Formulaic expressions in mind and brain: Empirical studies and a dual-processing model of language competence

Diana Van Lancker Sidtis, Ph.D., CCC/SLP

New York University

New York, New York 10012

and

Nathan Kline Institute for Psychiatric Research

Orangeburg, CA 10962

Formulaic expressions in mind and brain: Empirical studies and a dual-processing model of language competence

Background and purpose

After a long period of neglect and misunderstanding, formulaic language has finally come into its own (Coulmas, 1994; Cowie, 1998; Kuiper, 2004; Pawley, 2007; Wray, 2002, 2008a). Neglect arose from the myopic perspective that the speech formulas, idioms, and other conventional expressions known to a language community take the form of a not very interesting look-up list. As for misunderstanding, attempts to wrestle formulaic language down using generative linguistic approaches resembled, in the opinion of many today, trying to force a square peg into a round hole. These approaches not only failed to yield valid or useful descriptions; they also distorted the picture of a very large, very vibrant sector of language competence, which is worthy of examination on its own terms. Observations of formulaic language in diverse discourse contexts have greatly increased our perspective in the past decade, and theories have begun to mature. Much of this growth is attributable to the burgeoning interest in pragmatics--language use in everyday settings, and to the embracing of spoken text staunchly undertaken by sociolinguists (e.g., Schegloff, 1988; Tannen, 1989).

Our purpose in this article is to review progress in our understanding of formulaic language, beginning with a background discussion and continuing with a review of studies of normal use and incidence. Because many categories of formulaic language flourish in actual spontaneous language use, and because production material is the most difficult to capture and quantify, we focus on spoken texts, corpus studies, and discourse

analyses, supplemented by formal studies that attempt to probe speech production. The overview will then examine formulaic language in disordered speech. In the conclusion we propose a dual-process production model that accommodates both formulaic and novel language in a neurolinguistic context, whereby the accountable brain structures are described.

Definitions & Theory

Formulaic expressions carry meaning as whole units whose individual word parts do not necessarily reflect the meaning of the whole expression (Van Lancker Sidtis, 2004; Wray & Perkins, 2000). For example, the idiom "He put his foot in his mouth" has figurative meaning, as does the expression "That broke the ice." In the latter expression, the signifiers have more direct correspondence to the signified and so can be said to have more semantic transparency than the former expression. There are other forms of non-propositional expressions that are more transparent in terms of literal meaning. One category is speech formulas, for example "How are you?" or "After you" and such proverbs as "Better safe than sorry". Table 1 provides examples of formulaic expressions, each in a corresponding linguistic category:

Table 1. Non-propositional categories with an example of each (adapted from Van Lancker Sidtis, 2004)

Clichés: The pursuit of happiness
Conventional expressions: Pleased to meet you
Expletives: Gosh darn it.
Familiar proper nouns: Elizabeth Taylor
Indirect requests: It's awfully warm in here
Memorized expressions: i.e. lyrics, prayers, nursery rhymes
Pause fillers: Uh, um, like
Discourse elements: So, well
Proverbs: Look before you leap
Sentence stems: I'd like you to meet…
Serial speech: i.e. numbers, alphabet, days of the week
Slang: dead cert (Br.), far-out (Am.)
Speech formulas: See you later!

The best operational definition for formulaic expressions is an exclusionary one, applying "nonnovel" as the selection criterion; that is, formulaic expressions have in common that they are <u>not</u> newly created from the operation of grammatical rules on lexical items. They are holistically acquired and used in a language community based on shared knowledge of the stereotyped, canonical form, the conventionalized meaning, and conditions of use. The list of generally accepted categories has become well known: idioms (for which the lexical items do not carry their usual meanings: "She has him eating out of her hand"); speech formulas (conventional expressions used in specified

conversational contexts: "I'll get back to you later"); proverbs, expletives and exclamations ("Oh, my God," "Wow"); pause fillers ("uh, um"), discourse particles ("well" "so", "then"), conventional expressions ("that being said"), sentence stems ("I think"), and indirect requests ("it's getting warm in here"). Some treatments of formulaic language include nonreversible dyads and triads ("salt and pepper," "red, white and blue"), verb plus particle constructions ("back up" as in an argument), and specialized compound nouns ("truck driver"). These all have in common that they are not newly created but are produced and recognized by native speakers as unitary in form and meaning and conventionalized in usage. What remains to be further understood are the differences between these and other subtypes of formulaic language.

Stereotyped form, the first characteristic mentioned above, means that formulaic expressions, in their canonical form, contain precisely specified words in a certain word order spoken on a set intonation contour. Secondly, the meanings of formulaic expressions are conventionalized, which means the semantic aspects are idiosyncratic in various ways; they may be nonliteral, serve mainly as social signals, and/or they communicate a meaning that is greater than the sum of their parts—the special innuendos (Wray, 2002), which often deliver strong doses of affect and/or attitude. And third, as formulaic language lives in the realm of pragmatics, conditions of use constitute the third defining characteristic.

The flexible formuleme

A property of formulaic expressions that has wreaked havoc in linguistic circles is, paradoxically, their flexibility: many variants can and do appear. The canonical form

can be, and usually is, altered in actual usage. We argue below that the goal to discover cogent and explanatory generalizations underlying these variations, which will hold for a large set of formulaic expressions, has not been especially enlightening. It is more reflective of the situation to reference a canonical form or formuleme, and to accept that any alteration conforming to acceptable grammatical possibilities in the language is possible, as long as the canonical form remains recognizable. People intend for formulaic utterances to be recognizable, and they can allude to them in many ways. In reviewing versions of formulaic utterances given as "unacceptable" ("starred") exemplars in Kuiper & Everaert (2000), we could imagine playful or creative contexts allowing any of these. To alter formulaic expressions is merely to engage in theme and variation, one of the oldest and most pervasive practices in art and culture.

Throughout this article, we argue that people are free to apply linguistic rules on formulaic expressions for any purpose: communication, verbal playfulness, performing a new twist on a conventionalized meaning, allusion to a previous instance of a formula, conflation of two formulas, and so on. The speaker's meaning under these situations will be determined by linguistic and social context, as has been found in earlier studies (Gibbs, 1981). It is the context bound, creative basis of formuleme variation that has led to inconsistent results in the quest for linguistic regularity underlying versions of idioms and other kinds of formulaic language. From anecdotal observation, we submit that "acceptability" parameters for formulaic expressions work from particularly soft constraints, meaning that acceptability judgments are highly subject to numerous linguistic and contextual effects (Sorace & Keller, 2005).

Formuleme familiarity

As stated by Jackendoff (1995), a very large number of a broad range of formulaic expressions "*are familiar to American speakers of English; that is, an American speaker must have them stored in memory*" (p. 135). Formulaic expressions are "familiar" in the sense that a native speaker will recognize them as having this special status. For example, the sentence "He knew already a few days ago about the anniversary of the first actual moon landing" has probably not been said or heard before; this sentence has none of the properties of familiarity and predictability mentioned above (Pinker, 1995). On the other hand, the expressions "*Leave me alone*!" or "*You've got to be kidding*" or "*You don't say*!" (speech formulas); "*David spilled the beans*" or "*She's a real snake in the grass*" (idioms); and "*A bird in the hand beats two in the bush*" or "*The early bird catches the worm*" (proverbs) are all "familiar," in that native speakers recognize these utterances—they say that they do, and they demonstrate knowledge of their specialized meanings and appropriate contexts. This knowledge is demonstrated in more ways than can be enumerated here: ubiquitous allusions to formulaic expressions in the public media, echoing use in conversations, text, field and survey studies, and studies of word association. For example, people in a language community know not to say "See you later" on first entering a room; they do not give a "good morning" greeting twice in a row, when passing in the hall, to the same person." As an operational test, native speakers of a language accurately fill in blanks in such utterances, when key words are omitted (Van Lancker Sidtis & Rallon, 2004). Approximately a third of word association results is attributable to the "idiom effect," which means that given a target word, subjects produce words from a known formulaic expression (Clark, 1970).

Thus the effort to explain variants misses the main point: the stereotyped form (or "formuleme"), its conventional meaning and its conditions of usage are known to the language community. Any grammatical alteration (passivizing, movement, insertion, and so on)—that could legally apply to the corresponding nonformulaic structure--is possible in the communicative context, so long as enough of the formuleme remains identifiable and the context allows for correct interpretation. The key feature of formulaic expressions is their personal familiarity: people know them. Their status as common knowledge in a linguistic community forms the major portion of their raison d'être. This is a matter of native speakers' intuitions, which stand as valid linguistic evidence (Devitt, 2006).

