Sound Matters

–

The Affective Nature of Phonological Iconicity

DISSERTATION

in partial fulfillment of the requirements of the academic degree

Doktorin der Philosophie (Dr. phil.)

in Psychology

submitted to

the Department of Human Sciences

University of Kassel

by

Susann Ullrich

Berlin, August 2021

Supervisor and first reviewer: Prof. Dr. Ralf Rummer

Institut für Psychologie

Universität Kassel

Second reviewer: Prof. Dr. Judith Schweppe

Philosophische Fakultät

Universität Passau

Date of thesis defense: December 6th, 2021

"There are words which we feel instinctively to be

adequate to express the ideas they stand for,

and others the sounds of which are felt to be

more or less incongruous with their signification."

(Jespersen, 1922, p. 398)

Table of Contents

Summary

In linguistics, it is predominantly assumed that human language is arbitrary, i.e., is using symbols that do not bear any relationship to the meaning assigned to them. However, several linguists have repeatedly detected iconic relationships, i.e., a fit between certain language sounds and word meanings, within language corpora. Furthermore, numerous psycholinguistic studies have revealed empirical evidence for so-called iconicity patterns over the past 100 years. This suggests that iconicity is not just a negligible language phenomenon, but rather a functional mechanism of language that most likely played a critical role in the creation of language and during its evolution, and is still doing so. Accordingly, it is of scientific interest not only to prove the existence of such sound-meaning relationships, but also to examine the cognitive processes that lead to linguistic iconicity or are influenced by it. However, the approaches used by previous iconicity studies, investigating several different semantic dimensions, such as size, shape, brightness, weight, distance, speed, etc., do not easily allow for a holistic operationalization of iconicity in a language, here German, which would be necessary to investigate its function in general language processing. The semantic differential represents a possibility of mapping many different dimensions onto a few core dimensions. It reveals that the two dimensions valence and arousal account for most of the variance in word ratings for many semantic dimensions. This means that, in addition to the denotative meaning, connotative valence and arousal values resonate with most words. This is also plausible from a psychoneurobiological perspective, because dimensional theories of emotion assume that our reactions to environmental stimuli are mediated by valence and arousal. Accordingly, every stimulus that we perceive is quickly and automatically evaluated by our nervous system with regard to its valence and its arousal content, and the combination of these affective values results in a specific emotion that motivates the corresponding

behavioral reaction to the stimulus. Since words are associated with the mental concepts of the stimuli they designate in our environment (or within one's own body), affective evaluations also take place when words are perceived. Such affective evaluations can be collected through word ratings for valence as well as arousal which are then stored in affective word databases for scientific use. With affective stimuli from such word databases, extensive studies on affective word processing have been carried out over the past two decades. These have shown that the affectivity of words is evaluated automatically and early (about 200 ms after the word onset) and influences language processing. Since iconicity presupposes that there is also meaning coded at the level of speech sounds, the question arises whether valence and arousal also play a role at the sublexical level. In the introduction of this dissertation, I show for the affective dimensions valence and arousal that sound-meaning relationships can be substantiated across a German affective word database, from which a top-down operationalization of phonological iconicity is derived – the *sublexical affective potential*. In the first study, I use a lexical decision task, during which an electroencephalogram (EEG) was recorded, to show that this *sublexical affective potential* of words has a psychophysiological reality, i.e., it can be detected in the EEG. In fact, the EEG components of the *sublexical affective potential* and the *lexical affective content* of words are similar in the time of their occurrence (between 200 and 300 ms) and in their form (a posterior negativity), which speaks for a parallel affective processing of the word stimuli at both language levels. In the second study, the *sublexical affective potential* was contrasted with the *lexical affective content* of a word, that is, affectively iconic and non-iconic words were compared with each other. The EEG showed a significant interaction of the affective properties at the sublexical and the lexical level in the period between 280 and 430 ms after the stimulus onset: Iconic words, where affective values at the sublexical and lexical level are consistent, regardless of whether they are both negative and high-arousing or neutral and low-arousing, are processed more easily – we assume a facilitated lexical access – while inconsistent words show a higher negativity (N400) in the EEG, which indicates an increased processing effort. We were thus able to show a general relevance of iconicity for language processing, which underpins the results of other first studies in this direction and, overall, emphasizes the importance of iconicity as an active linguistic mechanism. An applied reading study with emotional poetry was carried out as a third study. Affective properties at the sublexical, lexical, and inter-lexical level of 57 poems were quantified and used as predictors in stepwise multiple regression analyses to explain the *general affective meaning* of the poems, which emerged from ratings. It was shown that affective properties of all linguistic levels influence the affective and aesthetic perception of the poems in a complementary way. All in all, the study results of this dissertation show that phoneme segments of the German language carry connotative affective meanings that can be used for a holistic operationalization of phonological iconicity on the basis of the affective dimensions of valence and arousal, which then allows the investigation of the role of iconicity in language processing.

Zusammenfassung

In der Linguistik wird vorwiegend davon ausgegangen, dass menschliche Sprache arbiträr ist, also Symbole benutzt, die in keiner Verbindung zu der jeweiligen ihnen zugewiesenen Bedeutung stehen. Jedoch haben Sprachwissenschaftler immer wieder ikonische Zusammenhänge, also eine Passung zwischen gewissen Sprachklängen und Wortbedeutungen, innerhalb von Sprachkorpora erkannt. Und auch durch psycholinguistische Studien konnten über die letzten 100 Jahre viele empirische Belege für sogenannte Ikonizitätsmuster gefunden werden. Dies legt nahe, dass es sich bei Ikonizität nicht nur um ein nebensächliches Sprachphänomen handelt, sondern vielmehr um einen funktionellen Mechanismus von Sprache, der mit hoher Wahrscheinlichkeit massiv an der Entstehung von Sprache und bei ihrer Evolution mitgewirkt hat bzw. es immer noch tut. Demzufolge ist es von Erkenntnisinteresse, nicht nur die Existenz solcher Klang-Bedeutungs-Zusammenhänge aufzuzeigen, sondern auch die kognitiven Abläufe zu untersuchen, welche sprachlicher Ikonizität unterliegen bzw. von ihr beeinflusst werden. Für eine dafür benötigte ganzheitliche Operationalisierung von Ikonizität in einer Sprache, hier der deutschen, stellt es jedoch ein Problem dar, dass in den bisherigen Studien viele verschiedene semantische Dimensionen, wie z.B. Größe, Form, Helligkeit, Gewicht, Entfernung, Geschwindigkeit, etc., betrachtet wurden. Eine Möglichkeit viele Dimensionen auf wenigen Basisdimensionen abzubilden stellt das semantische Differential dar. Bei dessen Anwendung zeigt sich, dass die beiden Dimensionen Valenz und Arousal einen Großteil semantischer Dimensionen abbilden können. Dies bedeutet, dass bei vielen Wörtern über die denotative Bedeutung hinaus auch stets konnotative Valenz- und Arousalwerte mitschwingen. Dies ist auch aus psychoneurobiologischer Perspektive plausibel, denn dimensionale Emotionstheorien gehen davon aus, dass unsere Reaktionen auf Umweltreize durch Valenz und Arousal vermittelt

werden. Demzufolge wird jeder Stimulus, den wir wahrnehmen, von unserem Nervensystem schnell und automatisch bezüglich seiner Valenz und seines Arousalgehaltes evaluiert, und aus der Kombination dieser affektiven Werte ergibt sich eine spezifische Emotion, die das entsprechende Reaktionsverhalten auf den Stimulus begründet. Da Wörter mit den von ihnen bezeichneten mentalen Konzepten der Stimuli in unserer Umgebung (oder auch innerhalb des eigenen Körpers) assoziiert sind, finden bei ihrer Wahrnehmung ebenfalls affektive Evaluationen statt. Diese können durch Wortratings für Valenz als auch Arousal explizit erfasst und in affektiven Wortdatenbanken dargestellt werden. Mit affektiven Stimuli aus solchen Wortdatenbanken wurden in den letzten zwei Jahrzehnten umfassende Studien zur affektiven Wortverarbeitung durchgeführt, welche u. a. gezeigt haben, dass die Affektivität von Wörtern automatisch und frühzeitig (etwa 200 ms nach dem Wortonset) evaluiert wird und die Sprachverarbeitung beeinflusst. Da Ikonizität voraussetzt, dass auch auf der Ebene der Sprachklänge Bedeutung kodiert ist, stellt sich nun die Frage, ob auch auf der sublexikalischen Ebene Valenz und Arousal eine Rolle spielen. In der Einleitung dieser Dissertationsschrift zeige ich, dass in einer deutschen affektiven Wortdatenbank Klang-Bedeutungs-Zusammenhänge für die affektiven Dimensionen Valenz und Arousal nachweisbar sind, von denen sich eine top-down Operationalisierung phonologischer Ikonizität ableiten lässt – das *sublexikalische affektive Potenzial.* In der ersten Studie zeige ich mittels einer lexikalischen Entscheidungsaufgabe, während der ein Elektroenzephalogramm (EEG) aufgezeichnet wurde, dass dieses *sublexikalische affektive Potenzial* von Wörtern eine psychophysiologische Realität besitzt, also im EEG nachgewiesen werden kann. Tatsächlich ähneln sich die EEG-Komponenten des *sublexikalischen affektiven Potenzials* und des *lexikalischen affektiven Wortgehalts* im Zeitpunkt ihres Auftretens (zwischen 200 und 300 ms) und in ihrer Form (einer posterioren Negativierung), was für eine parallele affektive Verarbeitung der Wortstimuli auf beiden Sprachebenen spricht. In der zweiten Studie wurden dann das *sublexikalische affektive Potenzial* und der *lexikalische affektive Gehalt* eines Wortes kontrastiert, also affektiv ikonische und nicht-ikonische Wörter miteinander verglichen. Im EEG zeigte sich dabei eine signifikante Interaktion der affektiven Eigenschaften auf der Klang- und der Bedeutungsebene im Zeitraum zwischen 280 und 430 ms nach dem Stimulusonset: Ikonische Wörter, bei denen die affektiven Werte auf der sublexikalischen und lexikalischen Ebene konsistent sind, egal ob sie beide negativ und aufregend oder neutral und unaufgeregt sind, können leichter verarbeitet werden – wir gehen von einem erleichterten lexikalischen Zugriff aus – während sich bei den inkonsistenten Wörtern eine erhöhte Negativierung (N400) zeigt, die auf einen erhöhten Verarbeitungsaufwand hindeutet. Somit konnten wir eine allgemeine Relevanz von Ikonizität für die Sprachverarbeitung zeigen, welche erste Ergebnisse anderer Studien untermauert und insgesamt die Bedeutung von Ikonizität als aktiven sprachlichen Mechanismus betont. Zur Abrundung wurde als dritte Studie eine angewandte Lesestudie mit emotionalen Gedichten durchgeführt, in denen affektive Eigenschaften auf der sublexikalischen, lexikalischen und inter-lexikalischen Ebene quantifiziert und als Prädiktoren in schrittweisen multiplen Regressionen zur Erklärung der *allgemeinen affektiven Bedeutung* der Gedichte, welche aus Ratings hervorging, verwendet wurden. Es zeigte sich, dass affektive Eigenschaften aller sprachlichen Ebenen in sich ergänzender Weise die affektive und ästhetische Perzeption der Gedichte beeinflussen. Insgesamt wird mit den Arbeiten im Rahmen dieser Dissertation gezeigt, dass Sprachklänge der deutschen Sprache konnotative affektive Bedeutungen tragen, die zu einer holistischen Operationalisierung phonologischer Ikonizität auf der Basis der affektiven Dimensionen Valenz und Arousal genutzt werden können und somit die Erforschung der Rolle von Ikonizität bei der Sprachverarbeitung ermöglichen.

1. Introduction

1.1. General introduction

Research on the human mind throughout the last decades has come to the conclusion that human cognition does not result from independent rational computations in the human brain, but is highly intertwined with emotions and the affective system in general (Eich, Kihlstrom, Bower, Forgas, & Niedenthal, 2000). Emotions and affect are the driving components of human behavior and influence each part of cognition – including language (Shanahan, 2008). Research on the affective meaning of words, sentences, and whole texts as well as affective prosodic expressions has flourished in the last years. However, one aspect of language has been mostly disregarded so far: the sublexical level, i.e., the level of syllables, subsyllabic segments and single phonemes – the sounds of language[1]. In opposition to the still prevailing assumption that the assembly of language sounds which form words is arbitrary (Saussure, 1916), this thesis follows the idea that language sounds already possess some type of coarse, in particular affective, meaning that influences language processing.

But how can language in general transfer meaning at all? Technically, different patterns of acoustic, visual, and even tactile features (mostly articulatory sounds, letters, or gestures, but also pictograms, morse tones, whistling and clicking sounds, or even tapping perceptions on the body) are conventionally combined and assigned with a specific meaning. The study of meaning in language is called semantics, the broader field of how signs in general can carry meaning is called semiotics. In his influential semiotic work, the philosopher Charles Sanders Peirce (1839-1913) defined a sign as a triadic relationship between the sign itself (for example

[1] Throughout this thesis, speaking of language sound(s) refers to phonemes or phoneme clusters (such as complex syllable onsets, nuclei, or codas – see chapter 1.6). Other language sound aspects such as prosody or variable pronunciations in different dialects are not dealt with in this thesis.

the word “tree”), the signified object (for example a particular tree), and its interpretant (the result of interpreting the object, i.e., a mental concept of the tree) (Peirce, 1931). The latter one is very central to his theory, as the meaning of a sign is mainly determined by the effect that it generates in a person (Atkin, 2013). For example, if a family have just bought a fir tree for Christmas, all family members stand in front of the same object and call it “Christmas tree”. However, everyone may interpret his or her percept of the fir tree somewhat differently. The mother might appreciate its seemingly perfect symmetrical shape and already imagines how beautifully she will decorate the tree. The father, who carried the tree and will place it in its stand, might in particular feel the needles pricking his skin. For the little boy who hardly remembers last year’s Christmas, the fir tree is impressively big and has an unfamiliar smell, whereas the teenage sister is already bored by the yearly tradition of decorating a tree and spending several days together with the family, and mainly cares about the gifts that will lie under the tree… Now, if there was no name for the object “Christmas tree” already, but every family member would be asked to invent one, presumably each one would come up with a different one. Thus, not the signified tree as a whole with all its properties would influence the name selection, but the individual mental concept, the interpretant in Peirce’s words. Wilhelm von Humboldt had a very similar notion of how different percepts of one object could have led to different language terms and thus to different languages in the long run.[2]

To get back to the idea that even language sounds can already carry some basic meaning, let’s continue the story of our name inventing family: If this was a contemporary family, the name suggestions would most likely be derived from already existing words that match their

[2] From Humboldt (1836, chapter 11): „und wenn z.B. im Sanskrit der Elephant bald der zweimal Trinkende, bald der Zweizahnige, bald der mit einer Hand Versehene heißt, so sind dadurch, wenn auch immer derselbe Gegenstand gemeint ist, ebenso viele verschiedene Begriffe bezeichnet. Denn die Sprache stellt niemals die Gegenstände, sondern immer die durch den Geist in der Spracherzeugung selbstthätig von ihnen gebildeten Begriffe dar […].“

individual associations, for example “shiny tree”, “pricky bitch”, “huge green thingy”, or “tree of boredom”. However, if we imagine a prehistoric family that is just developing a language and has so far only lived in scrubland and now sees a fir tree for the first time, their newly invented names might potentially include sounds that somehow mirror the perceived tree characteristics – for example an /i:/ for its beauty, some /t/ or an even more arousing /tr/ because of the pricky feeling, an /o:/ for its largeness, and perhaps a /w/ for the boredom. If all family members had to agree on one word as a memorizable and to be shared name for the fir tree, it might be something like “treow”[3]. After some sound changes over thousands of years, what might be left of this is the word form “tree”.

In what follows, I will first dive into the conceptual and empirical history of phonological iconicity in language, including definitions of the different terms used in this conceptual field. I will then introduce an account of affective dimensions of meaning and how it can be suitable at the sublexical language level to comprise the multitude of dimensions that phonological iconicity acts on. From there, I will develop my working hypothesis and the research questions that I want to answer with this thesis. By taking theoretical excursuses to aspects of the neurological processing of language, I will derive specific hypotheses and the methodological approaches that I use to test them.

[3] According to https://www.etymonline.com/word/tree this was the Old English predecessor word of “tree”. I am not an etymologist and fully aware that first languages appeared much earlier. I am just using this as a not too far off reality example to corroborate my story.

