while grammaticalization gives rise to new productive patterns that add to the grammar of a language, lexicalization produces seemingly isolated additions to the lexicon. Regularity, productivity, and transparency as components are therefore attributed to grammaticalized items, whereas idiosyncrasy, randomness and opacity are expected of lexicalized words.

According to Heine et al. (1991:95), the difference between the two terms can be summed up as follows: 'Assuming that both involve some kind of 'idiomization', the latter may be said to be morphologically productive in the case of grammaticalization but not in that of lexicalization.'

The above is representative of statements about lexicalization in grammaticalization literature as a whole, and it carries with it some subtle implications that are worth exploring. There is also an almost imperceptible shift between a diachronic use of the term to a synchronic use that has far reaching implications. An item is often seen to be more lexicalized if its componential makeup is relatively opaque, and less lexicalized if it is relatively transparent. This implies that full lexemic status ought to be accorded only to items whose derivation is synchronically unreadable. If we followed such logic to its ultimate conclusions, we would decide that a monomorphemic word is the most lexemic of all words. However, there are serious drawbacks to this approach when applied crosslinguistically to languages belonging to contrasting typologies.

In a language such as Hebrew, there are almost no monomorphemic words, since vowels and consonants play complementary roles in word formation. The majority of vowels found in a Hebrew word code derivational and inflectional contrasts, while a majority of consonants code lexico-semantic information. Morphemes are discontinuous, and every syllable carries parts of more than one morpheme. But despite the extremely fusional nature of the language, componential transparency is the norm, rather than the exception.

In fact, in Modern Hebrew, non-linear derivation patterns, involving the use of discontinuous roots and templates to form new words, are still extremely productive. (Bolzky 1999).

Mandarin is known as an isolating language, but it has a large stock of complex lexemes. (Li & Thompson 1998). The dependence level of these lexemes is low, in that they are not phonologically affixed and there is no reduction or resyllabification involved. In this sense, the morphologically complex lexemes of Mandarin are quite different from those of Hebrew. But both Mandarin and Hebrew enjoy a high level of morphological transparency. There is evidence from brain damaged speakers of Mandarin that

components are accessed according to their grammatical categorization, rather than the grammatical category of the word as a whole (Bates & Chen 1991).

Multimorphemic words in English are more tightly fused than those in Mandarin, but less so than in Hebrew. Multisyllabic lexemes in English are common, but there is by no means a one-to-one correspondence between morpheme and syllable. Speakers of English have great difficulty analyzing lexemes into their component morphemes. Even in such common words as *heal* and *healthy*, speakers of the language require instruction in its history in order to identify the common root, according to the introduction to a leading textbook on the history of English (Pyles & Algeo 1993:2) In other words, commonly used English lexemes are often componentially opaque for the average speaker.

The implications for the concept of lexicalization from the above observations of Hebrew, Mandarin and English are as follows:

- (1) Lexical status is not necessarily dependent on degree of fusion.
- (2) Lexical status is not necessarily dependent on opacity or semantic bleaching of components.
- (3) Opacity and fusion are independent of each other.

The significance of these initial observations is as follows: if lexical status can be established independently of both fusion and opacity, then lexeme formation (i.e., *lexicalization*) cannot be defined as a necessarily 'morphologically unproductive'.

Heine et al.'s statement (1991:95) might suggest that componential opacity is something that indicates a high level of lexicalization, where the term lexicalization is defined as incremental. The higher the opacity, the higher the degree of lexicalization. But if such an interpretation were used, most words in a language such as English might be judged to be more lexicalized than those in Hebrew and Mandarin. What would be the implications of such a finding?

One possibility would be to suppose the individual words have undergone a greater process of semantic erosion and that opacity results from such bleaching. But an alternative explanation is available: it is not so much that individual words have had their morphemes bleached, but rather that the speakers of the language, due to lack of overall systematicity in lexical patterning, have been conditioned to ignore componential analysis as a useful strategy.

This paper proposes to probe the issue by comparing psychological componential opacity (the psychological ability of speakers to analyze components) with circumstantial componential opacity (whether the requisite cues are synchronically there in the language to allow for such analysis).

2. PSYCHOLOGICAL COMPONENTIAL OPACITY VERSUS CIRCUMSTANTIAL COMPONENTIAL OPACITY. An extreme example of circumstantially opaque derivation in English is *hussy* [*<housewife*]. (Pyles & Algeo 1993:265) The phonological reduction and fusion that occurred over time have rendered the lexeme morphologically opaque

and therefore monomorphemic. Speakers of English cannot identify the component parts of the word, because no one could do so without knowing its history. The information is simply not available synchronically within the language as a whole. Therefore the opacity is circumstantial, not psychological.