Functions of formulaic language in everyday discourse

The pervasive nature of these types of expressions in our daily repertoire of communication has been examined in studies documenting the function and frequency of formulaic language in conversations and in literary texts and films (Van Lancker Sidtis, 2004). In comparison to isolated novel expressions, an utterance that contains a formulaic expression imbues the communication with emotional and attitudinal weight (Van Lancker Sidtis, 2008). Certain function-based divisions of formulaic language reflect these features; for example, thanking, apologies, requests and offers (Aijmer, 1996; Wray & Perkins, 2000). Others, falling under the heading of social interactions, include conversational maintenance and purpose (e.g. "How've you been?" "What's the occasion?" "I'm very sorry to hear that." "Really?" "No kidding!"). Formulaic expressions are used to achieve special purposes in communication (Tannen & Öztek,

1981), such as structuring talk (Fox Tree, 2006; Jucker, 1993) negotiating complaints (Drew & Holt, 1988); partnership solidarity (Bell & Healey, 1992; Bruess & Pearson, 1993), maintaining fluency in various contexts such as sport, weather forecasting, horse races, in the workplace, and auctions (Kuiper, 1991, 1992, 1996, 2006; Kuiper and Flindall, 2000; Kuiper and Haggo, 1985) and generally sounding like a native speaker of the language (Fillmore, 1979; Pawley and Syder, 1983; see review in Van Lancker, 2004). Because they are not newly formed, formulaic expression free up resources in the speaker to look ahead to the next part of the encoding process (Wray and Perkins, 2000; Kuiper, 2000). Production of ready-made phrases allows for more consistent rate of speech and greater fluency. It might be said that formulaic expressions buy time during speech production.

Dissecting these types of expressions reveals that they are tools with a purpose that includes, but also transcends, the act of conveying factual information. Vivid examples of this quality can be found in the formulaic expressions identified, classified and validated by Van Lancker and Rallon (2004) from the screen play “Some Like it Hot”: “Cut it out”; “I forgive you”; “I’m afraid not”; “Then hit’m with everything you’ve got"; “Quit stalling”; “It’s gone to his head”; “They wouldn’t be caught dead"; you got a lot of nerve. We note their properties of stereotyped form, conventional meaning (including affective and attitudinal meanings), and restricted pragmatic conditions. Native speakers recognize them as familiar and know implicitly how they function in conversation.

Quantity of formulaic expressions in everyday discourse

The actual number of these expressions is debated. Persons accumulating lists have not seen an upper limit. While experts trawling for formulaic expressions in the sea of language differ somewhat in how broadly to throw the net to establish formulaic categories, all agree that the number making up a typical speaker's repertory exceeds many tens of thousands, and some say the totals are upwards of hundreds of thousands (Kuiper, personal communication). All kinds of formulaic expressions occur in daily speech a great deal of the time, but a compiled sum remains difficult to come by at the present time, in part because analytic efforts examining speech samples have been fragmented and piecemeal.

Naturalistic speech is a notoriously difficult beast. Yet to understand formulaic language, our interests turn to the kinds of expressions and how many of them; in what social and linguistic contexts do they appear, by whom are they spoken and how frequently do they appear, and with which themes and discourse styles are they associated. This requires qualitative and quantitative examination of natural speech. Some field studies of incidence have appear. Jay (1980) tabulated use of cursing in specific populations (e.g., college students) and Gallahorn (1971) kept records of expletives emitted during team meetings of health care professionals on a psychiatric ward.

A brilliant contribution utilizing the fieldwork method came from Mathilde Hain (1951), who recorded over 350 proverbs in actual use among people who lived in a small German village, Ulfa, between the years of 1938 and 1943. Hain, a folklorist using sociolinguistic, field study methods, was interested in documenting veridical,

spontaneous production of proverbs, and in simultaneously recording the linguistic and social contexts of these specialized utterances. In this way, she could concern herself with quantity—how many different proverbial utterances appeared in the years that she lived in the village performing field work, and quality: how was each utterance used to communicate in the unique context? She examined the time of appearance of a proverb in the conversational setting (early, late, toward the end) and how the utterance was dealt with by the speaker and the listener. She set about to record the function of proverbs in everyday life as well as their position in a discussion. Here are some speech situations from the Ulfa village life recorded by Hain (1951, pp. 26-27 (trans.)). In the second example, the same proverb appeared in two settings:

1. The baptism of an infant with about 30 guests… Two older farmer women talk about the local young teacher, his performance in school and his meager salary. Mrs. X., especially, has a lot of information; she knows him and his modest means. X.: *Er ess hoard gscheit!* (er ist sehr klug) (Engl: he is very smart). At first K. listens pensively, then she responds slowly: *Aich saan als, wer de Hoawwern vedint hot, kritt en näid* (I always say, those who deserve the oats don't get them). Mrs. X. energetically agrees: *Joa, so ess!* (That's for sure!). And so this topic has ended; the individual case has become part of the general understanding.

2. The old dirt farmer Ch., in his late 70s, who always has a joke at hand, tells an acquaintance from the city about his daughter-in-law's long illness, including how long they tried to heal her and how much money they spent! It would not have been necessary if she would have gone for x-rays in Giessen right away. B. says: One only knows in hindsight. As he is getting ready to leave, Ch. replies loudly, so that you can even hear it across the street: *Joa, wann's Kend gehowe ess, gitts Gevadderleut* (ja, wenn das Kind aus der Taufe gehoben ist, gibt's Gevatterleute) (Engl: yes, when the child is baptized there will be godparents)!

I heard the same proverb in other situations. A dirt-farmer of about 50 years told me about her misfortune with her cows. When another one was infected with brucellosis the year before, she bought insurance for all animals. She ended her story calmly, almost wearily: *Wann's Kend gehowe ess, gitt's Gevadderleut!* In both situations the proverb offered a way out into a consoling generic situation.

Hain's labor-intensive study provides a rare record of ordinary usage of one particular category of formulaic expressions. From the examples provided throughout her fieldwork record, it is clear that each time a proverb is used, the speaker has an expectation that its stereotyped form and conventionalized meaning will be recognized by

the listener(s) as part of their linguistic knowledge. Meaning's were not glossed or explained. We reiterate three main facts about formulaic familiarity that must not be lost or ignored in any treatment: people know them; speakers using them expect listeners to know them and to appreciate the meaning conveyed; and, indeed, mutual familiarity with the specialized expressions forms an essential part of the reason for using them.

In the past few decades, discourse samples, or corpora, have been examined, using spoken samples and written texts. With the benefit of computerized searching, various texts have been analyzed using different algorithms to count incidence of formulaic expressions. Using a mathematical standard only, Altenberg (1998) listed 3-word combinations that occurred ten times or more in the London-Lund Corpus (Greenbaum & Svartvik, 1990). Other studies have utilized a human interface to classify utterance types. Some analyses have focused on a particular type of formula, such as proverbs. Cowie (1992) performed a study on "multiword lexical units" in newspaper language, differentiating idioms from collocations of various kinds. In an extensive treatment, Moon (1998) performed a descriptive study of formulaic expressions and idioms in an 18 million-word corpus of contemporary English, the Oxford Hector Pilot Corpus (Glassman et al., 1992), augmenting her analysis from other text sources Norrick (1985) reports only one complete proverb, plus a few proverbial allusions, in the 43,165 line corpus transcribed conversation published by Svartvik and Quirk (1980). A comparative frequency count of proverbs in French and English conversational corpora is described by Arnaud and Moon (1993).

Some quantitative and qualitative data on formulaic expressions come from studies of literary texts, especially in oral literature, (Kiparsky, 1976; Kuiper, 2000).

Tilley counted proverbs in the plays of Shakespeare (Tilley, 1950). Schweizer (1978) listed 194 idioms in 2,876 pages of six novels of Günter Grass, yielding an average of 14.8 idioms per page. Lord (1960) analyzed formulaic language in the Odyssey and the Iliad. For the qualitative contribution, numerous literary devices throughout Grass' writings involve idiomatic forms and meanings. The plays of Ionesco utilize an abundance of speech formulas to artistic effect as discussed by Klaver (1989). In his study of Homer's Iliad, Page (1959) estimates that about one-fifth of the poem is "*composed of lines wholly repeated from one place to another*" (p. 223), and that within the Iliad's 28,000 lines, there are approximately 25,000 repeated phrases. In the screen play "Some Like it Hot" idioms, proverbs, and speech formulas constituted 25% of the total number of phrases in the text (Van Lancker and Rallon, 2004).