1.2. Historical views on sound-meaning correspondences

Human language has evolved over presumably the last 100,000 years and nobody knows exactly how it originated (Atkinson, 2011; Jackendoff, 2006). This is why modern humans have again and again wondered and hypothesized about how the first naming of objects and later more abstract concepts might have happened. A first documentation of this topic represents Plato's dialogue "Cratylus" where Socrates and his scholars Cratylus and Hermogenes discuss their three theories of language (Jowett, 1892). Hermogenes holds the view that names are conventional – the names for concepts could be changed at will. He draws on the example of slaves' names, that when one buys a slave, he gives him a new name, while the character of the slave does not change. The same holds for different names for one and the same concept in different languages. Cratylus, however, argues that true names relate to the concept they are naming – for example, that they are formed by the imitation of ideas in sounds. Socrates offers a compromise in that he opts for a balance of the two forces. He supposes words to be originally formed by the imitation of ideas in sounds. But the vocal imitation of natural sounds may have been imperfectly executed, which is how elements of chance or convention entered (Jowett, 1892). In particular, to vocally imitate concepts that do not emit sounds must have led to a variety of descriptions depending on context and idiosyncratic associations – and thus to different languages. Also over time, names got changed by other speakers and their desire of euphony. Socrates further mentions the interactive influence of languages onto each other and also a certain element of chance. Thus, the original meanings of most words got obscured and these words now appear artificial. Yet, Socrates suggests that the truth of names can be found in the analysis of their elements (Jowett, 1892) and lists several examples: Alpha (/aː/) expresses size, Eta (/ɛː/) length, and Omikron (/o/) roundness, whereas Rho (/r/) stands for rushing or roaring, Lambda (/l/) for liquidity, and Phi (/f/) for wind.

In a similar manner, since then, several philologists and linguists argued for an inner connection between the concept of a word and the sounds the word is composed of. In the 17th century, in Baroque poetry, it was important that the letters' sounds of the words used in a poem suggest its meaning.[4] This even led to comprehensive manuals of how to use letters (nowadays referred to as phonemes) in poetry with reference to their nature. Example statements thereof say that /a:/ and /o:/ are strong-sounding letters, whereas /i:/ and /e:/ are soft and mild, with /u:/ coming halfway in between (Buchner, 1665, translated by Kjetsaa, 1974). Comenius (1592-1670), an educator and theologian who also studied several languages and attempted to create a universal language, as well held that /a:/ represented something vast and big, or that /l/ suggested something soft, and /r/ something hard (Kjetsaa, 1974, translating Comenii, 1966).

In the 18th century, the Russian polymath Mikhail V. Lomonosov (1711-1765) proposed that front vowels such as /i:/ were used in poetry for tender aspects, whereas back vowels such as /o:/ were rather used for things related to anger, fear, or sadness. A frequent repetition of the vowel /a:/ he assigned to magnificence or great space (Kjetsaa, 1974, translating Lomonosov, 1748).

Also famous scholars of the 19th century, such as Wilhelm von Humboldt or Jacob Grimm were convinced that the sound(s) of language accounted for the meaning of words, at least in the beginning of language development. Humboldt even recognized three types of iconicity in human language: (1) *direct imitations* of environmental sounds as close as the human

[4] From Opitz (1624): „Weil ein buchstabe einen andern klang von sich giebet als der andere / soll man sehen / das man diese zum offteren gebrauche / die sich zue der sache welche wir für vns haben am besten schicken. Als wie Virgilius von dem berge Etna redet / brauchet er alles harte vnd gleichsam knallende buchstaben." and „So / weil das L vnd R fließende buchstaben sein kan ich mir sie in beschreibung der bäche vnd wäßer wol nütze machen."

articulatory tract can manage (onomatopoeia[5]), (2) *symbolic denotations* where the object and the chosen language sound(s) share a common feature, for example, stability (such as the “st” in “stand”, “stiff”, “stable”, or “steady”), or fluctuation (such as the “w” in “wave”, “wind”, “waft”, or even “blow”), and (3) *sound analogies*, i.e., choosing similar language sounds for related concepts – however, without relating to the characteristics of the sound(s) themselves (Humboldt, 1836, pp. 78-81).

Grimm, one of the first philologists, claimed that nothing in language and nature was unnecessary, and hence, no letter has developed without an original meaning. He related the meaning of single language sounds to their place and way of articulation: "Every sound has its natural form, based on the organ producing it and readily available. […] Of the consonants, /l/ will designate what is gentle, and /r/ what is rough" (Grimm, 1864, p. 284).

Throughout the 19th century, linguistic studies took on form. The early linguist Georg von der Gabelentz referred to meaningful language sounds as etymological roots. He assumed that “in the original state of language, all roots, i.e., their sound units, had instinctively sound symbolic value for the speakers. Over time, this language instinct did not get lost, but just took on different directions” (von der Gabelentz, 2016, p. 237).

However, concepts such as "instinct" or "feel for language" did not play a role in the following era of modern linguistics, employing a purely rational scientific approach. Most of the formerly mentioned phenomena describing relations between language sounds and innate meaning were dismissed as ideocratic.

---

[5] Examples for onomatopoetic expressions would be: animals’ calls such as the *meowing* of cats or the song of the *coocoo* which gave him his name, or phrased nature sounds such as *dripping* water, the *rustling* of dry leaves, a *cracking* twig or the *hush* of the wind. Comic novels use onomatopoetic expressions to describe environmental sounds in their story settings, such as *bam, zoooom, crash, pow, puff, wham, slurp, biff, pop,* or *splash*.

In his seminal work “Cours de linguistique générale” (1916), the linguist Ferdinand de Saussure stated that the link between the signified (concept) and the signifier (word form) was arbitrary. “The word arbitrary […] implies that the signifier is unmotivated […] in relation to the signified, with which it has no natural connection in reality. […] In principle, any signifier could represent any signified. […] This is demonstrated by differences between languages, and even by the existence of different languages” (Saussure, 1916). Only onomatopoeia – the direct imitation of nature sounds – was accepted as exception from this rule, but dismissed as a marginal phenomenon. This view that arbitrariness was an important principle of language was not new, already Aristotle had denied a “natural connection between the sound of any language and the things signified” (Richards, 1932, p. 32).

However, despite being in the minority, opposing views have been voiced throughout. Immediately after de Saussure’s “Cours de linguistique générale” got published, Otto Jespersen criticized the “strong overestimation of the role of arbitrariness in language and the underestimation of the role of the sound-words” (Chang-Liang, 2018, translating Otto Jespersen, 1917). In his book “Language, its nature, development, and origin” (1922), Jespersen questions the logic behind the denial of sound symbolism in language, stating that it does not make sense to see “in our words only a collection of wholly accidental and irrational associations of sound and meaning”. He admits that many, actually most words may not be symbolic in their sense, but “that there are words which we feel instinctively to be adequate to express the ideas they stand for” (Jespersen, 1922, p. 398). Then, he lists substantial descriptions of linguistic examples that support the existence of sound symbolism in language: He classifies the examples into direct imitations and echoic words (onomatopoeia), impressions by movement (e.g., short vowels followed by a stop consonant indicating rapid striking movements – such as in *tap, knock, snap, or catch*), impressions by appearance (e.g., light being associated with high tones and darkness with low tones), states of mind (such as dislike being expressed by back vowels as in *dull, humdrum, sloppy,* or *numbskull*), size and distance (with a strong focus on the vowel [i] expressing smallness, weakness, insignificance, femininity, short duration, or close distance – also see his separate publication on the "Symbolic Value of the Vowel i", Jespersen, 1933), and finally the connection of word length with strength (with short word forms being used for imperatives, but also elongated word forms strengthening their meaning, such as *splendid* being exaggerated into *splendacious*, or mere prosodic amplifications by stretching the vowel such as in *huuuuge*).

1.3. Empirical research on sound-meaning correspondences

With the rise of empirical psychology since the 1920s, a growing body of evidence has been accumulated for the existence of an inherent relationship between the sound quality of phonemes and their semantic usage. It would be too much and not directly relevant to the actual question to be examined in this dissertation to present an exhaustive list of the studies conducted in the field of sound-meaning correlations (therefor see, e.g., Elsen, 2014; Hinton, Nichols, & Ohala, 1994b; Lockwood & Dingemanse, 2015; Magnus, 2013; Nuckolls, 1999; Svantesson, 2017), so I will cite the most important ones.

The most frequently used empirical approach uses invented pseudowords (to exclude the lexical meaning of existing words as a confounding influence) in which the sound characteristics are manipulated and subjects are asked to associate aspects of meaning to them. The relations of sound to size or shape were among the first and are still the most studied relations so far. In the 1920s, the linguist and anthropologist Edward Sapir constructed monosyllabic pseudowords that only differed in their vowel quality, for example "mil" and "mal". Then he asked subjects to rate whether these pseudowords denoted small or big items, for example, tables. More than 80 % of the subjects related the pseudowords containing high-frequency front vowels such as /i/ to small items and the pseudowords with low-frequency back vowels such as /a/ or /o/ to big items (Sapir, 1929). His scholar and colleague Stanley Newman confirmed these findings and elaborated on them, resulting in the following vowel continuum ranging from small size to big size: /i/ – /e/ – /a/ – /u/ – /o/ (Newman, 1933). Sound-size associations were resumed and refined in many more studies over the last decades (see for example Lowrey & Shrum, 2007, or Thompson & Estes, 2011). Ohala (1984, 1994) delivered a plausible theory – the *frequency code* – to explain such sound-size relations: Because in the animal kingdom small animals mostly produce higher sounds than large animals (compare, for example, the sounds of mice vs. bears), the linking of high-frequency sounds with small size and of low-frequency sounds with big size is evolutionary hardwired in our brains. Such associations thus might have led to the above-mentioned size-iconic patterns in language because /i:/ has a higher f_0 frequency than /a:/, /o:/, or /u:/ (Ohala, 1984). This might even transfer to associations with valence, as big animals are often more dangerous than small animals. Also within species, animals lower their voice to sound more threatening, or produce high tones to signal submission or positive affect. Among humans, children and women have higher voices than men due to their larynxes being smaller and higher in the throat than those of men, and women and children are generally perceived as more friendly and harmless than men (Ohala, 1984, 1994). Thus, the high-frequency vowel /i:/ in its iconic use also reflects positive meaning (also see Jespersen, 1933, and Rummer, 2014, 2019).

The psychologist Köhler drew two geometric figures, a rounded one with several curves and a spiky one with sharp edges (see Figure 1.1). Student subjects had to assign the names "maluma" and "takete" to these figures. More than 90 % of them assigned the name "takete" to the spiky figure and the "maluma" to the rounded figure (Köhler, 1947). Thus, voiceless plosives such as /t/ and /k/ seem to be strongly associated with edged and spiky shapes whereas continuant consonants such as /m/ and /l/ are associated with rounded objects and shapes.

Since the turn of the millennium a research revival regarding this phenomenon took place: Ramachandran & Hubbard (2001) replicated Köhler's shape naming study using "bouba" and "kiki" as pseudowords to be assigned to the spiky or round shape (95 % of their subjects assigned "kiki" to the spiky shape) which coined the term "bouba-kiki effect" for this type of iconic shape-sound relation. Their *sensorimotor synesthesia theory* states that sensory cortices in the brain are connected to motor areas and can thus cross-modally co-activate each other (Ramachandran & Hubbard, 2001, figure 8). For the "bouba-kiki" example this would mean that the neurons in the motor cortex which control the eye movement (or even neurons in the premotor cortex of hand steering that get activated just by imagining to draw the figures) register either an abrupt or a smooth movement and thereby co-activate neurons in the Broca area responsible for either smooth or abrupt articulation movements. Even without actually articulating anything, the auditory cortex registers the potentially resulting sounds via

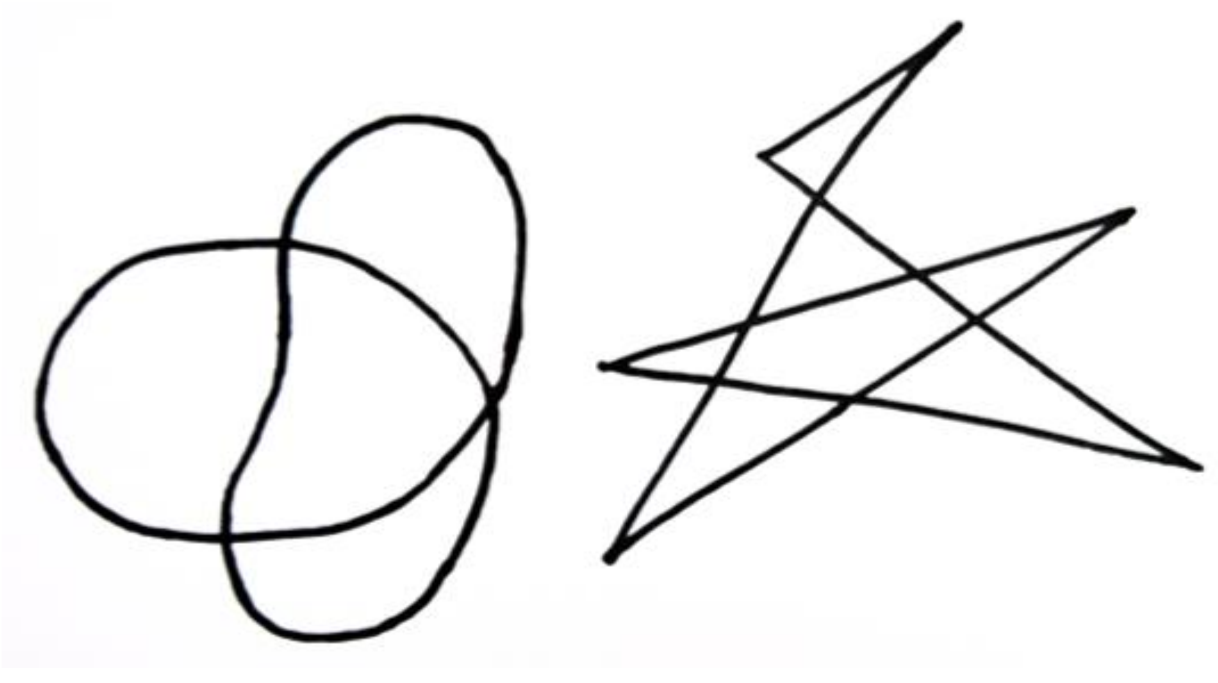

Figure 1.1: The shapes used by Köhler (1947), "maluma" (left) and "takete" (right) – own reproduction.

cross-wirings between the motor and auditory cortex that have been built up throughout the lifetime experience of talking and simultaneously hearing the resulting sounds. Comparing these sets of potential inner sounds with the sounds of the given pseudowords which have to become assigned to either of the shapes, people then choose the pseudoword with the highest match in sounds. As cross-modal percepts can be moderated by higher-level cognitive processes such as directed attention (Rich & Mattingley, 2010), we are generally not aware of such subtle associations in our everyday language use, but can make use of it when it comes to experimental settings. Regarding the "bouba-kiki" effect, some studies examined the role of the consonants in the used pseudowords (Nielsen & Rendall, 2011; Westbury, 2005), others focused on the vowel content (Spector & Maurer, 2013; Tarte & Barritt, 1971). Altogether, both vowels and consonants were found to matter in shape-sound correspondences (Ahlner & Zlatev, 2010; Nielsen & Rendall, 2013), although the effect of consonants might be slightly stronger (Fort, Martin, & Peperkamp, 2015). Behavioral studies with children aged 2.5 to 7 (Maurer et al., 2006; Tzeng et al., 2017) and eye-tracking as well as EEG studies with infants (Asano et al., 2015; Ozturk et al., 2013) show that the "bouba-kiki" effect is present already at a prelinguistic age, and suggest that sound symbolism scaffolds language learning from birth on.

The aforementioned experimental paradigms got further augmented over time: Apart from the usual size contrast, Taylor & Taylor (1962) included more dimensions of meaning such as movement, warmth, and pleasantness as anchors to which a pseudoword has to be assigned, and tested native subjects of four unrelated languages (English, Japanese, Korean, and Tamil). They could show that the phenomenon of phonetic symbolism, as they called it, existed in all tested languages and dimensions, but the patterns were different across the languages (Taylor & Taylor, 1962).