An example of psychological opacity in English can be found in the lexeme *rooster* [*<roost +er>*]. Here, the morpheme *roost* has separate lexemic status, and the suffix *-er* is productive throughout the language. Speakers synchronically have enough information to easily analyze the word. If they fail to recognize its components, this is due to psychological opacity. (That many do fail to recognize the components has been related to me by a number of native speakers of English who first learned of this word's composition in a morphology class.)

The methodology proposed has at its core the comparison of psychological versus circumstantial opacity. If a lexical item is both circumstantially and psychologically opaque, there is nothing surprising in such a situation. Likewise, if a lexeme is both circumstantially and psychologically transparent, there is nothing unusual about that. However, if there are lexemes that are psychologically opaque but circumstantially transparent, such a finding would have theoretical significance. Moreover, if a lexical item is circumstantially opaque but psychologically transparent, this would be a very interesting find.

3. PROPER NAMES AND THEIR IMPLICATIONS FOR LANGUAGE-WIDE COMPONENTIAL TRANSPARENCY. It might be supposed that opacity or transparency of lexemes would be largely determined by the necessity of engaging in creative acts of word formation. Since most speakers seldom have occasion to coin new words, their skills in the reverse process of analyzing the words with which they are presented might be atrophied.

The generative framework, for instance, equates creativity with the ability to generate novel utterances. (Chomsky 1966). The focus is on the speaker, not the hearer. Is transparency directly related to productivity? Some of the literature might suggest this.

Bolozky (1999:xi) tells us in his preface: 'Somewhere in the heart of every speaker of Israeli Hebrew, placed between a love of oriental food and a passion for politics, there is a constant urge to form new words.' He then proceeds to test productivity by a series of experiments in which speakers are requested to coin new words, and the templates and roots which they use are analyzed for productivity.

The approach of this paper is quite the reverse. It is speakers' comprehension of the componential makeup of commonly used lexemes that concerns us here. Nevertheless, the one time when most ordinary people have occasion to form a word is in the process of naming. A proper name is the most extreme form of lexicalization available. Since many speakers do engage in naming, they must also have some skill in deducing the meanings of names given by others. A brief examination is in order, in the spirit of a crosslinguistic comparison.

The following quoted material is of a non-academic nature, and it is presented not in order to prove any linguistic assertions contained therein, but to demonstrate the attitudes toward names displayed by ordinary lay people. In a short article, intended

for Chinese speaking students of English, Oscar Chiang explains: ‘The Chinese are more particular than most Americans when choosing personal names. Chinese names often have special meanings, while most American names do not. Chinese parents or grandparents usually select names for their children or grandchildren with great care in order to create a certain expectation or ideal for the child to follow. On the other hand, American parents adopt the names of relatives (and even friends) for their children, as a way of honoring that person.’ (Chiang 1999).

The facts on which the above lay perception is based are as follows: most names used by English speakers are not English words. Most are borrowings from other languages, and even many whose origin is Anglo-Saxon are not synchronically recognizable English words. As a result, English speakers have been conditioned to think of names as semantically empty labels. The attitude is so systematic for English speakers that it colors their perception even of names that are circumstantially transparent in modern day English.

There is evidence that at least some speakers of English, while being perfectly aware of the meaning of the word *rose*, appear to be concurrently unaware that the name *Rose* also means ‘rose’.

In a web search of name dictionaries, one such dictionary listed the name *Raisa* with the definition ‘rose’. The name *Rose*, however, was defined as: ‘Of the Rose Blossom’. In other words, the English speakers who compiled that name dictionary were not comfortable with having a name mean the same as a common word with which it is identical. At most, for many speakers of English, a name can have an etymology, but not a lexical meaning.

Since ‘meaning’ cannot be prescribed, and native speakers of a language are the only ones who determine the meaning of words in their language, another way to express this is to say that the name *Rose*, in the minds of at least some speakers of English, has no meaning directly associated with that of the common noun *rose*.

This assertion may seem slightly controversial, since many English speakers who are linguists have etymologically based perceptions of lexical meaning associated with word origin. Many literate speakers of English are aware that names have ‘meanings’—without knowing what those meanings are. The meanings form a set of arcane knowledge. English speakers often consult name dictionaries to find out what their names ‘mean’. What they are really looking up is etymologies, true or fabricated.

The same desystemetizing influence that we find in the proper noun inventory of English speakers can also be found in the list of common nouns. Borrowed words in English have elaborate morphology which is fairly systematic, but unintelligible to the untrained speaker. Most Germanic common nouns in English are monomorphemic. Barnyard animals make for a good subgroup: *cow*, *horse*, *dog*, *cat*, *duck*, *goose*, *hen*, *cock*. Words such as *chicken* and *rooster* are in fact bimorphemic—*chick-en*, *roost-er*. However, due to the lack of systematicity in the language, many speakers do not interpret them as such. A circumstantially transparent word such as *rooster* becomes psychologically opaque to the average English speaker, perhaps because of lack of priming.