Psycholinguistic studies

Although this article is focused on production of formulaic expressions, a brief look at comprehension studies of idioms offers some understanding of what makes formulaic language production special. Using idioms, speakers convey ideas and emotions using words that do not refer to their usual lexical meanings. "She's skating on thin ice" can be said without directly referring to skating, but to convey the idea of risky behavior. How the listener apprehends idiomatic meaning remains mysterious. Three models proposed to explain this process are literal-first (serial) processing, literal and idiomatic (parallel) processing, and direct access of idiomatic meaning (depending on the context). These models differ in their use of the notion of compositionality, as some are based on the assumption that idioms are not "composed," but are processed as cohesive

unitary items. The foundation for this idea comes from studies showing that people remember idioms as chunks rather than composite forms (Osgood & Hoosain, 1974; Pickens and Pollio, 1979; Horowitz & Manelis, 1973; Simon, 1974), a result also shown for Chinese idioms (Simon, Zhang, Zang, and Peng, 1989).

The two models of noncompositionality are The Idiom List Hypothesis and The Lexical Representation Model. The Idiom List Hypothesis (Bobrow & Bell, 1973) proposes that idioms are lexical items stored in memory and that upon encountering an idiom, the comprehension device first attempts a literal interpretation. After failing, the idiom retrieval mode kicks in and the idiom is selected from the look-up list. Thus, serial processing predicts greater response time latencies for idioms than for literal utterances because literal interpretation is the first step in any language task. A number of later studies refuted these findings by showing that visual classification of idioms is faster than literal phrases, forming the basis for the Lexical Representation model. Originally formulated by Swinney and Cutler (1979), the Lexical Representation model suggests that idiomatic meaning is processed in parallel with literal meaning, and that idioms are stored and retrieved whole, accounting for the faster reaction times.

However, the question remains as to how idioms can undergo syntactic and semantic modifications and maintain their pragmatic identity. Different studies have addressed these two separate underlying issues: Are idioms stored and accessed as whole units, and/or are their individual words and syntactic form taken into account during the retrieval process? Experimental approaches to these questions involve measuring production errors, response time, accuracy in recall and recognition memory tasks, and

various kinds of rating surveys. Fixed expressions that can undergo syntactic modifications yet maintain their conventional meaning are said to have syntactic flexibility (or productivity). For example, one could say "For years, she had been skating on really thin ice", and listeners, depending on context, could assume the nonliteral meaning. The notion "degree of compositionality," derived from native speakers' ratings, has been proposed to determine the flexibility of an idiom's comprehension (Gibbs et al, 1989; Gibbs & Gonzales, 1985). This finding is interesting although the reliability of the approach has been questioned (Titone & Connine, 1999; Cacciari & Tabossi, 1988). An idiom's "degree of compositionality" is likely not an all or none property, but instead viewed as a falling along a continuum, highly influenced by context. In addition, ratings in these surveys depend greatly on task instructions. Further, linguistic and situational contexts facilitate several aspects of nonliteral language comprehension, including how rapidly formulaic expressions (e.g. indirect requests and idioms) are processed, how they are interpreted (literally or figuratively), and how accurately and quickly they are retrieved from memory (Gibbs, 1980; Gibbs,1981). Social context plays a major role in judgments or perceptions of lexical transparency and syntactic flexibility.

The various findings of syntactic and lexical flexibility have led to hybrid psycholinguistic models. The earliest example is the Configurational model (Cacciari & Tabossi, 1988), which integrates literal-first and idiomatic-only approaches. This states that initially the idiom's potential literal meaning is activated until a key word is encountered which unlocks the idiom's figurative meaning. The idiom meaning is encoded within a specific word configuration that has weighted connections between its

lexical nodes. Idioms are considered to have either high or low predictability, depending on how early the key occurs in the string.

Several creative approaches to the production mode have been designed. From studies using speech-error elicitation experiment (Cutting and Bock, 1997), the authors conclude that idioms are not frozen or devoid of information about their syntactic and semantic structure. Interference in the form of blending errors was more likely to occur between idioms sharing the same syntactic form, and resultant word substitutions were in the same grammatical class. Another experiment showed that idiom production can result in activation of corresponding literal meaning. The authors put forth the argument that what is special about idioms is their relationship to a conceptual representation, which is a nonlinguistic entity, and that in normal language production, idioms are not special. Another study using a priming technique, in which a probe word is used to influence later on-line processing of a test word, also concluded that words in idioms stimulate word association networks as readily as words in literal sentences (Smolka, Rabanus, & Rösler, 2007).

We agree that syntactic and semantic operations can be performed on any formulaic expression, and that speakers can discern grammatical form and lexical meaning in these expressions. However, the additional properties of formulaic expressions reviewed above—among which are familiarity, stereotyped form, conventional, contextual-based meanings, and conditions of use are pragmatic-linguistic factors, which must be recognized in a veridical model of language competence. These are revealed in surveys, sentence completion studies, and association studies (Van Lancker Sidtis & Rallon, 2004; Clark, 1970). A series of idiom production experiments

by Sprenger (2003) using a priming paradigm with response time measurement addressed this question, yielded the notion of superlemma, which corresponds to the idea proposed in this article of canonical form or formuleme. Similarly, analysis of speech errors involving idioms suggest that "idioms are both compositional and noncompositional at the same time, at different levels of processing" (Kuiper, van Egmond, Kempen & Sprenger, 2007, p. 324). This superlemma or formuleme is stored and processed as a whole, but because it contains syntactic and semantic information of various kinds (as does any linguistic entity in the speaker's repertory), it can link to other parts of the lexicon and grammar.

The "dual model" of language competence, which has been previously mentioned and is proposed in this review, accommodates these findings (Wray and Perkins, 2000; Van Lancker Sidtis, 2008). The inconsistencies arising from idiom studies can be explained by the interplay of formulaic and novel processing. The dual-processing model states that formulaic expressions, which by definition have stereotyped form, conventionalized meanings, and are familiar in a language community (Kitzinger, 2000), exist in harmony with the grammar, which consists of rules and a lexicon. Formulaic expressions, depending on the intent and verbal creativity of the speaker and context, can be altered using standard grammatical processes. Because of this flexibility and context dependency, we submit that task demands in psycholinguistic approaches to the processing of formulaic expressions exert an overriding influence on the results. When subjects are involved in metalinguistic decisions about the interplay between formulaic and novel expressions, their performance will be highly influenced by contingencies of the experimental setting. The idiom studies do not lead to the conclusion that formulaic

and novel language are processed in the same way, but rather show that grammar can operate on any legitimate utterance, and that standard lexical meanings can be discerned in any phrase. Again, the interesting point is not that formulaic expressions can be semantically transparent and composed and are therefore alterable, but that there is a known entity to analyze and to alter.

Auditory contrasts of idiomatic and literal meanings

Another set of studies highlights speakers' knowledge of differences between literal and idiomatic utterances. A series of studies using American English ditropic sentences, those having both idiomatic and literal meanings, such as "He was at the end of his rope," showed that listeners can distinguish between these kinds of meanings from the acoustic signal alone. Listening studies confirmed the discriminability of these utterance types, and acoustic measures of rate, pitch mean and variability, and terminal pitch revealed significantly contributory auditory cues. These observations implied that native speakers articulate literal and idiomatic utterances differently, formulating consistent and stable auditory-acoustic cues that listeners use to distinguish between the two meanings (Van Lancker, Canter, & Terbeek, 1981). Later studies examined utterances in French and Korean, revealing differences and similarities in the use of acoustic-phonetic cues to convey contrastively idiomatic and literal interpretations of ditropic sentences (Abdelli-Baruh, Yang, Ahn, & Van Lancker Sidtis, 2007). It was found that Parisian French and American English speakers utilized the same cues to distinguish the sentences, but in an opposite manner: French idioms were significantly longer, while for English, the literal versions were longer. Pitch was also differently utilized to signal the meaning differences. When Korean ditropic sentences were

analyzed, it was seen that native speakers of Korean used duration, amplitude variations, and contrasting fundamental frequency in the last two words to mark utterances as either literal or idiomatic in Korean. Literal utterances have longer durations than idiomatic utterances, whereas idiomatic utterances are more varied in amplitude than literal utterances. Intonation contours differ for the two types of meanings: Literal utterances end more often with falling pitch whereas idiomatic utterances end with rising pitch. Thus formulaic and literal meaning contrasts, in utterances where the words and grammar do not differ, are successfully signaled by various acoustic-auditory cues known to the native speaker community.

In summary, many psycholinguistic studies have exploited the fact that formulaic expressions can be manipulated metalinguistically, but these studies have not yielded an consistently explanatory model. Demonstrating the interplay between formulaic and novel expressions using various performance tasks, some researchers have concluded that mental processes for these two types of utterances do not differ. As mentioned above, this conclusion misses the point of the status and value of formulaic expressions in everyday usage, and fails to acknowledge the dual-processing feature of language competence.

Disordered language in neurological conditions

A neurological account of formulaic language as a class of its own can be found in clinical observations of aphasic language (Espir and Rose, 1970). In clinical domains, the notion of formulaic language traces back to the well-known concept of "automatic speech" in aphasia, first identified and described by J. Hughlings Jackson (1874).