Alternatively, instead of artificial pseudowords, existing words of foreign unrelated languages (often antonyms) were presented in experiments, and subjects had to guess what those words unbeknownst to them could mean. In choosing from a selection of potential translations or allocating English antonym pairs to the foreign words, subjects guessed correctly higher than chance level. Over the years, this phenomenon could be shown for many languages unrelated to English: Chinese, Czech, and Hindi (Brown, Black, & Horowitz, 1955); Croatian and Japanese (Maltzman, Morrisett Jr., & Brooks, 1956); or Finnish, Kiwai, Old Hebrew, South Malaita, and Tongan (Gebels, 1969). Another nice example is the study of Brent Berlin (1994) who used fish and bird names of the Huambisa language (spoken by people living in the Peruvian rainforest) and presented them in pairs to native English subjects who had to check the foreign word that they believed to be a bird's name. With an accuracy rate of 58 %, the subjects guessed correctly significantly higher than chance level. Berlin also analyzed the phoneme content of the Huambisa words and found that the high front vowel /i:/ is more often present in the bird names, whereas the low central vowel /a:/ appears more dominantly in the fish names. He assumed this pattern to be related to the concept of quickness as well as sharpness vs. smoothness of the movements of birds or fish, respectively. But also aspects of size-sound correlations and an onomatopoetic influence cannot be excluded – different mechanisms seem to collude here.

To my knowledge, Brent Berlin also was the first one who tried a free word invention paradigm: He asked subjects to invent Droidese names in a CVCVCV pattern for the two shapes that Köhler had used for his study (see above, there presented as *takete* and *maluma*, but these names were not mentioned in this study). The names for the spiky *takete* shape contained more front vowels such as /i:/ and /e:/, whereas the names for the rounded *maluma* shape contained more back vowels such as /o:/ and /u:/. Regarding consonants, the *maluma*-like names contained more voiced stops such as /b/, /d/, or /g/ than the *takete*-like names (Berlin, 2006). Ralf Rummer and his colleagues used a similar approach in which subjects had to invent new names for faces of people expressing an emotion, or to name everyday objects with invented Swahili-like terms. With focus on the vowels /i:/ and /o:/, they could show that the vowel /i:/ appears more often in invented names for people showing a positive emotion or for positive objects, whereas the vowel /o:/ was found more often in names for negative objects or people displaying negative emotions (Rummer & Schweppe, 2019). This pattern, Rummer & Schweppe had predicted by their *Articulatory Feedback Hypothesis* which assumes a neural feedback loop between the articulatory muscle movements during emotion display and vowel pronunciation: Hence, /i:/ is related to positive emotion because the *zygomaticus major muscle* is involved in smiling as well as in the articulation of /i:/, whereas /o:/ is related to negative emotion because the antagonistic *orbicularis oris muscle* prevents smiling but is involved in the articulation of /o:/ (Rummer et al., 2014).

A last research paradigm to be mentioned here are word learning experiments: They show that iconic words are easier to learn than non-iconic words (Imai, Kita, Nagumo, & Okada, 2008; Kovic, Plunkett, & Westermann, 2010). For example, native English or Dutch speakers learnt Japanese words that contained some iconic mappings more easily when they were taught the right translation compared to when there were taught the opposite meaning (Lockwood, Dingemanse, & Hagoort, 2016; Nygaard, Cook, & Namy, 2009). The last two studies used Japanese ideophones as stimuli to be learnt. Ideophones are words of a special type with a highly depictive potential, mostly emphasizing actions. They exist in many languages, most comprehensively in West African and Austroasiatic languages as well as Japanese (Ahlner & Zlatev, 2010). An example would be the Japanese expression *koro-koro* denoting “a light object rolling” with the rolling character being depicted by the iconic use of /r/ and /o/ in combination with the duplication as amplifier. In contrast, *goro-goro* stands for “a heavy object rolling” – just the exchange of the voiceless /k/ against the voiced /g/ makes the movement appear iconically slower (Kita, 2008). While in English or German such ideophones are very rare[6], there still can be found a kind of action iconicity: verbs with stop consonants tend to refer to brief and abrupt actions, while words without stop sounds rather denote enduring actions (Żuchowski, 1998; Klenovšak, 2014). Also the speed of motion has been found to correspond to particular language sounds, for example back vowels representing slower speed (Cuskley, 2013).

Many more semantic properties apart from shape, size, valence, or action were found to be iconically associated with speech sounds: lightness/brightness (Hirata, Ukita, & Kita, 2011; also see Marks, 1974, 1987), weight (Kawahara, Noto, & Kumagai, 2018), distance (Ultan, 1978), taste (Gallace, Boschin, & Spence, 2011; Simner, Cuskley, & Kirby, 2010), surface feel such as roughness (Etzi, Spence, Zampini, & Gallace, 2016; Winter, Sóskuthy, & Perlman, 2017), gender (Cassidy, Kelly, & Sharoni, 1999; Cutler, McQueen, & Robinson, 1990; Sidhu & Pexman, 2015), and even abstract concepts such as personality (Sidhu, Deschamps, Bourdage, & Pexman, 2019), social dominance (Auracher, 2017), or abstractness/preciseness itself (Maglio, Rabaglia, Feder, Krehm, & Trope, 2014). Also grammatical categories such as word types (verb vs. noun) or plural forms have been shown to come along with phonological differences (Kelly, 1992), that guide children in their language acquisition (Monaghan, Chater, & Christiansen, 2005, 2007). It seems like all kinds of properties of the surrounding world that humans can perceive, are somehow phonologically presented in human language.

[6] One example would be the word *zigzag* (*Zickzack* in German) denoting a line abruptly changing its direction. In German, there are a few more examples, mostly appearing in spoken language to emphasize liveliness in story telling: *ratzfatz, holterdiepolter, pillepalle, schwipp schwapp, killekille, papperlapapp,* or *plemplem.*

1.4. Sound-meaning systematicities in language corpora

Apart from the associative relationships between many kinds of perceptible properties and language sounds that have been revealed empirically as shown, even more abstract systematicities between sound and meaning can be found in a language's lexicon: so-called phonesthemes (Firth, 1930). They are defined as frequently recurring sound-meaning pairings to be found in many languages (Bergen, 2004). Common examples are the onsets "gl-", referring to word meanings related to light and vision (e.g., *glisten*, *glimmer*, *glitter*, *gleam*, *glow*, or also in a more abstract sense *glass*, *glory* or *glamorous*), and "sn-", referring to meanings related to nose or mouth (e.g., *snore*, *snarl*, *snort*, *snout*, *sniff*, *sneeze*, or more abstractly *snuggle* or *snack*).

Here again, many different kinds of properties get represented phonesthetically in the lexicon: fastness is represented by the onset "fl-" (*fly*, *flow*, *flee*, *flutter*, or German *flitzen*, *flott*, or *fleißig*), linearity by the onset "str-" (*street*, *straw*, *strand*, *stretch*, *stripe*, or German *Strahl*, *Strebe*, or *striegeln*), small round things by the onset "kn-" (*knob*, *knee*, *knot*, *knead*, or German *Knospe*, *Knolle*, *knautschen*, or *knüllen*), negativity or weirdness by the onset "cr-" (*cranky*, *crazy*, *cringe*, *creepy*, or German *krank*, *Krieg*, *Krätze*, or *Kritik*), and many more.

But it does not always have to be the onset of a word, also rimes can bear meaning. For example, "-ash" stands for destructiveness (*smash*, *bash*, *crash*, *mash*, or *lash*), "-ap" for flatness (*map*, *cap*, *flap*, *lap*), and "-ake" denotes something unsteadily moving (*shake*, *quake*, *snake*, *rake*, or *wake*). Combining two phonesthemes can even result in a sensible joint meaning: "str-" for linearity and "-ap" for flatness together form "strap", a flat long kind of tape (Magnus, 1999). Or "tw-" for pinching (*twitch*, *tweak*, *tweak*, or *tweeze*) combined with "-irl" for spinning movements (*curl*, *furl*, *whirl*, or *swirl*) results in "twirl", a tight spin (Nuckolls, 1999). Thus, grammatically, phonesthemes come close to morphemes because of their meaning-bearing property but do not fulfill all of their characteristics (Bergen, 2004; Bolinger, 1950; Elsen, 2017a).

Some researchers hold the view that such systematicities in language are different from iconicity, assuming separate mechanisms behind these phenomena (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Elsen, 2017b; Thompson & Do, 2019). The rationale behind that refers to the seemingly arbitrary relation between the respective phoneme (cluster) and its meaning. An often-cited example is that it is not apparent why the phoneme combination "gl-" stands for words related to light or vision. Yet, using some imagination, a potentially iconic link can be found: A closer look at the phonesthetic words of this group – such as *glow, glitter, glisten, glare, glint*, or *gloss* – reveals that not light or vision per se are represented here but that it seems to be particularly about the perceptual experience of light being reflected from a very smooth or even liquid surface. Platon in his work "Cratylus" had proposed that "gl-" expressed the "detention of the liquid or slippery element" (Jowett, 1892) – such as if a liquid bounces against something. Accordingly, the *glug* sound of a liquid being stopped from time to time – when swallowing it or while unevenly pouring it out of a bottle – is a more direct form of onomatopoeia. From here, it is not too digressive to associate, for example, the light reflections from a lake's surface on a sunny day, when water waves are bouncing against a stone or a boat or gets irregularly aroused by wind, with similar "gl-" sounds.

Objectors against the phenomenon of phonesthemes often cite counter-examples with the respective consonant clusters that do not fit into the same semantic category as the phonesthemes (e.g., *glove*, *glad*, or *gloom* do not represent anything referring to light or vision). However, statistics proves them wrong: As an early example, Leonard Bloomfield compiled a comprehensive list of Germanic roots (Bloomfield, 1909, 1910) that allowed a quantification of the degree of such sound-meaning correlations. It became apparent that phonologically based groupings are not semantically random but hold throughout the entire vocabulary of Germanic (Magnus, 2013). Bloomfield had even detected the same vowel-meaning patterns that later empirical studies would confirm: “the forms with high-pitched vowel – the scale, running downward, is *i [y], e [ö], a, o, u* – represent high-pitched, clear, shrill sounds, fine, small, bright, flashing, quick, sharp, clear-cut objects or actions; the forms with low-pitched vowel express low, muffled, rumbling, bubbling, sounds and dull, loose, swaying, hobbling, slovenly, muddy, underhand, clumsy actions or objects. The *a* vowel will often express the large, the loud, the rattling, the open” (Bloomfield, 1909, p. 252).

Hence, both phenomena – iconicity and phonesthemes – seem to be related, but in most cases the link became obscured during language evolution: “It is likely that there do exist quite subtle iconic pressures toward associating certain types of images with certain types of articulations – pressures of a strength that can easily be washed out by sound change, borrowing, or other incidental factors. If this is the case, these factors can only be seen as statistical tendencies over large bodies of data extending over time and across unrelated languages“ (Rhodes, 1994).

Recently, sophisticated corpus-based studies are uncovering such subtle statistical patterns of iconicity, revealing, for example, that sensory words are more iconic than abstract words (Winter, Perlman, Perry, & Lupyan, 2017), or that particular language sounds are highly predictive of size in adjectives referring to size (Winter & Perlman, 2021), or of shape in object names (Sidhu, Westbury, Hollis, & Pexman, 2021). Hence, more and more types of sound-meaning systematicities are detected in real language corpora, not only in artificial pseudowords – further uncovering the vastness of iconicity patterns in language.

Several scholars and researchers meanwhile agree that there are different degrees of iconicity, which is reflected in the distinctive labels that they haven chosen for them: Margaret Magnus, for example, differentiates between *true iconism* and *clustering* (the latter one referring to phonesthemes) but suggests on the basis of her empirical data “that Clustering is still subject to the constraints of the inherent character of the sounds” (Magnus, 2001).

Taylor & Taylor (1965) distinguished *subjective phonetic symbolism* – which is detectable by observers – from *objective phonetic symbolism* – the overrepresentation of particular sounds in words of particular connotations. Their *Feedback Theory of Phonetic Symbolism* declares how both are interconnected: “if speakers of a language show subjective phonetic symbolism, then the language will come to show the same pattern of objective phonetic symbolism, and that the pattern of objective phonetic symbolism in the language similarly influences the development of subjective phonetic symbolism in the speakers” (I. K. Taylor & Taylor, 1965).

Also quite established is the differentiation between *primary* and *secondary iconicity* (Sonesson, 1997): Whereas primary iconicity results from the *perception* of an iconic ground between content and expression, secondary iconicity relies on the *knowledge* about the existence of an iconic ground between them. Ikegami & Zlatev suggested “that the distinction between primary and secondary iconicity is more of a cline, defined by the degree to which the ‘sign function’ (i.e. knowing what an expression represents) is necessary for perceiving the similarity involved in iconicity” (Ikegami & Zlatev, 2007).

Even de Saussure, whose statement that phonemes are mainly arbitrary coding items with no particular meaning attached to themselves still strongly dominates linguistic viewpoints, actually suggested an evolution of language from indexical or iconic forms towards conventionalization, with words losing their original motivation (Saussure, 1916). The same process can be observed in sign language, in which gestures and facial expressions carry the full communicative burden (Goldin-Meadow, 1999): Newly created signs do iconically represent objects or actions but then gradually lose the original resemblance between form and referent (Frishberg, 1975). Hence, highly associative sign- or sound-meaning correspondences, foremost onomatopoeia, are to be located at one end of the evolutionary language scale, whereas phonesthemes are "getting close to the arbitrary end of the language scale" (Hinton, Nichols, & Ohala, 1994a, p. 5).

Yet, what if there is not a unidirectional process going on but rather a bidirectional one? The linguist Otto Jespersen had stated that whilst iconicity may merge into arbitrariness during evolution, also new patterns of iconicity can develop (he called it echoism or symbolism, though): "Words that have been symbolically expressive may cease to be so in consequence of historical development, either phonetic or semantic or both. […] On the other hand, some words have in course of time become more expressive than they were at first. […] Sound symbolism, we may say, makes some words more fit to survive, […] languages in course of time grow richer and richer in symbolic words" (Jespersen, 1922, pp. 406–411).

Nowadays, most studies on iconicity assume a spectrum from iconicity to arbitrariness due to language being "a system structured by competing tendencies" (Nuckolls, 1999). An etymological study of Germanic language changes shows that "the processes of creating iconicity and conventionalization are continuously interplaying in language change" (Carling & Johansson, 2014). Potentially, one tendency drives language towards arbitrariness, in particular, when it comes to effectiveness in language use or to expanding the vocabulary to abstract concepts (Gasser, 2004). In contrast, the other tendency supports iconicity, which is especially relevant for intuitively learning a new or one's first language (Nygaard et al., 2009), or to make stories lively by depiction and embodiment, as for example with ideophones (Perniss, Thompson, Vigliocco, Christiansen, & Wilcox, 2010; Schourup, 1993). The cognitive mechanisms that made people pick meaningful sounds to form words in the first place seem to be still effective nowadays (as shown by the name invention studies of Berlin, 2006, or Rummer & Schweppe, 2019) due to the associative way the human brain works (see further elaborations on this in chapter 1.8).

Following these assumptions of general principles, traces of sounds that somehow fit a word's meaning should hypothetically be detectable in all languages. And indeed, some universal iconicity could be found in a study looking at the basic vocabularies of more than 4,000 languages: 30 different concepts are significantly associated with particular human speech sounds, e.g., *small* with /i/, *nose* with /n/, *tongue* with /l/, or *round* with /r/ (Blasi, Wichmann, Hammarström, Stadler, & Christiansen, 2016). That these patterns occur persistently across unrelated languages means that they must have evolved independently, presumably due to the common associative naming behavior of humans throughout all times.

Over the course of the research history on sound-meaning correspondences in language as depicted in the last three subchapters, many different terms came up for this phenomenon: *sound symbolism, phonetic symbolism, sound physiognomy,* or *phonological iconicity* as well as several variations thereof. I will briefly elaborate on some definitions and thereby derive the term(s) that will be used for the continuance of this thesis.

Sound symbolism is one of the most frequently used terms and stands for the "direct linkage between sound and meaning" (Hinton et al., 1994a). However, some confusion comes from the different understanding of the word "symbol". Ferdinand de Saussure stated that "it is characteristic of symbols that they are never entirely arbitrary. […] They show at least a vestige of natural connexion between the signal and its signification" (Saussure, 1916). Otto Jespersen borrowed the German term "Lautsymbolik" from Georg von der Gabelentz and translated it into English as "sound symbolism" (Chang-Liang, 2018). A later also frequently used synonym thereof is the term “phonetic symbolism”, specifying that language sounds in terms of phonemes are referred to and not just any type of sound.