4. COMPONENTIAL PRIMING IN CHINESE. In the case of Chinese the morphological components of a multimorphemic lexeme are not fused, by reason of language typology. The phonological independence of each morpheme has led scholars to suppose in the past that each syllable constituted a separate lexeme. It is by now well established that phonological word boundaries and lexemic word boundaries are not necessarily coterminous. (Li & Thompson 1998). Thus, a word such as 吃飯 *chi fan* 'eat' is a single unit for purposes of lexical insertion into a syntactic pattern, despite the phonological independence of its components, 吃 *chi* 'eat' and 飯 *fan* 'rice'. Nonetheless, psycholinguistic evidence shows that the component morphemes continue to enjoy psychological transparency, despite the lexical status of the compound. (Bates & Chen 1991).

This componential make-up of a majority of Chinese lexemes, both nouns and verbs, is systematic in terms of being widespread, but not in terms of generativity. This is derivation, not inflection. We can explain how a word is derived after the fact, but we cannot absolutely predict what derivational paths will be followed. In that sense, there is no anticipatory system for speakers to follow. The patterns are both transparent and productive, but not systematically generative.

Nonetheless, derivation in Chinese is by and large open and accessible to the average speaker. Relative to English, Chinese derivational morphology enjoys a high level of psychological transparency.

This does not, however, mean that every derivation is equally transparent. For instance, the word for China, 中華 *zhong hua*, is composed of two morphemes. Speakers will readily explain that 中 *zhong* means middle, and that China is so called because it is at the center of the world. But the second morpheme 華 *hua* is not as accessible. When questioned about it, speakers look confused and offer a variety of responses, but most admit they are not sure what it means.

As it turns out, the reason for the relative opacity of 華 *hua* directly correlates with word frequency. The word 中 *zhong* is a common word in Modern Mandarin, and hence compound words derived from it are also transparent in terms of the meaning of that component of the compound. The word 華 *hua* is somewhat archaic. It appears in compounds and especially in names, but its more literal meaning is bleached and obscured. Hence, the opacity of the component morpheme is entirely circumstantial. Speakers do not have synchronic access to the word 華 *hua*, and hence they do not know what it contributes semantically to the compound.

The circumstantial opacity described above contrasts nicely with the circumstantial transparency but psychological opacity that English speakers experience with the *roost-* of *rooster*. Chinese speakers do not know what 華 *hua* means, so they also do not know what it contributes to 中華 *zhong hua*. But there are English speakers who do know what *roost* means, yet do not recognize it when it serves as the root for a morphologically complex word.

The non-fusional nature of Chinese phonology may be credited with the ease with which speakers analyze the component morphemes of their multimorphemic words. However, this is not the case in Hebrew, where componential transparency is also the norm, despite the widespread use of discontinuous morphemes.

5. HEBREW MORPHOLOGICAL TRANSPARENCY VS. PHONOLOGICAL OPACITY. In the case of Hebrew, derivational patterns are so predictable and the paradigms are so widespread that one can almost assert that Hebrew derivation is not only systematic, it is downright generative! An average speaker can be given the task of generating a new lexeme using a particular template and root, and in as little time as it takes to conjugate a verb, he can produce a new lexeme. (Boložky 1999:xi).

Hebrew derivation is not as predictable as Hebrew inflection, of course, but it comes very close. The wide gap between inflection and derivation that European languages typically display is not present in a Semitic language such as Hebrew.

It is well established that Hebrew derivational morphology is extremely productive and transparent to speakers (Boložky 1999; Ephratt 1997). Therefore, in dealing with Hebrew, this paper will focus not on the question of psychological opacity, but rather on circumstantial opacity. It has been generally assumed that transparency in Hebrew morphology is directly correlated with systematic productivity. The question remains: how does the psychological transparency of a particular lexical item's components survive phonological changes that create circumstantial opacity? Is there circumstantial opacity in current day Israeli Hebrew?

Hebrew phonology has historically had the following stop phonemes /bdg ptk/ which were realized as their corresponding fricatives post-vocally, and as stops everywhere else. There has been some erosion of this rule since the revival of Hebrew, in that /t/, /d/, and /g/ are never realized as fricatives regardless of environment. Recently there has also been much inconsistency in the realization of /b/, /p/ and /k/ as stops or fricatives. Productively, the stops are not fricativized, but they appear as fricatives in many frozen expressions, including many derived lexemes. This is a situation which ordinarily leads to merger of phonemes and a high level of homonymy among roots. The expected result of such changes is circumstantial opacity of roots.

Thus a lexeme such as /miknasajim/, meaning 'pants', realized as [mixnasajim], may not be circumstantially interpreted as being related to a word such as /kneset/, meaning 'parliament', realized as /kneset/, even though both historically come from the root /k-n-s/, meaning 'to enter'.