Automatic speech includes "overlearned" utterances such as counting, speech formulas (salutations and conversational fillers), swearing, nursery rhymes, familiar lyrics and familiar songs, and other such expressions (Van Lancker, 1993; Code, 1989; Van Lancker & Cummings, 1999). Preservation of certain kinds of speech is dramatic when experienced in severe cases of language disturbance following left hemisphere stroke: the afflicted person cannot articulate a novel phrase or sentence, or can do so only with extreme effort and poor articulatory success. In contrast, he or she can fluently swear, count to ten, recite nursery rhymes, and produce a set of fixed expressions such as *how are you, good bye, I don't know*. This well known fact remained anecdotally transmitted among clinicians until survey studies performed for English (Code, 1982, 1989), Chinese (Chung, Code and Ball, 2004) and German speakers (Blanken & Marini, 1997), which systematically documented preserved expressions in severe aphasia.

Several group studies of production abilities in aphasia support impressions about preservation of formulaic expressions. Using three paired tasks, Lum and Ellis (1994) compared speech production in formulaic versus propositional contexts. First, counting was compared to naming Arabic numbers in nonconsecutive order; next, naming pictures with cues from formulaic expressions (e.g., Don't beat around the BUSH) was compared to naming pictures depicting novel phrases (Don't dig behind the BUSH); and third, formulaic and novel expressions were compared in a repetition task. Subjects performed better on formulaic subtests for number production and picture naming, with a slight advantage also for phrase repetition. A similar finding (Van Lancker & Bella, 1996) arose from comparing matched propositional and formulaic expressions in aphasic speakers also in the repetition and sentence completion tasks, again with weaker differences results in the repetition task.

It is clear that formulaic language has its most authentic presence in spontaneous speech. It is therefore desirable to examine the naturalistic speech of persons with left or right hemisphere damage due to stroke to determine the effect of localized damage on use of formulaic expressions. Recently, a study of formulaic expressions in conversational speech of right hemisphere damaged subjects and left hemisphere damaged subjects compared to normal controls was undertaken (Van Lancker & Postman, 2006). The incidence of formulaic language was significantly higher in left hemisphere damage resulting in aphasia than in matched normal control speakers, while in right hemisphere damage, formulaic expressions were significantly diminished in comparison to normal speech. A striking finding was the low percentage in right hemisphere damaged subjects' speech of pause fillers, *um* and *uh,* which serve semantic and pragmatic purposes in communication (Clark & Fox Tree, 2002). For left hemisphere damaged persons, these results suggest that formulaic expressions become a vehicle for communication in aphasia, and for right hemisphere damaged persons, who do not have phonological, syntactic, or linguistic-semantic deficits, these results may help clarify the clinical impression of abnormal pragmatics of communication.

Case studies and other limited observations in neurologically impaired speech have implicated subcortical structures in the brain. In one case report, a loss of formulaic speech production abilities followed damage in the basal ganglia, namely a right caudate stroke (Speedie, Wertman, T'air, & Heilman, 1993). Another examination of two individuals with damage confined to subcortical nuclei, utilizing discourse obtained from structured interviews, revealed a significantly smaller proportion of formulaic expressions when compared to similar interview settings obtained from normal control speakers matched for

education and age. In that same study, the aphasic speaker utilized a much greater proportion than all other groups (Van Lancker Sidtis, Canterucci, & Katsnelson, 2008). In another instance, which involved a motor speech disorder likely due to a subcortical stroke, a pathologically intrusive syllable (*sis*) occurred with greater frequency during recitation, counting, and other formulaic expressions than in novel speech (Van Lancker, Bogen & Canter 1983). Still other differences between formulaic and novel language production have been seen in analyses of speech samples from Parkinson's patients, who have diminished basal ganglia function, revealing reduced formulaic language expressions when speaking (Illes, Metterb, Hansonb & Iritanib, 1988). Further, in Parkinson speakers, acoustic measures of voice and articulation, as well as listeners' ratings, differ significantly for novel and formulaic vocalization tasks (Sidtis, Rogers, Katsnelson, & Sidtis, 2008).

The few functional brain imaging studies examining formulaic language have been inconsistent, and findings are not always in agreement with well established information derived from clinical observations and lesion studies. Earlier studies of cerebral blood flow using SPECT methodology associated bilateral hemisphere activation with automatic speech Larsen, Skinhoj, & Lassen, 1978; Ryding, Bradvik, & Ingvar, 1987), but with the proliferation of functional imaging of language studies, bilateral signal is reported for most language tasks, and the early SPECT results are no longer interpretable. Published studies of language processing in general typically report bilateral hemispheric blood flow responses, for reasons that are not yet well understood (Van Lancker Sidtis, 2006). More recently, Blank, Scott, Murphy, Warburton, and Wise (2002) reported bilateral activation for both propositional and automatic speech, which does not reveal an interesting contrast between the two language modes. Another study using PET imaging employed two automatic speech tasks: the months

of the year and the Pledge of Allegiance (Bookheimer, Zeffiro, Blaxton, Gaillard, & Theodore, 2000), compared to tongue movements and consonant–vowel syllable production. Continuous production of the Pledge of Allegiance showed activation in traditional language areas; reciting the months of the year selectively engaged language areas Brodmann areas 44 and 22. These studies did not examine counting, which has been the most widely used task in cortical mapping, and is the most frequent type of preserved aphasic speech. Van Lancker, McIntosh, and Grafton (2003) reported that counting and word generation differed in brain activity, with only word generation showing activation in Broca's area, or the left anterior frontal area, and counting associated with more diffuse brain activity, including some subcortical sites.

A case study a person with the diagnosis of "presenile dementia" was one of the first to suggest a special status for brain processing of formulaic expressions. A relative preservation of formulaic expressions was observed in a severely aphasic 59-year-old woman, who, after the onset of her illness, was never observed to produce a meaningful utterance, but could complete idioms and other familiar conventional expressions spoken to her with the last word missing (Whitaker, 1976). We studied an aphasic individual with a severe comprehension deficit and a diagnosis of transcortical sensory aphasia who was similarly unable to produce meaningful speech, but correctly completed 50% of idiomatic and other formulaic expressions presented verbally to him in a similar way (Van Lancker Sidtis, 2001). Idiom completion in another case of transcortical sensory aphasia was also reported by Nakagawa, Tanabe, Ikeda, et al. (1993).

Like these cases described above, observations in Alzheimer speech tend to corroborate a role of the basal ganglia in production of formulaic expressions. It is commonly

observed in the clinical setting that persons with considerable progression in the disease, with MiniMental State Examination scores as low as 7 (on a scale of 0-30), who have lost most cognitive capacity, continue to produce formulaic expressions, such as "Nice seeing you again," "Excuse me," and "Good-bye." Observations of conversational speech in Alzheimer's disease reveals a large proportion of formulaic expressions, although they are sometimes used inappropriately. An Alzheimer patient said "I haven't seen you for a while" to a stranger in the hallway while the two were waiting together for the elevator. In some speech samples, distortions of formulaic expressions appear, such as "But they were very good down there by me for," or a blend of two or more expressions, such as "put down my mind to it." The relative preservation of formulaic in comparison to novel expressions is likely attributable to the fact that Alzheimer's disease attacks the cortical layers, leaving the subcortical nuclei intact until very late in the disease, by which time the patient is mute.

Another example of preserved formulaicity in brain disease arises from observations of language behaviors in autism (Wray, 2008b). Clinicians and lay persons alike describe the frequent repetition of radio and television jingles taken from advertisements and serial shows by autistic and other developmentally delayed children (Prizant & Duchan, 1981; Cohen, 2002; Cohen & Volkmar, 1997; Bogdashina, 2005). Various forms of immediate and delayed echolalia, the repetition of verbal utterances of self or another speaker, are well known to occur in these children, alongside absent or severely impoverished spontaneous, novel speech production (Lord & Rhea, 1997). Prizant (1983) postulates that in many cases, echolalia becomes a communication strategy compensating for an inability to fully process language. On a more psychological level, stereotyped or excessively repetitive language may be the only means available for social contact (Howlin, 1997), which is often accompanied by similarly

mechanical and ritualistic behaviors (Paul, 2004). Studies have shown that echoed utterances may have communicative functionality, similar to that seen for formulaic expressions (Dobbinson, Perkins & Boucher, 2003). It has often been noted that many autistic spectrum disorder children exhibit a predominant development of holistic expressions with impoverished grammatical competence (Lord & Paul, 1997). Autistic echolalia has been referred to as the overuse of "gestalt language forms" that are normally seen in child language development (Prizant, Schuler, Wetherby, & Rydell, 1997). This perspective refers to a model of normal language development (e.g., Locke,1993, 1995; Peters, 1977, 1983; Wong Fillmore, 1979; Tomasello, 2003) in which two modes of language acquisition, holistic and analytic, unfold in different but interactive maturational schedules. As Dobbinson, Perkins & Boucher (2003, p. 305) suggest, "autistic formulaicity may be seen as the preferential use of a normative operation." In autistic children, the holistic process flourishes, while the mode of learning language that establishes grammatical competence is in varying degrees defective.