However, Charles Sanders Peirce, an influential philosopher of logic and semiotics at the turn of the 19th to the 20th century defined three types of signs: symbol, icon, and index (Peirce, 1931). Here, a symbol is an arbitrary sign, created by convention. For example, numbers and letters are symbols. An icon, however, somehow resembles the signified object or concept. Thus, a drawing of a tree would be an icon, or imitating a bird song would be iconic. With an index, the relationship between signifier and signified is a causal one that has to be learnt by experience. For example, smoke indexes fire, or leaves laying on the ground are indices for a tree nearby. With this differentiation in mind, the language phenomenon I am looking at in this thesis concerns iconic relations between the sound of a word and the associated concept. Stating that "any ordinary word, as 'give,' 'bird,' 'marriage,' is an example of a symbol", Peirce’s account of language (Peirce, 1931) initially did not differ from that of the linguist Ferdinand de Saussure who claimed that the relation between a signifier and the signified is arbitrary (Saussure, 1916). However, later in his life, Peirce conceded that any single sign may display some combination of iconic, indexical, and symbolic characteristics (Atkin, 2013). Referring this to the relation between word concepts and the language sounds constituting the respective word forms, it means that some degree of iconicity can also be found in mostly arbitrary word forms, even if this concerns just single sounds. According to Peirce’s nomenclature, such a resemblance – i.e. iconicity – between the qualities of the sounds in a word form – i.e. its phonological units – and the qualities of the word concept can be referred to as “phonological iconicity” (also see Schmidtke, Conrad, & Jacobs, 2014). In general, the term *iconicity* is already quite established for meaning-matching language phenomena with regard to different modalities, e.g., in sign language or research on gestures. Hence, I will use the terms *phonological iconicity* or just *iconicity* or – for more general statements – *sound-meaning correspondences* throughout the rest of this thesis.

Regarding the wide scope of iconicity degrees in language and the large quantity of dimensions of meaning that get reflected by these sound-meaning correspondences (see chapter 1.3 and the examples for phonesthemes in this chapter), the question arises, how this spectrum could be categorized or even quantitatively operationalized in a biopsychologically meaningful way to be used for experimental studies that are aimed at holistically investigating the role of phonological iconicity in language processing.

1.5. Affective dimensions of meaning

To approach the quantitative measurement of meaning, it makes sense to have a look at the lexical level first: Charles Osgood and colleagues developed an objective method to measure the meaning of words through connotations – the *semantic differential* (Osgood, 1952; Osgood & Suci, 1955). While a denotation is the literal meaning of a word that directly refers to the concept of the word (e.g., a table), each word also has several connotations, i.e., associative meanings depending on the context a word is used in or the subject that is using it (e.g., for *table* possible connotations would be *stable*, *wooden*, *heavy*, etc.). To investigate such connotations, Osgood et al. used 50 continuous scales with opposite adjectives as bipolar anchors such as *strong – weak, high – low*, or *hot – cold*. On these scales, subjects had to allocate different concepts – e.g., *tree*, *house*, *lady*, *flower, etc.* – derived from their immediate association. Factor analyses showed that many seemingly different meanings are actually equivalent and can be represented by a single dimension. All in all, most connotative meanings converged on three major dimensions: evaluation (or valence[7]), activity (or arousal), and

[7] Throughout subsequent literature on these terms, either as semantic or cognitive-affective dimensions, the used denominations have changed. The terms outside the brackets are the ones originally used by Osgood et al. (1952, 1955), the terms inside the brackets are the ones that will be used throughout this thesis.

potency, all of which together accounted for approximately 50 % of the total variance in the associative judgments (Osgood et al., 1952, 1955).

Interestingly, as early as 1904, Wilhelm Wundt had proposed a tripartite theory of subjective affect containing the dimensions *like – dislike*, *aroused – calm*, and *tense – relaxed* (Wundt, 1904, pp. 44 ff.). The last dimension especially refers to expectations (of external stimuli) and action readiness (tense because of unfulfilled expectation or action, relaxed after fulfilled expectation or action). Together, all three dimensions compose the quality of an affect resulting from a stimulus, and each affect represents a process that prepares an action (p. 59), whether this be an actual movement such as running or fighting, or just a facial expression, or the uttering of words. For Wundt, affect and language are so tightly interwoven, that he deals with his theory of affect in the first volume of his seminal "Völkerpsychologie" [folk psychology] entitled "Die Sprache" [language]. The parallels of Osgood's dimensions of meaning derived by the semantic differential and Wundt's dimensions of affect are manifest. If language is understood as a specification of human cognition it makes sense that the meaning of basic linguistic elements such as words (or even sounds – see chapter 1.6) is qualitatively composed in an analogous way as the meaning of the basic aspect of cognition – affect.

In 1974, Mehrabian & Russell also presented a framework of how environmental stimuli lead to behavioral responses by the mediation of affective responses in form of arousal, pleasure, and dominance. However, Russell saw dominance as a more cognitive factor and thus excluded it from the *Core Affect* (Russell, 1980). Thus, his *Circumplex Model of Affect* states that any affective experience – be it either free-floating mood or emotional feelings that can be attributed to internal or external causes – can be represented in the two-dimensional space of valence and arousal (Russell, 2003).

Since Margaret Bradley and Peter Lang had shown that direct ratings of valence and arousal for pictures or other types of stimuli correlate highly with the semantic differential method (Bradley & Lang, 1994), researchers have collected such affective word ratings for large language corpora in different languages: For English, there are, for example, the *Affective Norms for English Words* (ANEW) with 1,034 words (Bradley & Lang, 1999) as well as a more recent database containing ratings for 13,915 English lemmas (Warriner, Kuperman, & Brysbaert, 2013). The most extensive affective database for German words is the *Berlin Affective Word List* (BAWL) with more than 2,900 words (Võ, Jacobs, & Conrad, 2006; Võ et al., 2009), meanwhile available as an extended version with 5,695 words (Schmidtke & Conrad, 2018).

Affective word corpora have since been used extensively in computational and psycholinguistic research to understand affective word or text processing (for a review on the use of the BAWL database see (Jacobs et al., 2015). Computational linguistics' approaches – so-called sentiment analyses – have shown that a large portion of a text's overall emotional content can be reliably predicted from the affective connotations of the words in it (Anderson & McMaster, 1982; Bestgen, 1994; Paltoglou, 2014; Thelwall, Buckley, Paltoglou, & Cai, 2010). Psycholinguistic research, on the contrary, is mainly focusing on the perception and processing of affective words. Behavioral studies, for example, are measuring reaction times and error rates (Briesemeister, Kuchinke, & Jacobs, 2011; Kousta, Vinson, & Vigliocco, 2009), or eye movements (Võ et al., 2008), whereas neurophysiological studies use functional neuroimaging (Hsu, Jacobs, Citron, & Conrad, 2015; Kuchinke et al., 2005) as well as electroencephalography (EEG) to understand which neural substrates and what time windows are relevant for affective word processing (see the reviews by Kissler, Assadollahi, & Herbert, 2006, and Citron, 2012). Relevant findings of some EEG studies will get elaborated on in chapter 2.

In many studies, contrasting positive against negative words yields no processing differences, but both positive and negative words show processing advantages over neutral words (Kissler & Herbert, 2013; Kousta et al., 2009). Yet, this might have to do with the U-shaped distribution of arousal being plotted against valence in the above mentioned large affective word corpora (Bradley & Lang, 1999; Võ et al., 2009): The distribution shows that positive as well as negative words are more arousing than neutral words (see Figure 1.2). Hence, the mentioned effects of valenced over neutral words might actually be effects of arousal. Thus, high arousal values as well as high absolute distances of valence from neutral represent a higher affective salience of the respective word stimuli compared to less arousing or valenced stimuli (Lewis, Critchley, Rotshtein, & Dolan, 2007). Affective salience here means that attention gets automatically directed towards affective items as behaviorally relevant targets by which they receive prioritized processing (motivated attention model; Lang, Bradley, & Cuthbert, 1997; Schupp et al., 2004). Salience will play a relevant role in several aspects throughout this thesis.

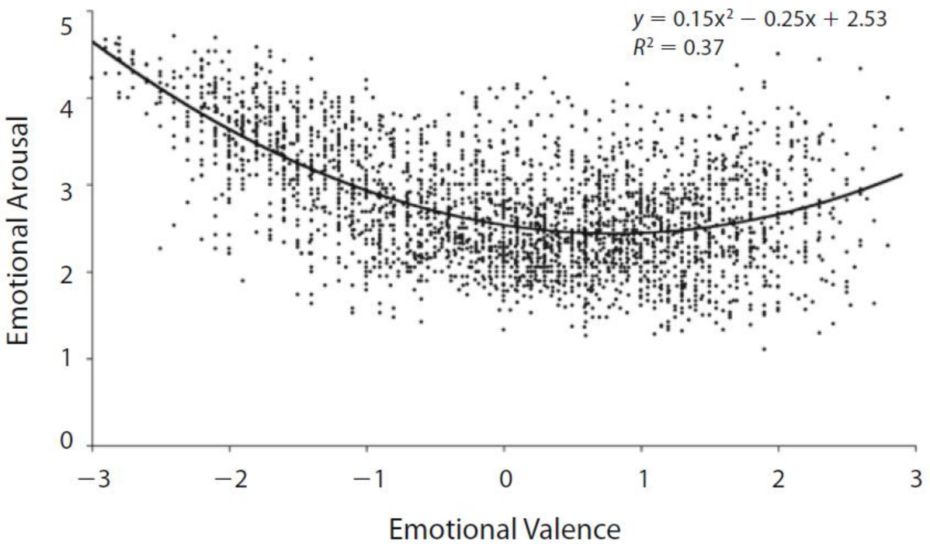

Figure 1.2: Distribution of mean values for all words of the Berlin Affective Word List Reloaded as a function of rated emotional valence and arousal (republished with permission of *The Psychonomic Society, Inc.*, from Võ et al., 2009, Cross-validating the Berlin Affective Word List. *Behavior Research Methods*, *38*(4), 606–609; permission conveyed through *Copyright Clearance Center, Inc.*).

1.6. Affective meaning at the sublexical level

If valence and arousal are the basic affective dimensions of meaning at the lexical level, why should the same affective categorization not be applicable to the level of language sounds? It does not seem far-fetched to assume that language sounds are loaded with affectivity: Looking at the animal kingdom, most terrestrial vertebrates communicate via acoustic signals within their species. They all have a larynx (or syrinx in birds), where contracting muscles can tighten specific tissue (such as the vocal chords) so that it starts to oscillate when air is exhaled (Titze, 1994). The fundamental frequency of the vocalization is subsequently filtered by the resonance conditions in the vocal tract. The physiological condition of the individual, first and foremost its arousal state, affects this process mainly by influencing the muscular tension and actions that are responsible for the produced sound qualities. For example, “broader affective-semantic relationships are common in the vocal communication systems of nonhuman primates, where harsh, noisy, and punctuate (i.e., strident) sounds are commonly produced in situations of high arousal and often also hostility and aggression, whereas smoother more harmonic (i.e., sonorant) sounds are associated with situations of lower arousal and also positive affiliation and contact” (Nielsen & Rendall, 2011). Similar patterns regarding arousal differences can be found in many other vertebrates: “with an increase in arousal, vocalizations typically become longer, louder and harsher, with higher and more variable frequencies, and they are produced at faster rates” (Briefer, 2012). For valence, however, the picture is less clear – positive situations and feelings in animals are harder to detect and mostly less intensive than negative ones (Boissy et al., 2007). Hence, they are confounded with arousal which is mostly low in comforting or calming interactions of individuals which leads to lower frequencies in their vocalizations as well (think, for example, of purring cats). That these patterns of arousal-sound correspondences are universal and can be perceived across species was shown by a study that let humans with different language backgrounds listen to pairs of animals' vocalizations – among them calls of birds, pigs, monkeys, elephants, and even frogs or alligators – and asked them to decide which of the two calls was the more aroused one. The proportion of correct answers was always higher than expected by chance (50 %) and even reached more than 80 % for six of the nine tested species – including 95 % for emotional human vocalizations as a reference (Filippi et al., 2017).

Coming from that perspective, it is very likely that affective vocalizations also marked the origin of human language long before the first word: "Thoughts were not the first things to press forward and crave for expression; emotions and instincts were more primitive and far more powerful" (Jespersen, 1922, p. 433). Already Aristotle had presumed that the human voice was primarily used to express affective symbols: "Now spoken sounds are symbols of affections in the soul, and written marks symbols of spoken words. And just as written marks are not the same for all men, neither are spoken sounds. But what these are in the first place signs of – affections of the soul – are the same for all; and what these affections are likenesses of – actual things – are also the same" (translation by Ackrill, 1963). A combination of uttering affective sounds to communicate inner emotional and motivational states and copying environmental sounds that carry affective information (hissing of snakes, roaring of lions, songs of birds, etc.) with the human vocal apparatus might have allowed our ancestors to communicate important information about potential danger or prey. These affective connotations of sounds may have survived in words for affective concepts until today.

However, not very many experimental studies on sound symbolism or phonological iconicity have dealt with affective associations to sounds so far. In the middle of the last century, after the method of the semantic differential had been developed with word concepts (Osgood & Suci, 1955), it was extensively used for other stimuli types as well, including phonemes.

Greenberg & Jenkins (1966) asked subjects to rate a series of consonants as well vowels on several adjective scales and conducted principal factor analyses as well as Varimax analyses. The factors they found using the principal factor analysis mostly yielded oppositions of the inherent qualities of the phonemes, for example, a stop vs. continuant dimension resulted for the consonants and a front vs. back dimension for the vowels. Yet with the Varimax analysis, for consonants factors emerged that are pretty similar to valence (*front, bright, free, open, good, light, smooth,* and *liquid* for the positive side), arousal (*passive, warm, mellow, rounded, still,* and *dull* for the low-arousing side), and potency (*empty, weak, thin,* and *small* for the low-potency side). For vowels, a potency and a valence dimension (called "pleasant-unpleasant" factor) were found as well as a factor that they called "acute-grave" but which could vaguely be compared to arousal – with *thick, falling, soft, dull, large, low,* and *wide* on the low-arousing side and *thin, rising, hard, sharp, small, high,* and *narrow* on the high-arousing side (Greenberg & Jenkins, 1966).

Around the same time, in another research group, adjectives were rated for their pleasantness as well as for their euphony. From these affective meaning and sound ratings, they derived the insight that stop sounds are overrepresented in bad as well as cacophonic words, whereas nasal sounds are more represented in good and euphonic words (Johnson, Suzuki, & Olds, 1964). From the knowledge they had gathered with these ratings about affective sound qualities, they created artificial words that were supposed to be good or bad, and paired them with positive or negative adjectives, so that for each adjective there were two artificial words with a different phoneme content to choose from. Hearing subjects matched the artificial words significantly in the predicted way, but deaf subjects only yielded chance level, which supports the notion of phonological iconicity (in the paper being called phonetic symbolism; Johnson et al., 1964).

Furthermore, some text-analytical studies using poetry – where emotions play an important role (also see chapter 1.10) – had looked at sound-emotion correspondences: Fónagy collected normative ratings for poems of the Hungarian poet Petöfi to categorize them as either "tender" or "aggressive". Then, he analyzed their phoneme content and found that tender poems contained more sonorants, such as /n/, /m/, or /l/, whereas aggressive poems contained more voiceless plosives, such as /k/ or /t/ (Fónagy, 1961). However, comparing happy versus sad poems, Auracher et al. found more plosives in happy poems and more nasals in sad poems – cross-linguistically verified among German, Russian, Ukrainian, and Mandarin native speakers (Auracher, Albers, Zhai, Gareeva, & Stavniychuk, 2011). This contradiction looks to me rather like an effect of arousal, as both tender and sad poems are low arousing whereas happy and aggressive poems are high arousing. Hence, harsher consonants such as plosives appear more often in high-arousing poetry and softer consonants such as sonorants in low-arousing poetry – which is in accordance with the characterization of differently aroused animal calls as described above.

Recently, affective dimensions are more and more getting back into the focus of iconicity research. Regarding valence, a corpus study analyzing five languages showed that valence is expressed in an iconic way mainly by the first phoneme of a word (Adelman, Estes, & Cossu, 2018). Experimental studies showed that particular vowels carry positive or negative meaning depending on their muscular articulation patterns overlapping with facial emotion display (Rummer & Schweppe, 2019; Yu, McBeath, & Glenberg, 2021; and forthcoming Körner & Rummer, 2021). Arousal, on the other hand, has been found to be mediating shape-sound correspondences such as in the bouba-kiki effect, in terms of spiky shapes and respective pseudowords being rated as more arousing than round shapes and pseudowords (Aryani, Isbilen, & Christiansen, 2020).