The approach to circumstantial opacity adopted here is somewhat akin to the issue of biuniqueness in phonology. Certain phonological theories assume a top down approach, looking at phonemic status from the point of view of the speaker. Requiring biuniqueness in a phonology, on the other hand, means looking at phonemic status from the point of view of the hearer. If the hearer has no way to determine from the sounds whether the [x] of [mixnasajim] is a *het* or a *kaff*, then under a biunique theory of phonology, all the [x] sounds have merged into a single phoneme, and the words derived from the root /k-n-s/ will also be split according to whether the *kaff* in a particular derivation is postvocalic or postconsonantal.

In terms of psychological transparency, such a split has not occurred. Speakers still experience the root /k-n-s/ as an accessible morpheme. They are able to generate new lexemes using this root and familiar templates, even though they may falter in deciding whether to realize the /k/ as a [x] or a [k]. Likewise, when they encounter

an established lexeme, the root /k-n-s/ has psychological reality for them, despite the circumstantial opacity created by phonological change. In fact, it is well known that Hebrew displays a concurrent semantic transparency in the face of phonological opacity (Barkai 1978, 1980). Speakers know what a root means and recognize its occurrence, even if they are a little fuzzy as to the phonemes of which it consists. Circumstantial opacity does not prevent them from experiencing psychological transparency.

The apparent implication for lexicalization is this: lexical status is not dependent on componential opacity. In turn, componential opacity or transparency seems to be a systemwide psychological phenomenon. Opacity is not necessarily brought about by the reanalysis of phonological boundaries. Sometimes it is a change in the entire system that renders a particular lexical derivation opaque, despite the circumstantial accessibility of its component parts. Speakers can perhaps tolerate a high level of circumstantial opacity, if derivation in the language as a whole is systematic and predictable.

6. CONCLUSION. Functional research in grammaticalization has equated lexicalization with low productivity, fusion and componential opacity. This paper has explored the possibility, using Chinese, Hebrew and English data, that circumstantial opacity and psychological opacity do not always correlate. We have seen that a lexeme can be psychologically transparent while circumstantially opaque (Hebrew), and conversely, that a lexical item can be circumstantially transparent but psychologically opaque (English). Likewise, Chinese morphological components of compound words were found to behave sometimes as more transparent and sometimes less, without regard to phonological fusion. Thus, the idea that fusion or opacity are predictable correlates of lexemic status has been shown to be questionable.

Lexicalization, as a process, can occur as a result of reanalysis, but reanalysis itself does not necessarily lead to componential opacity. The opacity or transparency of the components of a word may be a psychological effect brought about by language-wide phenomena. Perhaps speakers are conditioned by the language as a whole to use componential analysis as a productive strategy or to fail to make such use.

The above observations are preliminary in nature. I am currently working on an experimental research project funded by the National Science Council of Taiwan. My investigation of componential transparency and opacity is ongoing. Analysis of the data I am currently collecting may eventually support or disprove the initial hypotheses set forth in this paper.

REFERENCES

- BARKAI, M. 1978. Phonological opacity vs. semantic transparency: two cases from Israeli Hebrew. *Lingua* 44:363–78.
 ———. 1980. Aphasic evidence for lexical and phonological representation. *Afro-asiatic linguistics* 7:163–1187.

- BATES, ELIZABETH & SYLVIA CHEN. 1991. The noun-verb Problem in Chinese aphasia. *Brain and language* 41:203–33.
- BERK, LYNN M. 1999. *English syntax: From word to discourse*. Oxford: Oxford University Press.
- BOLOZKY, SHMUEL. 1999. *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: E. J. Brill.
- CHIANG, OSCAR. 1999. More about names. *Time for students* 15283:10–11.
- CHOMSKY, NOAM. 1966. *Topics in the theory of Generative Grammar*. The Hague: Mouton.
- EPHRATT, M. 1997. The psychological status of the root in Modern Hebrew. *Folia Linguistica* 31:77–103.
- HOPPER, PAUL & ELIZABETH TRAUGOTT. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- HEINE, BERND, ULRIKE CLAUDI & FRIEDERIKE HUNNEMEYER. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
- LI, CHARLES & SANDRA THOMPSON. 1998. *Mandarin Chinese: a functional reference grammar*. University of California Press.
- PYLES, THOMAS & JOHN ALGEO. 1993. *The origins and development of the English language*. New York: Harcourt, Brace Jovanovich.
- TRAUGOTT, ELIZABETH & BERND HEINE. 1991. *Approaches to grammaticalization*. Amsterdam: John Benjamins.
- ZHANG, XIUHONG. *Evidence for a neurological distinction between nouns and verbs*. ms.