The relative preservation of formulaic expressions as Alzheimer's disease progresses to the severe state, while semantic functions and newly created spontaneous speech deteriorate, supports the notion of a dual-process model of language. Similarly, observations in autism reveal a more successful development of a holistic mode of language processing, in contrast to defective grammatical and semantic functions, again pointing toward the viability of a model of language that describes both holistic-configurational and analytic-compositional modes of processing in acquisition and use.

How are formulaic expressions acquired?

Another provocative source that supports the dual-process model arises from developmental language studies, in infants' first and in adult second language acquisition. As mentioned above, researchers in child language document acquisition of holistic "chunks" of speech which evolve into compositional structures. While unitary utterances are utilized by children early on, acquisition of formulaic expressions at adult levels lags behind acquisition of grammatical competence (Kempler, Van Lancker, Marchman, & Bates, 1999). This suggests that the two processes, holistic and analytic, perform different roles at different stages of language acquisition, and, further, that different maturational schedules are in play for novel versus formulaic language knowledge. Similarly, in adult second language acquisition, the difficulty posed by formulaic expressions is well known. It is likely that critical periods for native-like acquisition exist for various types of language competences, including for acquisition of formulaic expressions.

The question of acquisition of formulaic expressions in the native speaker has seldom been seriously posed. It is generally assumed that a major force in acquisition of formulaic expressions is frequency of exposure, such that many repetitions of a formulaic expression eventually make a lasting impression in the developing child's or adult's memory. A companion assumption is that, because our brains are finite, novel language is learned primarily by abstract rules. The point behind emphasizing the finiteness of our brains is that memory capacity is severely limited, when considering the very large set of phrases and sentences we command. We submit that both of these assumptions—frequency of exposure and limited memory capacity—are questionable, and fail to

provide a basis for either a viable model or for empirical observations in formulaic language competence.

The "slow exposure" assumption is most likely wrong because, as mentioned above, the number of known expressions is very high and most of them occur infrequently. Native speakers acquire precise forms and complex semantic and pragmatic meanings of these expressions; unlike newly created sentences, they are all learned "by heart." If memory is poor for exact verbal replicas, then very many exposures of each expression are required. It is not a logical possibility that a sufficient number of repetitions of 100,000 (or, as some researchers say, 500,000) exact forms and contexts are provided for each language user, adequate to satisfy the requirements of the incremental learning process, leading to the huge repertory of pristinely stored formulemes.

It follows from this point, then, the "highly limited memory" assumption is also misguided. As the first bit of evidence, we point out again that people know tens of thousands of multiword expressions. Of interest here are recent findings by Gurevich and Goldberg (submitted) showing a successful verbatim memory for language. In these studies, participants reliably recognized and recalled full sentences that they were exposed to only once. It is likely that this fundamental ability to capture verbal material exactly is greatly heightened by the properties of formulaic language, and that this ability follows a unique maturational schedule. Similar and related findings are reported for incidental retention of voice identity information, such as gender (Geiselman , 1977; Geiselman & Crawley, 1976). For subjects listening to spoken words, voice attributes were clearly retained in memory (Goldinger, 1996). Using a memory model called the

MINERVA 2, studies suggest that separate episodic memory traces are retained for each stimulus (Hintzman, 1986). These approaches reveal specificity of memory processes and allow for the possibility that brief exposures suffice to encode a large number formulaic expressions. These expressions differ from novel expressions in their relationship to grammatical and semantic functions, social context, and use. These significant differences likely provide cues to signal a different type of encoding and retention process.

We propose that a specialized form of knowledge acquisition may be operative for formulaic expressions. For example, something comparable to one-trial learning, a special case in classic learning theory, should be considered. In this type of learning, rather than acquiring a piece of information over through repeated exposure or trials, organisms "learn" (acquire a conditioned response) immediately or very quickly as the result of a "strong contingency" reinforcement (Lattal, 1995). Another kind of instantaneous "learning" that is well studied is imprinting. Brain biochemistry and localization for imprinting have been extensively investigated in birds (Horn, 1985; Knudson, 1987). It is known that forebrain hemispheres with strong connections to striatal (subcortical) and brainstem structures are operative in imprinting, and these structures are also involved in visceral and endocrine functions as well as emotional expression (Horn, 1985, p. 243). Genetic determinants control the basic neuronal circuitry in some species studied, but flexibility allows for experience to shape the perceptual system (Knudsen, 1987). There are critical time windows but these sensitive periods are often labile and flexible (Marler, 1998). Also worth mentioning as a rapid memory process are "flash bulb memories," in which an unusual amount of experiential

detail is retained in memory in association with receiving surprising information (Brown & Kulik, 1977). Features of the processes include arousal, surprise, and personal relevance, in addition to affect, attention, distinctiveness and poststimulus elaboration (Christianson, 1992). It has been suggested that this function is part of an automatically preattentive mechanism (Neisser, 1967). Arousal levels may arise from hormonal or other chemical influences (Gold, 1992), which results in immediate and detailed acquiring of a stimulus.

Our perspective is that acquisition or learning of formulaic expression occurs according to a unique learning process comparable to the three relatively instantaneous types described above. We propose that throughout the early lifespan, formulaic expressions are acquired in a manner quite different from processes involved in learning the rules and lexicon that underlie generation of novel expressions. (In view of the difficulties encountered by adult second language learners, it is likely that the rapid acquisition ability diminishes with age.) These three examples of extraordinary learning and memory functions are offered to provoke new ways of looking at acquisition of formulaic expressions. These examples of near-instantaneous acquisition of information all demonstrate the coordinated roles of procedural, episodic and declarative memory, as well as arousal and attention, rendering a special status to some kinds of knowledge acquisition. An alerting mechanism may be engaged by the fact that formulaic expressions pattern differently in the conversational setting and are tightly bound to social context. In summary, we propose that formulaic expressions, under specialized circumstances not yet understood, ascend quickly and suddenly into a native speaker's language competence.

Dual-Processing Model of Language Competence

Evidence for a dual-process model of language processing comes from several sources. Neurological damage can disturb, diminish or enhance formulaic language. Novel and formulaic language are affected differently by different types of brain damage: left hemisphere damage leads to selective impairment of novel language and relative preservation of formulaic language, while right hemisphere and /or subcortical damage lead to selective impairment of formulaic language, sparing novel language (Van Lancker Sidtis & Postman, 2006). Enhancements or selective presentation of formulaic language use are seen in aphasia, Tourette's syndrome, and Alzheimer's disease, while diminution is observed in right hemisphere and subcortical disease. It is likely that such differences will be more extensively and accurately documented as information about formulaic language is disseminated into clinical practice. Recognition of the important role of formulaic expressions in evaluation and recovery in aphasia and other neurological disorders has barely begun, despite the "automatic speech" tradition extending more than a hundred years into the past.

The notion of two such processing modes has emerged from studies of learning and memory, comparing, for example, procedural and declarative knowledge (Squires, 1986; Mishkin, Malamut & Bachevalier 1984). Subcortical structures have been associated with complex motor planning and execution (Lieberman, 2000; Baev, 1997; Marsden, 1982) "chunking of action repertoires" (Greybiel, 1998) and "habit learning" (Knowlton, Mangels, & Squire, 1996). These perspectives have been aligned with hierarchical levels of the central nervous system, such that automated motor gestures are

accommodated by subcortical structures, which developed phylogenetically earlier in human evolution (Koestler 1967). Correspondingly, it has been suggested that the origin of human language might be located in initial use of formulaic expressions (Jaynes, 1976; Code, 2005; Wray, 1998, 2000; Wray & Grace, 2007).

Relationship of formulaic language to animal vocalization and role in evolution

The functional importance of subcortical nuclei in formulaic verbal behavior in humans is especially interesting when considering studies of animal vocalization. For many decades, investigations of the neurology of nonhuman animal vocalization have identified subcortical sites in the initiation and production of calls. While formulaic utterances and animal calls are obviously different in many ways, they are similar in being formal and conventionalized and in having essentially social and emotional functionality. In humans, stereotyped vocalizations have occurred when subcortical sites are electrically stimulated during stereotaxic surgical techniques, usually for treatment of epilepsy, (Schaltenbrand, 1965; Petrovici, 1980). Hyperfunction of the basal ganglia/limbic system, as in persons with Tourette's syndrome, gives rise to semi-compulsive emotive utterances (called coprolalia or "foul speaking"). Conversely, as mentioned above, subcortical hypofunction due to stroke or degenerative disease is associated with diminution of formulaic expressions.