If language sounds carry affective meaning, the question of how to measure such affective meaning of a phoneme or phoneme cluster comes up. Using an approach similar to words being rated on the two or three main dimensions that resulted from the original semantic differential, one study determined the semantic features of 24 English and German consonants by letting people rate them on six scales including evaluation, activity, and potency (Gnatchuk, 2015): Several statistical analyses were conducted with the obtained ratings, identifying significant associations and comprehensive semantic profiles for each consonant which are suggested to be used for the creation of iconic brand names for industrial goods. However, such studies should be viewed with some skepticism as the paradigm of rating single phonemes is a very artificial approach. Language sounds rarely exist in isolation but get influenced by the co-occurring sounds in a word. Thus, to analyze phonemes in isolation might not yield representative results as, for example, the subjects who have to rate them might rather think of interjections. Yet these underlie different mechanisms – compare, for example the meanings of /i:/ and /o:/ as interjections: /i:/ is an expression of disgust whereas /o:/ is uttered in amazement. But from text-analytical and experimental iconicity studies we know that /i:/ is generally perceived as positive and /o:/ as rather negative. Looking back at the study by Hanna Gnatchuk (2015), one conspicuous finding is, for example, that she identified /m/ as the consonant with the highest symbolic (iconic) potential in the German language but with the lowest symbolic potential in the English language. This strikes me as very unreliable due to the facts that English and German are closely related languages sharing many cognates, and that /m/ has been found to be one of the strong universal associations for concepts such as *mother* or *breasts* (Blasi et al., 2016).

The interwovenness of potential phoneme meanings with word meanings makes it hard to develop a clear bottom-up approach to quantify affective properties on the phonological level.

Hence, one can also turn the perspective around and approach the issue in a holistic top-down manner, starting with affective words, under the assumption that the properties of the phonemes composing these words contribute to their affectivity. As early as 1966, David R. Heise showed that certain phonemes occur systematically in words with particular affective meanings (also referring to the three basic dimensions of affect evaluation, activity, and potency) by calculating affective scores for each phoneme based on the affective ratings of the words (from a word list of 1,000 frequently used English words) this phoneme appears in. Furthermore, he predicted affective sound scores for each word by averaging the affective scores for each phoneme contained in a word (each of the affective dimensions being treated separately), and could show that the predicted sound scores correlate significantly with the affective rating scores (Heise, 1966).

Cynthia Whissell followed a similar approach in her extensive work on phonosymbolism. In 1994, she published her Dictionary of Affect, comprising 4,746 words rated on valence and arousal. Investigating the phoneme distributions in this dictionary, she correlated the number of times each phoneme occurred in each word with the ratings for that word. The resulting correlations were weak but significant (see the "Eighth Step" in Whissell, 1999, p. 36). She concluded that "phonemes do not have the power to induce full-blown emotional responses of the lock-and-key type under any and all conditions, but they do have some of this power some of the time" (Whissell, 1999, p. 37).

Inspired by such approaches, my cognitive psychology colleagues at the Freie Universität Berlin looked for sound-meaning correlations in an extended version of the BAWL (*Berlin Affective Word List*) – since its first publication with 2,900 words (Võ et al., 2006, 2009) normative affective ratings for more words have been collected continuously so that it amounted to 5,695 words (Schmidtke & Conrad, 2018, supplemental material) when the work for this thesis was conducted. This extended version of the word database will be called BAWL-E throughout this thesis. Merging the BAWL-E with the CELEX database (Burnage, 1990) made the phonemic transcriptions for each word available.

We found that particular phoneme clusters are associated with particular connotative meanings – herewith referring to the affective ratings for valence and arousal: For example, looking at all words starting with /kr/ – 50 in total (see Figure 1.3A) – 27 are negative and medium to high-arousing (valence ratings below -0.5 on a scale from -3 to +3 and arousal ratings above 2.4 on a scale from 1 to 5) but only 12 are positive (valence rating above 0.5) and medium to high-arousing. The remaining 11 words are of low arousal (thereof 1 negative, 3 neutral, and 2 positive) or medium arousal (all 5 of neutral valence). Hence, one could say that the word initial unit /kr/ has an affective meaning of rather negative valence and high arousal. On the opposite, the word initials /fr/ show a very different distribution (see Figure 1.3B): From 48 words, only 13 words are negative (1 high-arousing, 10 of medium arousal, and 2 low-arousing), but the double amount – 26 words – are positive (1 high-arousing, 17 of medium arousal, and 8 low-arousing). Thus, the overall affective meaning of /fr/ is rather positive with medium arousal.

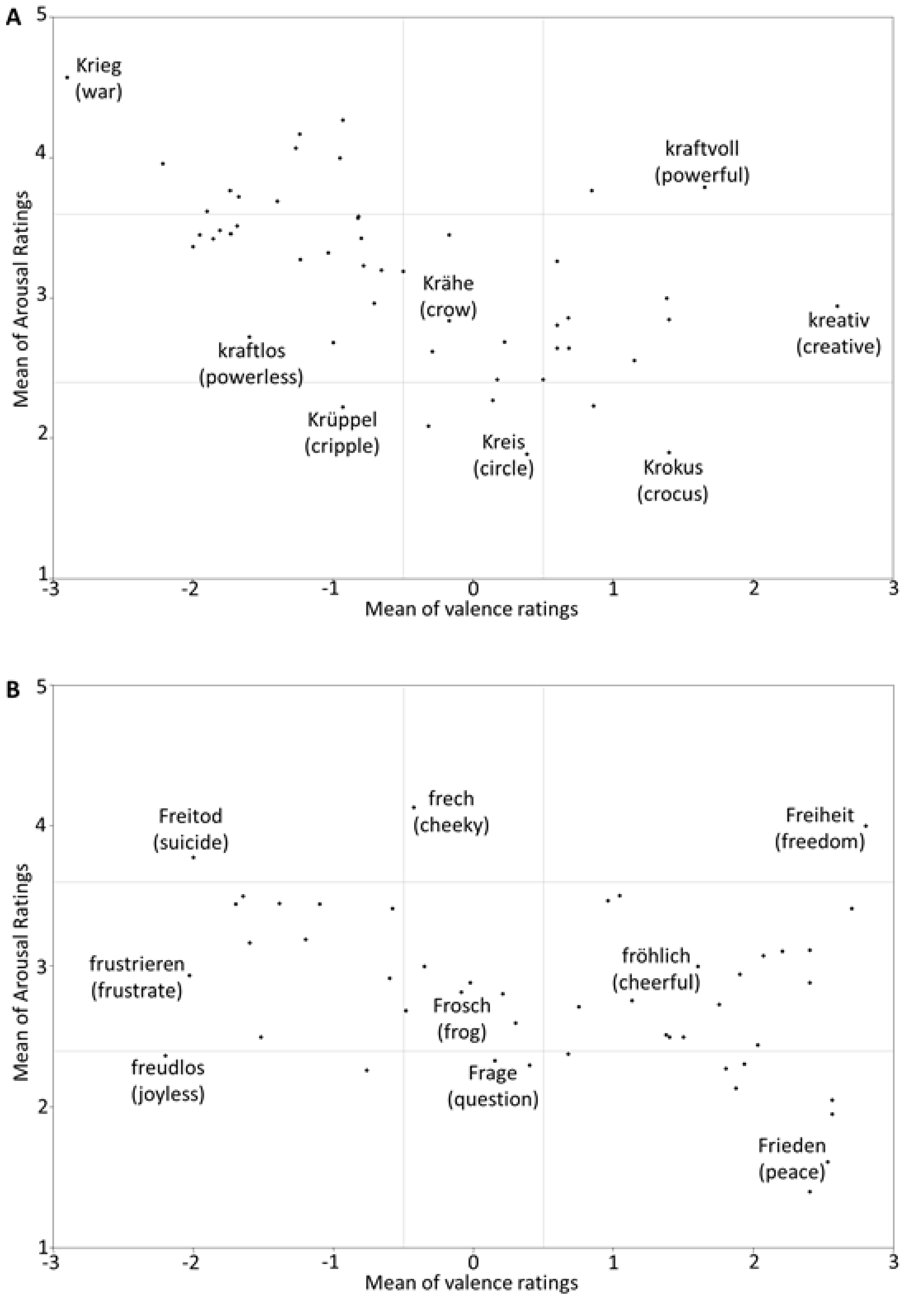

Figure 1.3: Distributions of mean valence ratings vs. mean arousal ratings of all words in the BAWL-E database starting with /kr/ (A) and /fr/ (B), some example words are quoted.

To quantify the affective value for each sublexical unit holistically, we calculated the mean of the lexical affective values (the normative valence and arousal ratings, respectively) of all the words in the BAWL-E that contain this particular sublexical unit in the same position – as syllabic onset, nucleus, or coda. We used such subsyllabic units as sublexical units and not single phonemes, for the problem with phonemes is that they are not pronounced one by one but get co-articulated with the phonemes that precede or follow them in a word (such as /kr/ is not articulated like /k/ and /r/). Thus, depending on its respective phonetic context the same phoneme is not always pronounced consistently. Also, experimental studies of language processing support the importance of subsyllabic segments as perceptual units encoding phonology (Nuerk, Rey, Graf, & Jacobs, 2000), with syllable onsets being particularly relevant for the reading process (Brand, Giroux, Puijalon, & Rey, 2007). As could be seen in chapter 1.4, also phonesthemes are mostly consonant clusters, for example functioning as a word's onset. Sorting all sublexical units by their calculated affective values – valence as well as arousal – allows one to look for known patterns from the iconicity research. For valence we know from the literature, for example, that the vowel order /i:/ > /e:/ > /a:/ > /o:/ > /u:/ reflects the spectrum from positivity on the left to negativity on the right end of the scale with /a:/ as a rather neutral vowel. However, this is not reflected in the valence ranking of the *sublexical affective values* – here, we rather find /u:/ as the most positive vowel, followed by /e:/, then /i:/ > /a:/ > /o:/. But /u:/ and /e:/ also show the smallest frequencies of appearance in the database which might have led to distortions – if we just look at the most frequent vowels /i:/, /a:/, and /o:/, the familiar order shows up.

For arousal, it can be seen among the onsets that single consonants are usually to be found at the low-arousing end whereas high-arousing onsets are mostly complex consonant clusters such as /dr/, /Sr/, /st/, /kr/, or /Str/. I will come back to that type of potential salience in chapter 2. Among single consonants, the following order from high to low arousal can be found: /r/ > /f/ > /t/ > /p/ > /k/ > /w/ > /h/ > /n/ > /s/ > /g/ > /b/ > /d/ > /m/ > /l/ > /z/. While the order is not exactly as in the literature (especially in the middle field the arousal values are very similar to each other and might change positions if further words get included in the database), the strong opposition of /r/ at the high-arousing end and /l/ at the low-arousing end immediately reminds us of Jacob Grimm's statement: "Of the consonants, /l/ will designate what is gentle, and /r/ what is rough" (Grimm, 1864, p. 284). Also, Fónagy (1963) had asked subjects about the symbolic relations of /r/ and /l/, and the majority found /r/ to be "wild, pugnacious, manly, rolling, and harder". Furthermore, the unvoiced plosives /p/, /t/, and /k/ being more arousing than the voiced plosives /g/, /b/, and /d/ is in line with findings from other studies (Fónagy, 1961). Also associations of voiced plosives with spiky shapes (Köhler, 1947; Ramachandran & Hubbard, 2001) could be interpreted as belonging to the dimension of arousal (Aryani et al., 2020). Thus, arousal seems to be the affective dimension that is somewhat better represented by the *sublexical affective values*.

To derive an approximate affective sound value for each word in the BAWL-E, the affective values of all sublexical units in a word were averaged. To avoid circularity issues, the rating values for the word whose affective sound was to be calculated were not included in the calculation for the affective values of the sublexical units it consists of. By implication of the averaging approach, the resulting scale widths of the affective word sound values were much narrower (valence: −0.7 to +0.7; arousal: 2.5 to 3.2) than those of the normative rating scales. Because such an affective word sound value is just a simple mathematical approximation of the potentially real affective word sound, it will be called the *sublexical affective potential* of a word throughout this thesis.

The *lexical affective content*, as presented by the normative affective word ratings, and the *sublexical affective potential* were found to correlate in a highly significant manner (arousal: $r = 0.13$, $p < 0.0001$; valence: $r = 0.1$, $p < 0.0001$). Using a linear regression model fit, it can also be said that the *lexical affective content* is predictable by the *sublexical affective potential* (arousal: $R^2 = 0.02$, $F(1,5734) = 95.09$, $p < 0.0001$; valence: $R^2 = 0.01$, $F(1,5734) = 63.03$, $p < 0.0001$). Although only a very small portion of the variance in the affective word ratings can be explained by the *sublexical affective potential*, the model is highly significant. This means that generally, there is a very subtle but highly significant affective iconicity present in the BAWL-E corpus.

For the *lexical affective content*, we find the characteristic U-shaped distribution when plotting the words' arousal ratings as a function of their valence ratings (see chapter 1.5 and Figure 1.2). However, does the same relation exist at the sublexical level of affectivity? Plotting the sublexical arousal potential as a function of the sublexical valence potential (see Figure 1.4) results in a quite condensed circular aggregation of the majority of the words with a seemingly linear extension of some outliers to the negative high-arousing corner as well as to the positive side – in a more distributed manner. Yet, a quadratic regression model still yields a slightly better fit ($R^2 = 0.28$, $F(2,5733) = 1113.75$, $p < 0.0001$) than a linear regression model ($R^2 = 0.25$, $F(1,5734) = 1910.73$, $p = 0.00$). Therefore, a similar valence-arousal relation as among the lexical affective content can be assumed at the sublexical affective level.

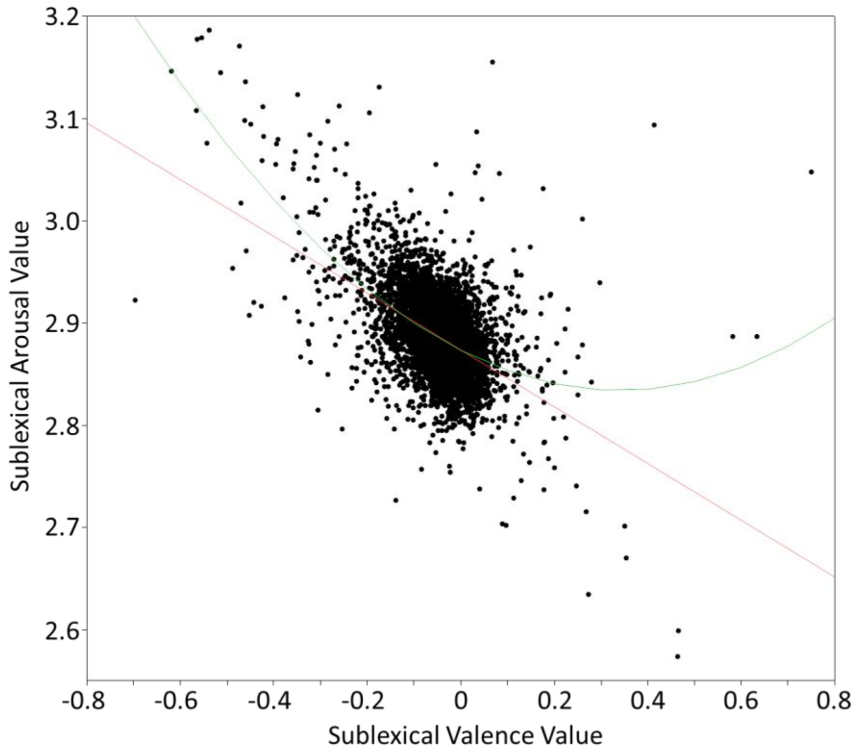

Figure 1.4: Distributions of the sublexical valence and arousal potentials of all words in the BAWL-E database.

All in all, our holistic top-down approach can only be seen as a very rough approximation to the quantification of iconicity in the German lexicon – yet it serves as a starting point into that complex subject. An analogy to painting may demonstrate the level of abstraction: On highly expressionistic pictures one cannot identify any depicted objects or subjects anymore, but one can still get the basic affective tone of that picture: A sad picture would not be painted in bright colors, for example. Or a vivid scenery would not be illustrated with one big, resting square. In the way that basic arousal and valence potentials even speak from abstract, expressionistic pictures, the *sublexical affective potential* used for the following studies reflects a very basic affective tone of words.

1.7. Working hypothesis & research questions

The **working hypothesis** for this thesis is derived from the corpus studies in the BAWL-E as explicated in the previous chapter and supports Jakobson's postulation that "sound is necessarily linked with, and thus informs about, meaning" – this sound-meaning relation being mostly of an indirect, or mediating, nature (Jakobson & Waugh, 1979). All the following studies assume a basic iconic relationship between the affective qualities of sublexical units in the German language (= *sublexical affective potential*) and the affective content of the words.

Research question 1: The first question is concerned about the psychophysiological reality of the *sublexical affective potential*. Our top-down approach towards the calculation of the *sublexical affective potential* of a word is merely a statistical procedure, looking at phonological units being associated with particularly affective words. The only psychological aspects in here, so far, are the valence and arousal values for each word in the database, being derived from normative ratings. Thus, the open question is whether the statistically derived *sublexical affective potential* of a word, being constituted by potentially affective phonological units, poses a construct with a measurable psychological reality. Do words whose *sublexical affective potentials* differ in terms of valence or arousal ranges show different reaction times or error rates in classic psycholinguistic tasks such as the lexical decision task? And what are the underlying mechanisms in the brain as far as can be examined with todays neuroimaging techniques? I will explicate some theoretical aspects of how language and (affective) meaning are processed in our minds to develop a hypothesis to those questions. However, it has to be kept in mind that this first research question about the psychophysiological reality of the *sublexical affective potential* is a purely explorative one, as to my knowledge, no one else had tried a similar approach before.

Research question 2: If a psychophysiological reality of the *sublexical affective potential* can be shown empirically, the next question asks whether the s*ublexical affective potential* also has an influence on language processing at other language levels. This question is inspired by Roman Jakobson's postulation that the distinctive features of language sounds function as signs – as words do, too – and that each level of language signs brings new particularities of meaning, which participate in the meaning of the whole (Jakobson, 1980, p. 20).

2a) First, I will examine if the affective features at the phonological level do specifically interact with those at the word level. A model of distributed and parallel language processing will be used to derive a hypothesis explaining why and how the affective features of the sublexical and the lexical language level could interact.

2b) If an interaction between the sublexical and the lexical language level can be shown, of course, it leads to the question of whether those effects can further "climb the ladder" to higher language levels. To be able to grasp how meaningful features at the sublexical level, the lexical level, as well as inter-lexical levels could interact to create the general affective meaning of a text, an applied approach using poetry will be pursued.

To be able to derive respective hypotheses and suitable methodical approaches to answer the research questions above, I will discuss theories about how the human brain processes language, in particular regarding the lexical and the sublexical level, in the following three chapters.

1.8. Deriving hypothesis 1 from the DMF model

Human language is processed in the brain. Hence, the principles of language must be closely related to the functioning principles of the human brain. As the brain is composed of neurons organized in a complex network structure, our semantic knowledge is also assumed to be organized as a distributed network (Patterson, Nestor, & Rogers, 2007). Neuroimaging research has shown that concepts that have to do with perception (like seeing an object) or action (such as handling an object) are represented in those brain areas that are responsible for perceiving and acting. Thus, when a concept becomes mentally present (by thinking, talking or dreaming about it), this emerges from weighted activity of neurons within or connecting between those property-based regions (Martin, 2007). These semantic neural networks have developed throughout our lifespans due to Hebbian learning mechanisms: "Any two cells or systems of cells that are repeatedly active at the same time will tend to become 'associated,' so that activity in one facilitates activity in the other." Or in short: "What fires together, wires together," (both citations from Hebb, 1949). This way, different aspects and contexts of our experiences become connected to each other.

When it comes to associating a concept's meaning with a spoken or written word form, the concept's neural network and the neurons being involved in the perception of the acoustic or visual word form properties are active at the same time and rhythm and thus become associated with each other in the form of an extended network. Hence, the word form, for example the word "rose", is now connected to (and thus can evoke) a mental picture of the flower, its smell, the pain of getting stung by its thorns, as well as the emotional context of the learning situation, such as the smile on the face of a loved one receiving a bunch of roses. Throughout the situational contexts in which a concept gets neurally evoked, the network that includes the concept properties as well as the word form properties also acquires the affective properties of the context, a process which is called *contextual learning* (Braun, 2015; Fritsch & Kuchinke, 2013).

Hence, there is no semantic lexicon for language items only (such as a network of words) – the mental semantic network must be understood as a part of the whole cognitive architecture including long-term memory (Engelkamp & Rummer, 1998). The DMF theory of Herrmann & Grabowski (1994) gives a structured overview about which aspects might be connected in such a mental semantic network: DMF stands for *dual*, *multimodal*, and *flexible* – three properties that characterize the conceptual structures of the mental semantic network. First of all, the DMF theory divides between concepts and words – similar to the semiotic theory of Peirce – which is the *dual* aspect. Concepts are the knowledge elements that humans have stored cognitively about things, events, people, and further facts and experiences – reaching from concrete objects to abstract ideas. Words are the word forms – spoken, written, and potentially in multiple languages – associated with each concept. But as shown above, words and concepts are not only associated with each other but represent an aggregate of several cross-linked components. These components can be categorized regarding their modality – whether they are sensoric, motoric, abstract, or emotive-evaluative. This is the *multimodal* aspect of the theory, and even further modalities would be possible. Sensoric components represent sensoric sensations such as the smell, sound, touch, taste, or look of a concept. But also a word form has a look (its written form) and a sound (when it is pronounced), and can even leave a tactile impression if it is written in Braille, for example. Motoric components represent the imagination of handling an object or articulating or writing a word, whereas the abstract components refer to all the semantic knowledge about a concept or word. Finally, the emotive-evaluative components represent the affective assessments obtained through contextual learning – whether I like an object or a person, or had bad experiences with it. For word forms, apart from having favorite words or letters due to idiosyncratic reasons, emotive-evaluative factors could be implicit connotations of the sounds constituting a word, subsumable under the basic affective dimensions valence and arousal – our *sublexical affective potential*. Thus, getting back to the mental network picture, when an external or internal stimulus is perceived, the relevant multimodal components become activated and the activation spreads to further connected components. Thus, a concept or a word that comes to mind in a specific situation actually is the neuronal activation pattern of several simultaneously activated components. But not all of the components that are related to a word or a concept in such a mental semantic network always get activated – the strength of the connections between the components (and hence the threshold for single components to become activated) can be different depending on the learning history of the individual. The *flexible* aspect of the DMF theory refers to these temporally differential patterns of activated components and underlines that the situational context determines which part of the network gets activated and which is inhibited.

My first research question concerns the psychological reality of the statistically derived *sublexical affective potential*. In a lexical decision task, I want to contrast words with different affective potential at the sublexical level. This task targets the *dual* aspect of the DMF theory: The subjects are presented with several existing German word forms as well as pseudowords – orthographically valid German letter strings. To be able to decide whether the presented stimulus is a real word or a pseudoword the concept behind a real word needs to become activated. Only if a concept for a particular letter string can be found in the mental semantic network, can this letter string be classified as an existing word form. Hence, only for real words, will both the word form nodes and the concept nodes become activated in the mental semantic network.

Then, for the words, the *multimodal* aspect of the DMF theory becomes relevant: Here, the focus is especially on the emotive-evaluative component of both the word concept and the word form. For the word concept, the emotive-evaluative component is operationalized by the normative affective ratings from the BAWL-E database (valence as well as arousal), whereas the *sublexical affective potential* (also valence as well arousal values) functions as the operationalization of the emotive-evaluative component of the word form. As each word activates a neuronal pattern that simultaneously includes the emotive-evaluative components of both the word concept and the word form, this allows to comparatively measure the behavioral effects and the neurophysiological correlates of both the *lexical affective content* and the *sublexical affective potential*. As both are affective measures, similar patterns could be expected. **Hence, my first hypothesis is that contrasting differences in the *sublexical affective potential* should yield similar results as same-type contrasting differences in the *lexical affective content.***

Regarding the *flexible* aspect of the DMF theory, however, one cannot be sure that both of the emotive-evaluative components become activated in the same strength during a lexical decision task, which per se does not focus on emotional aspects. Yet, a task that would try to make a presumably implicitly perceived feature such as the *sublexical affective potential* explicit – for example, by rating or making decisions regarding the perceived word sound affectivity – might have distorting effects on the measurements. If one does not know the basic effects of a psychological construct, it might be hard to properly interpret further manipulations. Hence, there remains a risk that effects of the *sublexical affective potential* might be too small or too insignificant to be measured with a lexical decision task. Thus, a suitable neuroimaging method needs to have a high resolution in a relevant dimension to be measured. In comparison to neuroimaging methods such as functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) which are good at localizing processes in the brain, the electroencephalography (EEG) can resolve brain waves at a very high temporal resolution (at the scale of milliseconds) and has been used to assess the processing steps of affective word processing. Thus, even if no behavioral effects regarding reaction times or error rates might be found for the *sublexical affective potential*, in the EEG, very small effects could still become visible. The EEG experiment used to investigate the psychophysiological reality of the *sublexical affective potential* is described in chapter 2.

1.9. Deriving hypothesis 2a from the Interactive Activation Model (IAM) and Predictive Coding

The previous section has highlighted the network character of the semantic lexicon. Thinking of an object and its name is represented in such a distributed semantic network as a specific pattern of simultaneous neural activity across several nodes. The dynamics of how the nodes of such a pattern become activated, i.e., in which order and direction, are relevant for my research question 2a regarding potential interactions of affective features at the phonological level with those at the word level.

The *Interactive Activation Model* (IAM) by McClelland & Rumelhart (1981) is a word recognition model which assumes a spreading activation from one excited node of the semantic network to all the nodes that are connected to it. For example, if someone is listening to someone else speaking, the auditory perception of an acoustic feature of the spoken word activates the respective node and from there all connected nodes. Apart from the matching phonemes and phoneme clusters that contain this acoustic feature, also the connotations of these phonemes become activated – such as their *sublexical affective values*. When the phoneme (cluster) is fully understood, the co-activation of similar phonemes is inhibited, but words that potentially contain this phoneme (cluster) become pre-activated. When the next phoneme (cluster) is perceived and its node excited, only the nodes of words that still contain the complete phonemic information stay activated. Finally, when there is no ambiguous network activation anymore, the word is fully perceived, i.e., the lexical access is completed, and the established neural activation pattern does not only contain the word form and its components' nodes but also all the components' nodes that represent the respective word concept and its connotations.

Yet, most importantly, the activation can move forward as well as backward between the activated nodes of the network. Due to these bidirectional feedback options this activation model is called "interactive". Figure 1.5 displays the potential interconnections between the linguistic levels from the acoustic or visual feature level up to the word form level. This means, for example, an auditorily perceived word can also elicit an inner perception of the orthographic information about this word, for the letter nodes become excited either by the phonemes' or the word's nodes. And vice versa, when reading a word via the visual pathway, the parallel co-activation of the phonemes encoded by the letters or the whole word form leads to a mental auditory percept of the word – the so-called inner voice (also see chapter 2). Because of this parallel and highly interactive language processing, information from different hierarchical levels is continuously compared and integrated. Paul Bouissac (1995) realized that connectionist models can be advantageous in explaining iconicity in language: "What seems to make these models particularly interesting for handling the problem of linguistic iconicity is that they enable us to relate linguistic output to complex knowledge structures based on the coordination of a large number of different components through which the interactions of simultaneous pieces of information or constraints are computed".

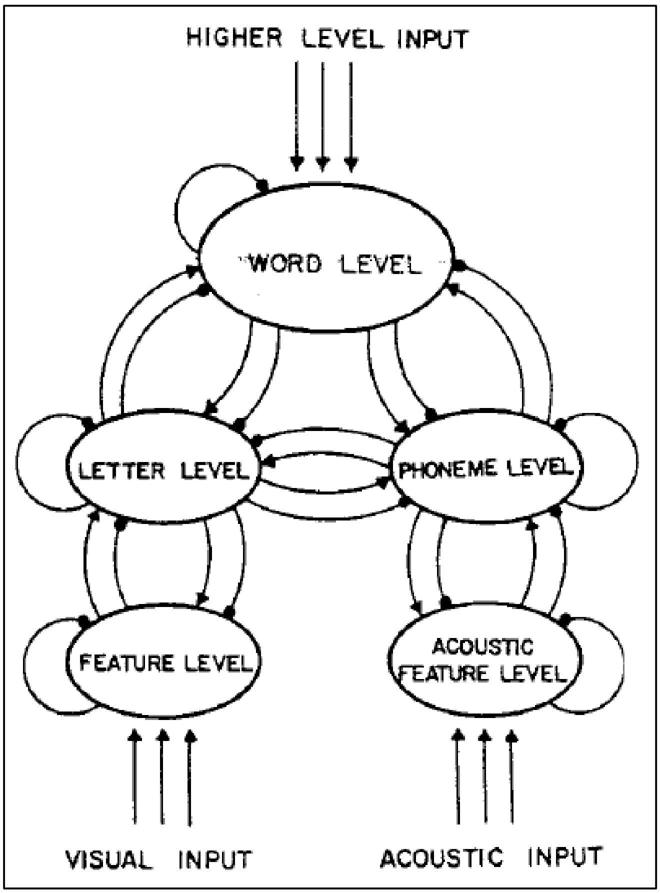

Figure 1.5: Processing levels involved in visual and auditory word perception. (From McClelland, James L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account on Basic Findings. *Psychological Review*, *88*(5), 375–407. Copyright © 1981 by *American Psychological Association*. Reproduced with permission.)

Most relevant for developing a hypothesis to the question of interaction between the sublexical and lexical level is the fact that in the IAM the phoneme and the word level continuously interact in both directions. **This leads to my hypothesis that, if there is psychophysiologically relevant affectivity at the sublexical level (see research question and hypothesis 1), it should directly interact with the affectivity at the lexical level.**

This leads to the question of what such an interaction would look like in terms of measurable output. Although there is an overall correlation between the *lexical affective content* and the *sublexial affective potential* (as shown in chapter 1.6), there are words for which the affective categorizations at the lexical and sublexical level match, but also many for which they do not.[8] Hence, in the first case, there is an iconic relationship at the affective level, but in the second

[8] In the BAWL-E database we did not determine the affective consistency of all words, but only of those words in the negative to neutral spectrum that were considered for the stimulus selection of the EEG experiment. Among these, 1,058 words were consistent and 768 words were inconsistent in their lexical and sublexical affectivity.

case there is none – or even an anti-iconic relationship. As Jespersen did say: “There are words which we feel instinctively to be adequate to express the ideas they stand for, and others the sounds of which are felt to be more or less incongruous with their signification” (Jespersen, 1922, p. 398). Such an explicit mismatch between the sound and meaning of a word should have a measurable psychophysiological reality, too – again a case where the term *salience* seems appropriate. Per definition, salience refers to stimuli that are prominent with regard to their context – often this is produced by novelty or unexpectedness (Taylor & Fiske, 1978). According to the rather new *Predictive Processing Framework* (a continuation of the parallel and interactive processing account), our brains are continuously generating predictions regarding incoming sensory information. Such predictions facilitate perception and cognition, as only prediction errors need to be transmitted from sensory cortices to higher cognitive processing – thus reducing redundancy and saving energy (Bar, 2007; Friston, 2010; Huang & Rao, 2011). Therefore, sensory features that were not predicted by the brain, in our case affective speech sounds that are inconsistent with the lexical meaning, should lead to salience or surprise and might delay the lexical access – something that can be measured with behavioral and electrophysiological methods. At the behavioral level, the influence of affective sound on lexical meaning has been shown by two studies of colleagues: In a letter search task, Schmidtke & Conrad (2018) found that in general, letter strings of high arousal were responded to faster than low-arousing letter strings, no matter if they actually appeared in the superimposed words or not. In particular, subjects were faster to identify that a letter string did not occur in a target word (i.e., a non-match) when the respective letter string was of high sublexical arousal while the word in which one had to look for it had a low-arousing meaning. This interaction between arousal at the sublexical and the lexical level shows that affective iconicity influences language processing.

Also in judging words' lexical arousal levels in a two-alternative forced choice task (i.e., whether a word means something exciting or calming), subjects responded faster if the relation between the lexical and sublexical arousal of the words was iconic than when it did not match (Aryani & Jacobs, 2018). Hence, the subjects realized unconsciously whether there was a sound-meaning correspondence in the words and were slowed down in their judgements when opposed affective properties at different language levels led to incorrect predictions that needed to be resolved.

Hence, my further refined hypothesis would be that at the neuronal level affective consistency between the sublexical and the lexical level can ease word processing (due to a low prediction error), whereas affective inconsistency would hinder word processing (due to a high prediction error). An EEG study contrasting word stimuli with an either consistent or inconsistent manipulation of the *sublexical affective potential* and the *lexical affective content* to test these hypotheses is described in chapter 3.