A system of neuronal organization in the central nervous system emerges from studies of vocalization with higher level control in subcortical structures in nonhuman primates, and with cortical representation occurring only in humans. In humans, two levels of nervous system control for vocalization may be described: an older, subcortical

system, which is capable of emotive and formulaic speech behavior; and a newer system that is cortical, unilateral, and involved in voluntary, novel and planned speech (Robinson, 1987; Ploog, 1975; Jürgens, 2002). Formulaic expressions might be seen as an evolutionary counterpart of nonhuman vocal behavior. Some have proposed that this evolutionarily older system also continues to perform in the emotional and routinized vocal behaviors seen in formulaic expressions (Jaynes, 1976; Jespersen, 1933; Patel, 2008; Wray, 2002; Code, 2005).

Thus studies in primates and humans lend credence to the notion that formulaic language in humans emerges from a evolutionary history different from that proposed for novel language, and is differently structured and controlled in the brain.

Conclusion

The formulaic phrase has unique properties: it is cohesive and unitary in its canonical structure, while easily subject to variation. Like the phoneme, the morpheme, and the lexeme (abstract classes of sounds, units of meaning, and words respectively), any emitted formulaic expression constitutes the instantiation of a formuleme. A formuleme may have aberrant grammatical form. It is often nonliteral or deviant in meaning properties and conveys a nuanced, connotational meaning that transcends the sum of its (lexical) parts. As an essentially pragmatic dimension of language, it is intimately tied to social context and has subtle specifications of usage. Most importantly, the canonical form of the expression, the formuleme, is known to native speakers. A formulaic expression functions differently in form, meaning and use from a literal, novel, or propositional expression (Lounsbury, 1963).

Despite scattered attempts at treatments over the years, formulaic expressions have not found a niche in generative theory (Kuiper, 2000). Our view is that this failure arises from the fact that fixed, familiar, unitary expressions are informed by principles and properties that cause them to be essentially different from newly created sentences. They cannot be handled by the same analytic apparatus. While there is no doubt that different types of formulaic utterances express the principles of formulaic language in different degrees, we reject the claim that formulaic and novel processes are processed in essentially the same manner in the language user (Gibbs, 1980, 1989, 1994). Instead, we have presented evidence that formulaic and novel language are more accurately characterized as two disparate modes of language competence. They have the ability to be in continuous interaction, and generative rules can operate on any formuleme, which can take any grammatically allowed variant providing it remains identifiable. Formulaic expressions are learned according to a distinctive maturational schedule that differs from learning principles that are operative for acquiring the rules of grammar, and they draw on instantaneous rather than incremental memory processes. These perspectives have relevance for models of language competence, language learning, and language loss in neurological disorders.

References

Abdelli-Baruh, Yang, Ahn, & Van Lancker Sidtis, (2007). Acoustic cues differentiating idiomatic from literal expressions across languages. American Speech-Language Hearing Association, Boston, MA, November 15-17.

Aijmer, K. (1996). *Conversational routines in English*. London, New York: Longman.

Altenberg B. (1998). On the phraseology of spoken English: the evidence of recurrent word-combinations. In Cowie, A. P. (Ed.), *Phraseology: Theory, analysis and application* (pp. 101-124). Oxford: Clarendon Press.

Arnaud, P., & Moon, R. E. (1993). Fréquence et emploi des proverbes anglais et français. In C. Plantin (Ed.), *Lieux Communs: Topoï, Stéréotypes, Clichés* (pp. 323-341). Paris: Kime.

Baev, K. (1997). Highest level automatisms in the nervous system: a theory of functional principles underlying the highest forms of brain function. *Progress in Neurobiology, 51,* 129-166.

Bell, R.A. & Healey, J.G. (1992). Idiomatic communication and interpersonal solidarity in friends' relational cultures. *Human Communication Research 3*, 307-335.

Blank, S. C., Scott, S., Murphy, K., Warburton, E. & Wise, R (2002). Speech production: Wernicke, Broca and beyond. *Brain, 125*, 1829-1838.

Blanken, G. & Marini, V. (1997). Where do lexical speech automatisms come from? *Journal of Neurolinguistics*, *10*, 19-31.

Bobrow, S., & Bell, S. (1973). On catching on to idiomatic expressions. *Memory and Cognition, 1*, 343-346.

Bogdashina, Olga. (2005). *Communication issues in autism and Aspergers syndrome. Do we speak the same language?* London: Jessica Kingsley Publishers.

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T. A., Gaillard, P. W., & Theodore, W. H. (2000). Activation of language cortex with automatic speech tasks. *Neurology, 55*, 1151-7.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition,* 5, 73-93.

Bruess, C. J. & Pearson, J. C. (1993). 'Sweet pea' and 'pussy cat': an examination of idiom use and marital satisfaction over the life cycle. Journal of Social and Personal Relationships, 10, 609-615.

Cacciari C., & Tabossi P. (1988). The comprehension of idioms. *Journal of Memory and Language*, *27*, 668-683.

Christianson, Sven-Åke. (1992). Do flashbulb memories differ from other types of emotional memories? In E. Winograd and U. Neisser, (Eds.), *Affect and accuracy in recall: Studies of "flashbulb memories"* (pp. 191-211). Cambridge: Cambridge University Press.

Chung, K. K. H., Code, C., & Ball, M. J. (2004). Lexical and non-lexical speech automatisms in aphasic Cantonese speakers. *Journal of Multilingual Communication Disorders, 2*, 32-42.

Clark, H. H. (1970). Word associations and linguistic theory. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 271-286). Baltimore: Penguin Books.

Clark, H.H. & Fox Tree, J.E. (2002). Using uh and um in spontaneous speaking. *Cognition 84(1),*73-111.

Clark, R. (1974). Performing without competence. *Journal of Child Language 1*, 1-10.

Code, C. (1982). Neurolinguistic analysis of recurrent utterance in aphasia. *Cortex, 18*, 141-152.

Code, C. (1989). Speech automatisms and recurring utterances. In C. Code (Ed.), *The characteristics of aphasia* (pp. 155-177). London: Taylor and Francis.

Code, C. (2005). First in, last out? The evolution of aphasic lexical speech automatisms to agrammatism and the evolution of human communication. *Interaction Studies, 6*, 311–334.

Cohen, Donald J., & Volkmar, Fred R. (Eds.) (1997). *Handbook of autism and pervasive developmental disorders.* New York: Wiley & Sons.

Cohen, Shirley. (2002). *Targeting autism: What we know, don't know, and can do to help young children with autism and related disorders.* Berkeley, CA: University of California Press.

Coulmas, F. (1994). Formulaic Language. In R.E. Asher (Ed.), *Encyclopedia of language and linguistics* (pp. 1292-1293). Oxford: Pergamon.

Cowie, A.P. (1992). Multiword lexical units and communicative language teaching. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 1-12). London: Macmillan.

Cutting, J. C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory-Cognition, 25(1)*, 57-71.

Devitt, M. (2006). Intuitions in linguistics. *British Journal for the Philosophy of Sci*ence, *57,* 481-513.

Dobbinson, S., Perkins, M. R. & Boucher, J. (2003). The interactional significance of

formulas in autistic language. *Clinical Linguistics and Phonetics*, *17(4),* 299–307.

Drew, P., & Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems, 35*, 398-417.

Espir, L., & Rose, F. (1970). *The basic neurology of speech*. Oxford: Blackwell Scientific Publications.

Fillmore, C. (1979). On fluency. In C. J. Fillmore, D. Kempler & W. S-Y Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85-102). London: Academic Press.

Fox Tree, J.E. (2006). Placing like in telling stories. *Discourse Studies: An Interdisciplinary Journal for the Study of Text and Talk, 8(6),* 723-43.

Gallahorn, G. E. (1971). The use of taboo words by psychiatric ward personnel. *Psychiatry, 34,* 309-321.

Geiselman, R. E. (1977). Incidental retention of a speaker's voice. *Memory and Cognition, 6,* 658-665.

Geiselman, R. E. and Crawley. J. M. (1976). Long-term memory for speaker's voice and source location. *Memory and Cognition, 4* (15), 483-489.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory and Cognition, 8 (2)*, 149-156.

Gibbs, R. W. (1981). Your wish is my command: Convention and context in interpreting indirect requests. *Journal of Verbal Learning and Verbal Behavior, 20*, 431-444.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding.* New York: Cambridge University Press.

Gibbs, R. W., Jr., Nayak, N.P., Bolton, J.L. and Keppel, M.E. (1989). Speaker's assumptions about the lexical flexibility of idioms. *Memory and Cognition 17 (1)*, 58-68.

Gibbs, R.W., Jr., & Gonzales, G.P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition 20*, 243-259.

Glassman, L., Grinberg, D., Hibbard, C., Meehan, J., Reid, L.G., van Leunen, M.-C. (1992). Hector: connecting words with definitions. *SRC Report 92a*; Digital Equipment Corporation Systems Research Center, Palo Alto, CA.

Gold, Paul E. (1992). A proposed neurobiological basis for regulating memory storage for significant events. In E. Winograd & U. Neisser, (Eds.). *Affect and accuracy in recall: Studies of "flashbulb memories"* (pp. 141-161). Cambridge: Cambridge University Press.

Goldinger, Stephan D. (1996). Words and voices: Episodic trace in spoken work identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166-1183.

Graybiel AM. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory 70*, 119-136.

Greenbaum, S., & Svartik, J. (1990). The London-Lund corpus of spoken English. In J. Svartik (Ed.), *The London-Lund corpus of spoken English: Description and research* (pp. 11-45). Lund: Lund University Press.

Gurevich, O., & Goldberg, A.E. Incidental verbatim memory for language. Submitted.

Hain, M. (1951). *Sprichwort und Volkssprache*. Giessen: Wilhelm Schmitz Verlag. English trans., in D. Sidtis & S. Mohr (Eds.), *Formulaic language in the field*, Anja Tachler, translator. Copyright.

Hintzman, Douglas L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*(4), 411-428.

Horn, Gabriel. (1985). *Memory, imprinting and the brain*. Oxford Psychology Series No. 10. Oxford: Clarendon Press.

Horowitz, L. M., & Manelis, L. (1973). Recognition and cued recall of idioms and phrases. *Journal of experimental psychology*, *100*, 291-296.

Howlin, P. (1997). Autism: *Preparing for adulthood*. New York: Routledge.

Hughlings Jackson, J. (1874). On the nature of the duality of the brain. In J. Taylor, (Ed.), *Selected writings of John Hughlings Jackson* (Vol. 2, pp. 129-145). London: Hodder & Stoughton.

Illes, J., Metterb, E.J., Hansonb, W.R., S. Iritanib, S. (1988). Language production in Parkinson's disease: Acoustic and linguistic considerations. *Brain and language, 33,* 146-160.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van der Linden, A. Schenk & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 133-166). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jay, T.B. (1980). Sex roles and dirty word usage: A review of the literature and a reply to Haas. *Psychological Bulletin, 88* (3), 614-621.

Jaynes, J. (1976). *The origin of consciousness in the breakdown of the bicameral mind*. Boston: Houghton Mifflin.

Jespersen, O. (1933). *Essentials of English grammar*. London: George Allen and Unwin, Ltd.

Jucker, A.H. (1993). The discourse marker 'well': a relevance-theoretical account. *Journal of Pragmatics 19*, 435-52.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26* (2), 235-258.

Kempler, D., Van Lancker, D., Marchman, V. & Bates, E. (1999). Idiom comprehension in children and adults with unilateral brain damage. *Developmental Neuropsychology, 15*(3), 327-349.

Kiparsky, P. (1976). Oral poetry: some linguistic and typological considerations. In B.A. Stolz, R. S. Shannon (Eds.), *Oral literature and the formula*. Michigan University, Center for Coordination of Ancient and Modern Studies: Trillium Pr.

Kitzinger, C., (2000). How to resist an idiom. *Research on Language and Social Interaction, 33*, 121–154.

Klaver, E., (1989). The play of language in Ionesco's play of chairs. *Modern Drama 32*, 521–531.

Knowlton, B., Mangels, J. & Squire, L. (1996). A neostratal habit learning system in humans. *Science, 273*, 1399-1402.

Knudsen, Eric I. (1987). Early experience shapes auditory localization behavior and the spatial tuning of auditory units in the barn owl. In Joseph P. Rauschecker & Peter Marler, Eds. (1987). *Imprinting and cortical plasticity* (pp. 3-7). New York: John Wiley & Sons.

Koestler, A. (1967). *The ghost in the machine*. Chicago: Henry Regnery Company.

Kuiper, K. (1991).The evolution of an oral tradition: Race-calling in Canterbury, New Zealand. *Oral Tradition 6*, 19-34.

Kuiper, K. (1992). The English oral tradition in auction speech. *American Speech, 67*, 279-289.

Kuiper, K. (1996). *Smooth Talkers: The linguistic performance of auctioneers and sportscasters.* Mahwah, NJ: Erlbaum Association Publishers.

Kuiper, K. (2000). On the linguistic properties of formulaic speech. *Oral Tradition, 15*(2), 279-305.

Kuiper, K. (2004). Formulaic performance in conventionalised varieties of speech. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 37-54). Amsterdam: John Benjamins.

Kuiper, K., & Everaert, M. (2000). Constraints on the phrase structural properties of English phrasal lexical items. In B. Roswadowska (Ed.), *PASE Papers in Language Studies* (pp. 151-170). Wroclaw: Aksel.

Kuiper, K., & Flindall, M. (2000). Social rituals, formulaic speech and small talk at the supermarket checkout. In J. Coupland (Ed.), *Small talk* (pp. 183-207). London: Longman.

Kuiper, K., & Haggo, D. (1985). The nature of ice hockey commentaries. In R. Berry & J. Acheson (Eds.), *Regionalism and national identity: Multidisciplinary essays on Canada* (189-197). Australia and New Zealand, Christchurch: Association for Canadian Studies in Australia and New Zealand.

Kuiper, K., van Egmond, M., Kempen, G. & Sprenger, S. (2007). Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon 2* (3), 313-357.

Larsen, B., Skinhoj, E., & Lassen, H. A. (1978). Variations in regional cortical blood flow in the right and left hemispheres during automatic speech. *Brain, 10,* 193-200.

Lattal, K. A. (1995). Contingency and behavior analysis. *Behavior Analyst 24*, 147-161.

Lieberman, P. (2000). *Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought*. Cambridge, MA: Harvard University Press.

Locke, J. L. (1993). The child's path to spoken language. Cambridge, MA: Harvard University Press.

Locke, J. L. (1995). Development of the capacity for spoken language. In P. Fletcher and B. MacWhinney (Eds.) *The handbook of child language* (pp. 278– 302). Oxford: Blackwell.

Lord, A. (1960). *The singer of tales*. Cambridge, MA: Harvard University Press.

Lord, Catherine & Paul, Rhea (1997). Language and communication in autism. In Donald J. Cohen and Fred R. Volkmar (Eds*.). Handbook of autism and pervasive developmental disorders* (pp. 195-225). New York: John Wiley & Sons.

Lounsbury, F.G., 1963. Linguistics and psychology. In: Koch, S. (Ed.), *Psychology: Study of a science* (pp. 553–582). McGraw-Hill, New York.

Lum, C. C., & Ellis, A. W. (1994). Is nonpropositional speech preserved in aphasia? *Brain and Language, 46*, 368-391.

Marler, P. (1987). Sensitive periods and the role of specific and general sensory

stimulation in birdsong learning. J.P. Rauschecker & P. Marler (Eds.). *In: Imprinting and cortical plasticity* (pp. 99-135). John Wiley & Sons, New York.

Marler, P. (1998). Animal communication and human language. In G. Jablonski and L. C. Aiello (Eds.). The origin and diversification of language. *Wattis Symposium Series in Anthropology* (pp. 1-19). Memoirs of the California Academy of Sciences, No. 24, San Francisco: California Academy of Sciences.

Marsden, C. D. (1982). The mysterious motor function of the basal ganglia: The Robert Wartenberg lecture. *Neurology, 32*, 514-539.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh & N. M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65-67). New York: The Guilford Press.

Moon, R. E. (1998). Fixed expressions and text: a study of the distribution and textual behaviour of fixed expressions in English. *Oxford studies in lexicology and lexicography*. Oxford: Clarendon Press.

Nakagawa, Y., Tanabe, H., Ikeda, M., Kazui, H., Ito, K., Inoue, N., Hatakenaka, Y., Sawada, T., Ikeda, H. & Shiraishi, J. (1993). Completion phenomenon in transcortical sensory aphasia. *Behavioural Neurology, 6*(3), 135-142.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman & Co.

Norrick, N.R., 1985. *How proverbs mean: Semantic studies in English proverbs*. Mouton, Berlin.

Osgood, C., & Hoosain, R. (1974). Salience of the word as a unit in the perception of language. *Perception and Psychophysics, 15,* 168-192.

Page, D.L., 1959. *History and the Homeric Iliad.* University of California Press, Berkeley.

Patel, Aniruddh D. (2008). *Music, language, and the brain*. Oxford: Oxford University Press.