1.10. Deriving hypothesis 2b from the Neurocognitive Poetics model (NCPM)

Even if the psychophysiological reality of the *sublexical affective potential* and an interaction thereof with the *lexical affective content* of single words can be shown in a lexical decision task and the respective EEG data, what does that mean for real-life use of language? Can the very small and presumably implicit effects of the *sublexical affective potential* affect, for example, our reading experience in its entirety? A more applied study in which subjects read and appraise real texts, here poems, will be used to give the answers to those questions.

The renowned linguist Roman Jakobson not only proposed that all phonic components of language sounds are equipped with meaning, but also that "each level above brings new peculiarities of meaning: they change substantially by climbing the ladder which leads from the phoneme to the morpheme and from there to words […], then go through various levels of syntactic structures to the sentence, then to the groupings of sentences into the utterance and finally to the sequences of utterances in dialogue" (Jakobson, 1980, p. 20).

But most studies looking at interactions of different text levels start at the word level and investigate interactions with higher affective text properties (Bayer, Sommer, & Schacht, 2010; Ding, Wang, & Yang, 2014; Jiang et al., 2014; Hsu et al., 2015; Lüdtke & Jacobs, 2015). Thus, guided by Roman Jakobson's proposition, I will explore whether the *sublexical affective potential* of the words in a text, for example a poem, influences the text's affective perception over and above the influence of affective lexical and inter-lexical features.

Reading in general is perhaps "the most remarkable specific performance that civilization has learned in all its history" (Huey, 1908). Yet, it not only transfers factual information and ideas across time and space, but also emotional content. Although reading misses the highly affective aspects of direct communication, such as prosody, facial display, and gestures, it still has a strong potential to evoke emotions and feelings due to the cross-modal nature of our neural semantic networks (as described above according to the DMF theory). For example, reading about a blossoming garden evokes memories of the different looks and even smells of roses, violets, or hyacinths, and raises aesthetic appreciations of beauty and peacefulness. Whereas reading about someone breaking up with someone else can induce strong feelings such as sadness, anger, or despair, if the reader is deeply involved in the preceding story and has built up affective connections to the protagonists.

Besides these immersive emotions arising from the plot of a story, reading can also elicit aesthetic emotions, triggered by textual features and style, that contribute to the overall affective impact. This is the part where sublexical affectivity might come into play as well, especially if it elicits stylistic salience, for example, by a frequent occurrence of rare or complex consonant clusters in a text, or by the usage of words that have a *sublexical affective potential* which is inconsistent with their *lexical affective content*. Yet, the question remains whether such presumably very small effects of the *sublexical affective potential* can be distinguished and measured among all the different influential factors.

A genre of literature that allows easy access to aesthetically affective perceptions might be poetry. Poetry is less restricted by conventions, frequently deals with emotional or affective issues, and often plays with sounds, for example rhymes or alliterations. In this regard, poetry is the realm of written language where iconicity has the biggest potential to come forward, and poets know how to utilize it: "The /ou/ sound may suggest the sound of beasts, may invoke the images of darkness and mourning, /l/ and /m/ may have a soft effect, […] the sounds salvage a fraction of the genuine experience, of reality. We get a plosive for an explosion, the /sch/ sound for a storm, the /r/ sound for the stream of air from a brawl" (Fonagy, 1961). In their book "The Sound Shape of Language" Jakobson & Waugh (1979) state that "nowhere is the direct interplay of sound and meaning more salient than in poetry, which is based on similarity (and equivalence) at all levels" (pp. 4-5). This citation connects the topic of iconicity in language with its effectiveness at different language levels. Hence, for a first exploratory study on potential interactions of affectivity at different text levels, using poems as text material seems promising.

With regard to how poetry is perceived cognitively by the reader, the Neurocognitive Poetics Model (NCPM, Jacobs, 2015a, 2015b) proposes two neural pathways which are processed in parallel: On the one hand, there is a fast and implicit processing of so-called *backgrounding* elements. Such *backgrounding* processes are responsible for detecting familiar features, for example known word forms, which in accumulation can lead to fluent reading and immersive states (Jacobs, 2015a, 2015b). On the other hand, there is a slower and explicit processing of so-called *foregrounding* elements (Hakemulder, 2004; Miall & Kuiken, 1994; van Peer, 1986).

Foregrounding processes are mostly elicited by salient text features. This could be infrequent words, unusual syntax, metaphors, or lyrical stylistic devices such as rhymes, alliterations, oxymorons, etc. Their salience leads to a slow-down in processing, as the unexpected needs to be made meaningful (which would correspond to a higher prediction error in the terms of *Predictive Coding* – see the previous chapter).

As can be seen from Figure 1.6, foregrounding elements can be found at each text level and among different text feature groups such as phonological, metric, morpho-syntactic, or semantic features (and more are possible, yet not listed here). The Neurocognitive Poetics Model includes the sublexical level, reflecting the fact that at the time of its publication, several studies on phonological iconicity, including mine for this thesis, were already ongoing in this research group. Hence, the *sublexical affective potential* – being either congruent or incongruent to the *lexical affective content* – could lead to salience at a semantic level, or single phonemes or phoneme clusters could become salient at the phonological level due to a very frequent appearance in the poem and/or their highly affective sublexical values.

However, in a whole text, such as a poem, the potentially small effects of sublexical features could easily get lost in the multitude of further influences from other text levels and features. Only some rather strong and unidirectionally effective sublexical affective characteristics could provide a whole poem with something like a basic affective tone. Such a *basic affective tone* has been quantified by colleagues of mine using a new statistical operationalization of affective sound at the sublexical level (Aryani, Kraxenberger, Ullrich, Jacobs, & Conrad, 2016), taking into consideration the frequencies of occurrence of the subsyllabic segments in each poem (Aryani, Jacobs, & Conrad, 2013).

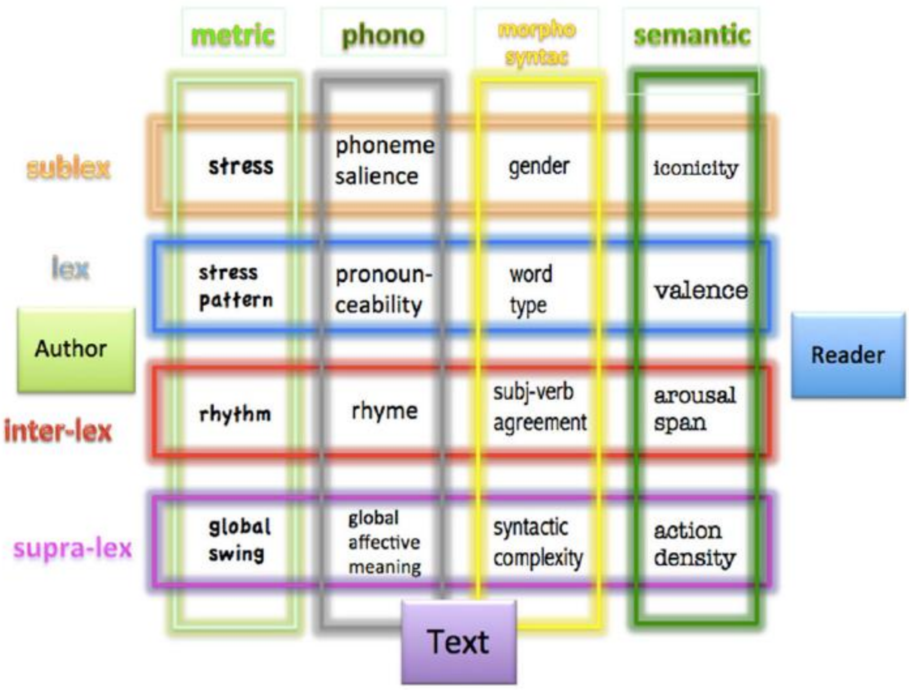

Figure 1.6: Examples of foregrounding elements organized in a 4x4 matrix crossing feature groups with text levels of poetry perception. (From Jacobs, A. M. (2015). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, Article 186, doi:10.3389/fnhum.2015.00186, licensed under CC BY 4.0.)

Hence, only subsyllabic segments whose observed frequencies in a poem were significantly higher than the respective frequencies that could be expected on the basis of a linguistic corpus (SUBTLEX-DE, Brysbaert et al., 2011) were considered as salient and were scored with the respective *sublexical affective values* as derived from the normative affective ratings in the BAWL-E database (valence as well as arousal separately). Furthermore, the weighted mean of the respective *sublexical affective values* of all salient segments in a poem were divided by the standard deviation of a corresponding null model. The resulting sigma values reflect the extent of deviation from an expectable value (see methods' section in Aryani et al., 2016). The respective sigma values of both valence and arousal represent the *basic affective tone* of a poem.

For 57 modern-style poems (unrhymed, free verse) from Hans Magnus Enzensberger‘s volume “verteidigung der wölfe“ (defense of the wolves; 1957) the *basic affective tone* was calculated and its relation to both the author’s affective categorization of the poems (19 friendly, 21 sad, and 17 spiteful poems) and readers’ affective perception of the poems (assessed by ratings) was examined. Results show that the *basic affective tone* differs significantly between the three poem groups in accordance to the author’s labels. Also, the *basic affective tone* alone already accounts for up to 20 % of the variance in the affective ratings of the poems. Thus, the salient subsyllabic segments in connection with their *sublexical affective values* do indeed function as foregrounding elements that affect the emotional reading experience.

Now, my research question concerns how this *basic affective tone* interacts with affective features from other text levels. According to the IAM, the processing of affective information from the sublexical, lexical, and presumably also inter-lexical level takes place in a parallel manner which also coincides with Jakobson’s statement (see above). **Thus, my hypothesis is that foregrounded affective features from the different text levels influence each other and thus the global affective perception of the poems by the readers.** This is tested in a study (see chapter 4) using the same poem volume by H. M. Enzensberger and the same poem ratings – plus two additional ones – as in Aryani et al. (Aryani et al., 2016). In addition to the *basic affective tone*, further variables from the lexical level – sigma values of the valence and arousal means of the words in a poem – as well as the inter-lexical level – affective peaks, ranges, and dynamic changes within the lexical affective content – are used in a regression analysis to predict the general affective meanings of the poems, and also to examine the specific influences of the affective features from the three text levels.

2. EEG study part 1 – Phonological iconicity electrifies: An ERP study on affective sound-to-meaning correspondences in German

This chapter is published as:

Ullrich, S., Kotz, S. A., Schmidtke, D. S., Aryani, A., & Conrad, M. (2016). Phonological iconicity electrifies: An ERP study on affective sound-to-meaning correspondences in German. *Frontiers in Psychology, 7,* Article 1200. doi:10.3389/fpsyg.2016.01200,

2.1. Abstract

While linguistic theory posits an arbitrary relation between signifiers and the signified (de Saussure, 1916), our analysis of a large-scale German database containing affective ratings of words revealed that certain phoneme clusters occur more often in words denoting concepts with negative and arousing meaning. Here, we investigate how such phoneme clusters that potentially serve as sublexical markers of affect can influence language processing. We registered the EEG signal during a lexical decision task with a novel manipulation of the words' putative *sublexical affective potential*: the means of valence and arousal values for single phoneme clusters, each computed as a function of respective values of words from the database these phoneme clusters occur in. Our experimental manipulations also investigate potential contributions of formal salience to the *sublexical affective potential*: Typically, negative high-arousing phonological segments – based on our calculations – tend to be less frequent and more structurally complex than neutral ones. We thus constructed two experimental sets, one involving this natural confound, while controlling for it in the other. A negative high-arousing *sublexical affective potential* in the strictly controlled stimulus set yielded an early posterior negativity (EPN), in similar ways as an independent manipulation of

lexical affective content did. When other potentially salient formal features at the sublexical level were not controlled for, the effect of the *sublexical affective potential* was strengthened and prolonged (250-650 ms), presumably because formal salience helps making specific phoneme clusters efficient sublexical markers of negative high-arousing affective meaning. These neurophysiological data support the assumption that the organization of a language's vocabulary involves systematic sound-to-meaning correspondences at the phonemic level that influence the way we process language.

Keywords: sublexical, lexical, affect, emotion, language, EEG, event-related brain potentials (ERPs), phonological iconicity, sound symbolism, sound-to-meaning correspondences.

2.2. Introduction

Most people would probably agree that not all words sound "neutral". But is it just personal taste or idiosyncratic individual experience that some words sound nicer and others rather harsh to us? Or do, on the contrary, sublexical phonological patterns possess systematic affective connotations? And if so, might these relate systematically to the meaning of words? A potential associative or even physical resemblance between sound and meaning of a word is called *phonological iconicity* in terms of Peirce's typology of semiotic elements (Peirce, 1931; see also Aryani, Jacobs, & Conrad, 2013; Perniss, Thompson, & Vigliocco, 2010; Schmidtke, Conrad, & Jacobs, 2014), challenging the conventional linguistic view that the relationship between the signifier and the signified be arbitrary (de Saussure, 1916). Note that our use of the term "sound" in this paper refers exclusively to phonological constituents of words themselves, not to speaker related issues such as prosody or the speaker's identity or affective state (for research on the latter ones see, for example, Belin, Bestelmeyer, Latinus, & Watson, 2011; Hellbernd & Sammler, 2016). This conforms with the traditional literature on *sound symbolism,* which also posits that specific speech-sounds – phonemes – words are made of, may carry specific meaning (Allot, 1995; Jakobson, 1937).

Internal relations between phonological aspects and semantic meaning of words show most directly and prominently in onomatopoetic expressions (that typically describe acoustic phenomena by mimicking them): e.g., bears growl, snakes fizz, babies babble, or water splashes, sprinkles, squirts, drops, or drizzles. On a more abstract level, e.g., phonaesthemes involve the correspondence of specific sublexical patterns (typically word initial phoneme clusters) to specific semantic word fields (Firth, 1930). For instance, many English words related to vision and light start with “gl-”: glance, glitter, gloom, glisten, glare, or gloss – while many words related to the nose start with “sn-”: snore, sniff, snort, snuff, snoop, or sneeze (Wallis, 1699; Bloomfield, 1933). Although the reasons for the evolution of phonaesthemes remain somewhat opaque, Bergen (2004) could show in priming experiments that these subtle statistical associations influence language processing. Other systematic sound-to-meaning correspondences have also been found to support word learning (Lockwood, Dingemanse, & Hagoort, 2016; Nygaard, Cook, & Namy, 2009).

That the sound of a word and its signified semantic concept may, in general, share a common quality has already been discussed by Socrates in Plato’s Cratylus (Plato, transl. 1892). Throughout the last century, a number of empirical psychological studies have investigated how potential correspondences between sublexical language sounds and attributes of meaning influence human perception of, e.g., size, shape, lightness, pleasantness, or excitement. For instance, back vowels (a, o) are perceived as bigger, heavier, or darker than front vowels (i, e), as has been shown, for example, by Sapir (1929) who asked people to connect pseudowords such as MAL and MIL with either a large or a small object. Other researchers replicated and refined these findings on vowels and extended them to consonants, showing, for example, that people perceive front consonants as smaller and more pleasant than back consonants, or voiced consonants as darker and larger than unvoiced consonants (Newman, 1933; Folkins & Lenrow, 1966). In general, such phenomena subsumed under the terms *sound symbolism* or *phonological iconicity* (for reviews see Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Perniss et al., 2010; Perniss & Vigliocco, 2014; Schmidtke et al., 2014) involve the view that that the sound of a word and the signified concept share a common quality (see already Humboldt, 1836, or Plato, transl. 1892). As a potential cause, it has been proposed that language may have phylogenetically evolved from the imitation of natural sounds (Darwin, 1871; Plato, transl. 1892). Cross-language replications of, e.g. the kiki-bouba phenomenon – people, including toddlers, consistently match pseudowords such as *kiki* or *takete* preferentially to spiky shapes, vs. *bouba* or *baluma* to rounded shapes (Davis, 1961; Köhler, 1929; Maurer, Pathman, & Mondloch, 2006; Werner, 1934, 1957; also see Westbury, 2005) – suggest *phonological iconicity* to be a common feature of language in general, spurring theories about the biological origin of language (Ramachandran & Hubbard, 2001).