Paul, R. (2004). Autism. In R. D. Kent (Ed.), *The MIT handbook of communication disorders.* Cambridge, MA: MIT Press.

Pawley, A. (2001). Phraseology, linguistics, and the dictionary. *International Journal of Lexicography*, *14(2)*, 122-134.

Pawley, A. (2007). Developments in the study of formulaic language since 1970: a personal view. In P. Skandera (Ed.), *Phraseology and culture in English*, (pp. 3-45). Berlin: Mouton de Gruyter.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J. C. Richards, R. Schmidt (Eds.), *Language and communication* (pp. 191-225). London: Longman.

Peters, A. (1977). Language-learning strategies: does the whole equal the sum of the parts? *Language, 53*, 560-573.

Peters, A.M. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.

Petrovici, J-N. (1980). Speech disturbances following stereotaxic surgery in ventrolateral thalamus. *Neurosurgical Review, 3*(3), 189-195.

Pickens, J. D., & Pollio, H. R. (1979). Patterns of figurative language competence in adult speakers. *Psychological Research, 40*, 299-313.

Pinker, S. (1995). *The language instinct*. New York: Harper Collins.

Ploog, D. (1975). Vocal behavior and its 'localization' as prerequisite for speech. In K. J. Zülch, O. Creutzfeldt, & G. C. Galbraith (Eds.), *Cerebral localization*. Berlin: Springer-Verlag.

Prizant, B. M. (1983). Language acquisition and communicative behavior in autism: Toward an understanding of the 'whole' of it. *Journal of Speech and Hearing Disorders*, *48*, 296–307.

Prizant, B. M. and Duchan, J., (1981). The functions of immediate echolalia in autistic children. *Journal of Speech and Hearing Disorders, 46*, 241–249.

Prizant, Barry M., Schuler, Adriana L., Wetherby, Amy B., & Rydell, Patrick. (1997). Enhancing language and communication development: Language approaches. In Donald J. Cohen and Fred R. Volkmar (Eds.). *Handbook of autism and pervasive developmental disorders* (pp. 572-605). New York: John Wiley & Sons.

Robinson, B.W. (1987). Limbic influences on human speech. *Annals of the New York Academy of Sciences, 280*, 761-771.

Ryding, E., Bradvik, B., & Ingvar, D. (1987). Changes of regional cerebral blood flow measured simultaneously in the right and left hemisphere during automatic speech and humming. *Brain, 110*, 1345-1358.

Schaltenbrand, G. (1965). The effects of stereotactic electrical stimulation in the depth of the brain. *Brain, 88,* 835-840.

Schegloff, E. (1988). Discourse as an interactional achievement II: an exercise in conversation analysis. In D. Tannen (Ed.), *Linguistics in context: Connecting observation and understanding* (pp. 135-158). Norwood, New York: Ablex.

Schweizer, B.-M. (1978). *Sprachspiel mit Idiomen: eine Untersuchung am Prosawerk von Günter Grass*. Zürich: Juris Druk Verlag.

Sidtis, D., Rogers, T., Katsnelson, D & Sidtis, J.J. (2008). Effects in Parkinson's subjects ON vs OFF deep brain stimulation on overlearned and serial speech. Motor Speech Conference, March 6-9, Monterey, CA.

Simon, H.A. (1974). How big is a chunk? *Science*, *183*, 482-448.

Simon, H.A., Zhang, W., Zang, W., Peng, R., 1989. STM capacity for Chinese words and idioms with visual and auditory presentations. In H. A. Simon (Ed.), *Models of thought, II* (pp. 68–75). Yale University Press, New Haven and London.

Smolka, Eva, Rabanus, Stefan, Rösler, Frank. (2007). Processing verbs in German idioms: Evidence against the configuration hypothesis. *Metaphor and Symbol, 22* (3), 213-231.

Sorace, A. & Keller, F. (2005). Gradience in linguistic data. *Lingua 115*, 1497-1524.

Speedie, L.J., Wertman, E., Tair, J. & Heilman, K.M. (1993). Disruption of automatic speech following a right basal ganglia lesion. *Neurology, 43*, 1768-1774.

Sprenger, S.A. (2003). *Fixed expressions and the production of idioms*. Doctoral dissertation, University of Nijmegen.

Squires, L.R. (1986). Memory and the brain. In Sarah L. Friedman, Kenneth A. Klivington, & Rita W. Peterson (Eds.), *The brain, cognition, and education*. Orlando: Academic Press.

Svartik, J. & Quirk, R. (1980). A corpus of English conversation. *Lund studies in English*. Lund: CWK Gleerup.

Swinney, D A. & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18,* 523-534.

Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse.* Cambridge: Cambridge University Press.

Tannen, D., & Öztek, P.C. (1981). Health to our mouths: Formulaic Expressions in Turkish and Greek. In F. Coulmas (Ed.), *Conversational routine. Explorations in standardized communication situations and prepatterned speech* (pp. 37-57). The Hague: Mouton.

Tilley, M. P. (1950). *A dictionary of proverbs in England in the 16th and 17th centuries*. Ann Arbor, MI: University of Michigan Press.

Titone, D.A. & Conine, C.M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics 31*, 1655-1674.

Tomasello, M. (2003). *Constructing a language. A usage-based theory of language acquisition*. Cambridge, MA, & London: Harvard University Press.

Van Lancker Sidtis, D. (2001). Preserved formulaic expressions in a case of transcortical sensory aphasia compared to incidence in normal everyday speech. *Brain and Language, 79* (1), 38-41.

Van Lancker Sidtis, D. (2004). When novel sentences spoken or heard for the first time in the history of the universe are not enough: Toward a dual-process model of language. *International Journal of Language and Communication Disorders, 39* (1), 1-44.

Van Lancker Sidtis, D. (2006). Has neuroimaging solved the problems of neurolinguistics? *Brain and Language, 98*, 276-290.

Van Lancker Sidtis, D. (2008). Formulaic and novel language in a dual-process model of language competence: Evidence from surveys, speech samples, and schemata. In R. Corrigan, E. Moravcsik, H. Ouali, and K. Wheatley, (Eds.). *Proceedings of the 2007 University of Wisconsin-Milwaukee symposium on formulaic language*. Amsterdam: John Benjamins (to appear).

Van Lancker Sidtis, D. & Postman, W.A. (2006). Formulaic expressions in spontaneous speech of left- and right-hemisphere damaged subjects. *Aphasiology, 20* (5), 411-426.

Van Lancker Sidtis, D. & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: methods for classification and verification. *Language and Communication, 24*, 207-240.

Van Lancker Sidtis, D., Canterucci, G. & Katsnelson, D. (submitted). Impairment of formulaic language in cortical and subcortical brain damage.

Van Lancker, D. (1993). Nonpropositional speech in aphasia. In G. Blanken, J. Dittmann, J. Grimm, J. C. Marshall, C-W. Wallesch (Eds.), *Linguistic disorders and pathologies: An international handbook* (pp. 215-225). Berlin: Walter de Gruyter.

Van Lancker, D. & Bella, R. (1996). The relative roles of repetition and sentence completion tasks in revealing superior speech abilities in patients with nonfluent aphasia. *Journal of the International Neuropsychological Society, 2,* 6.

Van Lancker, D. & Cummings, J. (1999). Expletives: neurolinguistic and neurobehavioral perspectives on swearing. *Brain Research Reviews, 31*, 83-104.

Van Lancker, D., Bogen, J. E. & Canter, G. J. (1983). A case report of pathological rule-governed syllable intrusion. *Brain and Language, 20*, 12-20.

Van Lancker, D., Canter, J. & Terbeek, D. (1981). Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech and Hearing Research, 24*, 330-335.

Van Lancker, D., McIntosh, R. & Grafton, R. (2003). PET activation studies comparing two speech tasks widely used in surgical mapping. *Brain and Language, 85*, 245-261.

Whitaker, H. (1976). A case of the isolation of the language function. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in Neurolinguistics* (Vol. 2, pp. 1-58). London: Academic Press.

Wong Fillmore, L. (1979). Individual differences in second language acquisition. In C. J. Fillmore, D. Kempler, W. S-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 203-228). London: Academic Press.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication, 18*, 47-67.

Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. In Chris Knight, James R. Hurford and Michael Studdert-Kennedy (Eds.). *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 285-302). Cambridge: Cambridge University Press.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, A. (2008a). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

Wray, A. (2008b). Formulaic sequences in language disorder. In M. Ball, M. Perkins , N. Muller & S. Howard (Eds.). *The handbook of clinical linguistics* (pp. 184-197). Oxford: Blackwell Publishers Ltd.

Wray, A. & Grace, G. W. (2007) The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua, 117*(3), 543-578.

Wray, A., & Perkins, M. (2000). The functions of formulaic language: an integrated model. *Language and Communication, 20*, 1-28.