As communication of affect could be seen as a primordial feature of human communication (Jackendorff, 2002), *phonological iconicity* may well extend to affective meaning communicated through language – potentially since its very origins (see Darwin, 1871; Kita, 2008; Morton, 1977; Perniss & Vigliocco, 2014). The basic dimensions of affective meaning in the most influential emotion models (Bradley et al., 1992; Russell, 1978, 1980, 2003; Watson & Tellegen, 1985; Wundt, 1896) are those of valence and arousal, accounting also for a major amount of variance of semantic meaning according to semantic differential techniques (Osgood & Suci, 1955). Interestingly, analyzing the phonological content of 1,000 English words rated for valence and arousal, Heise (1966) found that certain phonemes occur significantly more often in words of a specific affective meaning (see also Whissel, 1999, 2000). Conrad, Aryani, Schmidtke, & Jacobs (in preparation) recently applied this approach to a large-scale database of over 6,000 German words rated for valence and arousal (see also Aryani, Kraxenberger, Ullrich, Jacobs, Conrad, 2015). Their analyses reveal systematic sound-to-meaning correspondences concerning the use of certain phonemes or phoneme clusters in words of specific valence and arousal ranges – in particular representing a combination of high arousal and negative valence that might be summarized as denoting potential threat. To quantify these patterns, they computed *sublexical affective values* (SAVs) for single sub-syllabic phoneme clusters – representing syllabic onsets, nuclei, and codas – by averaging valence and arousal values of all words these units are part of in the database. The choice of these subsyllabic phonological segments instead of single phonemes is motivated by linguistic theories of syllable segmentation (Davis, 1982; Hall, 1992; Wiese, 1996). Accordingly, both experimental (Brand, Giroux, Puijalon, & Rey, 2007; Nuerk, Rey, Graf, & Jacobs, 2000) and simulation studies (Jacobs, Grainger, Rey, & Ziegler, 1998) of language processing support the importance of those segments as perceptual units encoding phonology in terms of syllabic onsets, nuclei and codas. Within the German database, SAVs for a number of such phonological segments show significant deviations from neutral global means (Conrad et al., in preparation[9]), suggesting an intrinsic affective potential of specific language sounds, which might accordingly serve as sublexical markers of affect, in particular concerning threat. Following this rationale, the average of SAVs for all phonological segments in a word – henceforth called *sublexical affective potential* – might predict the affective appeal of the

[9] This paper has so far (August 2021) not been published and, insofar as I am informed, will not be published in the originally planned version anymore. The German database (referred to as BAWL-E in chapter 1) was published as supplementary material of Schmidtke & Conrad, 2018. How the *sublexical affective values* were derived and how they correlate with lexical affective ratings is shown in chapter 1.6 of this thesis.

whole phonological word form at a sublexical level. Indeed, Conrad et al. (in preparation) reveal significant correlations of this *sublexical affective potential* with lexical valence and arousal ratings across the entire respective word database. These findings interestingly point toward *phonological iconicity* with regard to affective content as a systematic feature determining the organization of language (see also, Aryani et al., 2015).

2.2.1. The present study

In this study, we address the question of whether these numerical measures of sublexical affective values – derived from a large-scale normative database for the German language, reflecting systematic sound-to-meaning correspondences within this database – possess any psychological reality concerning the perception of language. In particular, we ask whether these sound-to-meaning correspondences or the underlying affective *phonological iconicity* of the German language would have any neuroscientific correlates during a standard lexical decision task using EEG measurements. If anything like sublexical markers of affective content, in particular threat, exists, those phonological segments typically occurring in words of high arousal and negative content should leave an impact on brain activity strong enough to be traceable with neuroscientific methods during the time course of language perception.

Furthermore, our study focuses on the potential role of formal salience for processes related to phonological iconicity. Concerning sublexical phonological units presumably encoding – according to the analyses of our database – negative high-arousing content, we consistently found structurally rather complex phonological segments (i.e., more than one consonant in a syllabic onset or coda) and phonological segments of low frequency of occurrence to appear preferentially in words of negative and high-arousing meaning. As high arousal is thought of as an early alert indicator attracting attention to potentially relevant stimuli (see Recio, Conrad, Hansen, & Jacobs, 2014, for ERP effects disentangling valence and arousal effects during visual word recognition), it seems intuitive that formal salience could be crucial for making a sublexical unit a most efficient “sign of threat” at the conceptual level.

Event-related potentials (ERPs) obtained via EEG measurement with its high temporal resolution are most suitable to study if, when, and how such phenomena influence cognitive processes. A number of psycholinguistic studies have already investigated effects of lexical affective content during visual word recognition using ERPs. Two main ERP components were found to be modulated by the affective meaning of words: The *early posterior negativity* (EPN), a component that is larger for emotion-laden words compared to neutral ones (Conrad, Recio, & Jacobs, 2011; Herbert, Junghöfer, and Kissler, 2008; Keuper et al., 2014; Kissler, Herbert, Peyk, & Junghöfer, 2007; Kissler, Herbert, Winkler, & Junghöfer, 2009; Schacht & Sommer, 2009), appears around 200-300 ms after stimulus onset. It was first reported in the context of emotional face and picture processing (Junghöfer, Bradley, Elbert, & Lang, 2001; Schupp, Hamm, & Weike, 2003; Schupp et al., 2004), hence presumably reflecting general, modality-independent affective processing. The EPN is assumed to mirror fast and effortless detection of emotionally significant stimuli and thereby indexes natural selective attention (Olofsson, Nordin, Sequeira, & Polich, 2008). MEG studies reported that the neural loci of cognitive functions such as semantic memory, attention, and evaluation of emotional stimuli are involved in the formation of the EPN (Keuper et al., 2014). Furthermore, the late positive complex (LPC), appearing around 400-700 ms after stimulus onset, also proved sensitive to differences in the affective meaning of words (Conrad et al., 2011; Dillon, Cooper, Grent, Woldorff, & LaBar, 2006; Kissler et al., 2009; Schacht & Sommer, 2009). This late component is assumed to indicate more elaborated and task-dependent cognitive processing of affective or emotional stimuli. This includes, for example, continued stimulus evaluation such as categorization or memory updating. Useful reviews on ERP emotion effects in visual word recognition have been provided by Citron (2012) or Kotz & Paulmann (2011).

To investigate potential effects of affect encoded at the sublexical phonological level within the framework of known general emotion effects during visual word recognition, we used a design including a classical manipulation of *lexical affective content* together with a novel manipulation of *sublexical affective potential* in a standard visual lexical decision task.

Most theoretical reasoning on *phonological iconicity* assumes phonology as the source of respective effects. If these effects exist, they should, though, also show and might most effectively be studied during silent reading which has been shown to involve mandatory phonological processing (e.g., Abramson & Goldinger, 1997; Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009; Conrad, Grainger, & Jacobs, 2007; Van Orden, 1987; Ziegler, Jacobs, & Klüppel, 2001). The visual lexical decision task is the most standardized and most used research tool in the field of psycholinguistics. German is a shallow orthography with high grapheme-to-phoneme consistency, i.e., the presentation of specific German letter strings would evoke unambiguous phonological activations regardless of context and of whether a letter string is a word or not. Using a standard visual lexical decision task appears thus a reasonable initial step for the investigation of phonological iconicity effects in German. It provides both a methodological match to the available literature on emotion effects quoted above as well as an optimally standardized experimental context excluding potential distortion through auditory effects of, e.g., affective prosody or speaker identity.

At both the lexical and the sublexical level, our manipulations of *affective content* or *potential* involve the contrast between high arousal in combination with negative valence on the one hand, and low arousal combined with neutral valence on the other hand. This has both pragmatic and theoretical reasons: As already evident from Võ et al. (2009) and Schmidtke,

Schröder, Jacobs, & Conrad (2014), valence and arousal values of German words are characterized by a very tight correlation within the range of overall negative valence, but not within the positive valence range. That is, increasingly negative valence of concepts is generally associated with increasing arousal, whereas positive concepts can be either calm or exciting. As the sublexical affective values (SAVs) we use for the operationalization of the *sublexical affective potential* represent the average values of words containing a given phonological segment, it goes – to some extend – by itself that comparable correlations are given for SAVs. That is, the majority of phonological segments with negative valence also have rather high arousal levels, whereas positive valence and arousal SAVs are less related. Further, the combination of negative valence and high arousal fits best the assumed reason underlying these *phonological iconicity* phenomena: the encoding of threat at a sublexical level (see Conrad et al., preparation). Most of the phonological segments that might in general serve as icons of affective content – displaying statistically significant deviations from global neutral means – in the database of German words indeed follow this pattern of combining negative valence with high arousal. That is why the combination of negative valence and high arousal contrasted against neutral valence and low arousal allows for a most pronounced contrast – potentially leading to most pronounced effects – for this novel manipulation of *sublexical affective potential* taking into account both dimensions of the affective space.
As already mentioned, when considering phonological segments of syllabic onsets, nuclei, and codas rather than single phonemes, affectively deviant segments of negative valence and high arousal often also are structurally more complex – i.e., contain more phonemes – and of lesser frequencies of occurrence as compared to affectively neutral ones. To account for both types of effects – intrinsic SAVs on the one hand and formal salience on the other – as two potentially additive sources of *phonological iconicity* influencing affective processing during

language perception, we prepared two separate experimental stimulus sets to be presented in one and the same experimental session (see Conrad et al., 2007, and Conrad, Carreiras, Tamm, & Jacobs, 2009, for detailed elaboration of the methodological advantages of this approach):

- Set 1 involves the natural confound of SAVs with formal salience to capture a most natural picture of effects of affective *phonological iconicity* or *sublexical affective potential* – just the way they arise in the lexicon.
- Set 2 controls for this confound to allow for a clearer attribution of possible *sublexical affective potential* effects, disentangling them from phenomena of structural complexity or frequency of occurrence.

We predict effects of the sublexical manipulation to be strongest when SAVs are allowed to co-vary with formal salience. Further, if any effects at all would still be obtained for the sublexical manipulation controlling for formal salience, these effects might – with even more confidence – be considered evidence for sublexical encoding of affectivity, especially if they resembled ERP effects established so far for general emotion processing during lexical decision, and predicted for our second factor – affective content at the lexical level. In particular, such effects might be expected similar to an early posterior negativity (EPN), because sublexical effects should occur rather early during the time course of the reading process – or at least not later than lexical effects.

2.3. Materials and methods

2.3.1. Participants

41 native speakers of German, university students of the Freie Universität Berlin, participated in the experiment after giving informed consent. All were right-handed (Oldfield, 1971) with normal or corrected-to-normal vision. None of them reported neurological or language problems. Six participants were excluded from the final data analysis due to bad signal-to-noise ratio of ERP data so that data from 35 subjects (21 women; age range: 18 – 36 years, *M* = 26.7 years, *SD* = 4.2) were submitted to analyses. All participants received financial compensation.

2.3.2. Stimuli and design

We selected two separate sets (set 1: *maximally manipulated*; set 2: *maximally controlled*) of 312 German words each – containing between one and three syllables, with a maximum of nine letters length – from the extended BAWL database (Võ et al., 2009; publication of the extended version in preparation) as stimuli for the two experimental sets. Both sets involved twofold, independent manipulations of these two factors (each factor cell comprised 156 stimulus words):

- *lexical affective content* (negative valence and high arousal vs. neutral valence and low arousal)

  and

- *sublexical affective potential* (negative valence and high arousal vs. neutral valence and low arousal, based on mean SAVs per word)

Lexical affective content was closely controlled for between the two cells of *sublexical affective potential* and vice versa. *Lexical affective content* is operationalized in the database in form of rating values of valence on a scale from -3 to 3, and of arousal on a scale from 1 to 5. A word was entered in the *negative high-arousing lexical affective content* condition when the mean of its valence ratings in the database was more negative than -0.8 (furthermore, the sum of mean and standard deviation of the valence ratings for a word did not exceed 0) and its arousal ratings higher than 2.8. For the *neutral low-arousing lexical affective content* condition the valence ratings of the words had to be between -0.8 and 0.8 (and the standard deviation below 1) and the arousal ratings lower than 2.8.

The factor *sublexical affective potential* was operationalized as follows: We computed hypothetical affective values for sublexical segments (the aforementioned *sublexical affective values* – SAVs) as a function of the affective values of the words they occur in in our database of over 6,000 German words (Conrad et al, in preparation): We calculated valence and arousal SAVs for all given syllabic onsets, nuclei, and codas by averaging the rating values of words they form part of. We then averaged these values for all segments found in a single given word to obtain an estimate of the *sublexical affective potential* of this word. Naturally, the resulting scale widths for valence (-0.7 to 0.7) and arousal (2.5 to 3.2) of these *sublexical affective potential* values per word were much narrower than those of the *lexical affective content* rating scales. A word was entered in the negative high-arousing *sublexical affective potential* condition when its valence value was more negative than -0.05, and its arousal value higher than 2.9. For the neutral low-arousing *sublexical affective potential* condition the valence value of a word had to be between -0.04 and 0.45, and the arousal value lower than 2.9. Specifically for the sublexically neutral low-arousing words, additional attention was paid to the following selection criteria: If words contained single very negative or high-arousing phonological segments – albeit the overall mean fit in the neutral low-arousing category – they were excluded, for we assume that such single salient phonological segments could already attract enough attention to not let the whole word sound affectively “neutral” anymore. Stimulus characteristics are shown in Table 2.1. While our manipulation of *sublexical affective potential* is based on numerical mean SAVs across all phonological segments in a word, this certainly involves that specific segments are more likely to occur in one condition, e.g., negative/high arousal *sublexical affective potential*, than in the other (neutral/low arousal). To make our manipulation more transparent to the reader, Table 2.2 lists how many times specific phonological segments were used across conditions.

In both sets a large number of variables that are known to influence visual word processing (see Graf, Nagler, & Jacobs, 2005, for an overview) were controlled for between cells of the two factors (see also Table 2.1):

- word frequency (in terms of dec. logarithms + 1 of the word frequencies in the SUBTLEX database, Brysbaert et al., 2011)
- word length in terms of number of letters (max = 9) / phonemes / syllables (max = 3)
- imageability ratings
- word class (nouns, verbs, adjectives)
- stress pattern (on which syllable)
- composita patterns (classification of prefixes, suffixes, composita of two words, loanwords)
- number of orthographic and phonological neighbors (Coltheart, Davelaar, Jonasson, & Besner, 1977)
- frequency of orthographic and phonological neighbors (in terms of the dec. logarithm + 1 of the sum of the frequencies of all neighbors)
- specifically the number of orthographic and phonological neighbors with higher frequencies

In the *maximally controlled set* we further controlled for the following sublexical variables:

- syllable lengths (separately for each of the maximal three syllables and separately for orthographic and phonological syllables)
- token frequency of the first syllable (dec. logarithms + 1; for first syllable frequency effects see Carreiras, Álvarez, & de Vega, 1993; Conrad & Jacobs, 2004; Hutzler et al., 2004)
- token frequencies of all syllable segments (onset 1–coda 3, respectively, dec. logarithm + 1)
- morphological (CVC) structure of the onsets, nuclei, and codas respectively in all syllables
- combined consonant complexity patterns of each syllable (possible combinations: onset and coda simple [coded as 1], onset complex and coda simple [coded as 2], onset simple and coda complex [coded as 2], onset and coda complex [coded as 3])
- lengths of the nuclei vowels in each syllable (short vs. long)
- positional token frequencies (dec. logarithm + 1) of all bigrams and biphons in a word
- token frequeny (dec. logarithm + 1) of the respective last bigram and biphon of a word
- token frequency (dec. logarithm + 1) specifically of those bigrams spanning syllable boundaries

To assure best overall comparability between data for the two sets, all stimuli were presented in a unique experimental session to the same participants. Overlapping items, i.e., stimuli that were used in both manipulations, entered the final stimulus set only once to avoid repetition. Thus, a total set of 521 stimulus words was presented together with 535 pseudowords that were matched to word stimuli in length and number of syllables. Pseudowords included pseudo-homophones to assure a sufficiently difficult overall task environment where participants actually had to achieve lexical access for stimulus words. The pseudoword material involved a different experimental manipulation not addressed in the present study. All results presented in this paper refer exclusively to the word material possessing affective values at both the lexical and (hypothetically) the sublexical level.

2.3.3. Procedure

All Stimuli were presented visually in randomized order using “Times New Roman” font, size 24, in white letters on a black background in the center of a 17’’ computer screen with 80 cm distance to the participant’s eyes. Each trial began with the presentation of a fixation cross (500 ms) followed by a blank screen of 500 ms. The pseudo-randomized single word and pseudoword items were presented for 500 ms each and were followed by a blank screen that lasted until the key response had been carried out, followed by a scattered inter-stimulus interval of 700 to 1500 ms. The task of the participants was to decide whether the presented stimulus was a “word” or a “non-word” by pressing one of two respective push-buttons on a Playstation remote control. The labels “Wort” (word) and “Nichtwort” (nonword) were counterbalanced between left- and right-hand responses across participants. They were encouraged to respond as fast but also as accurately as possible. Before the actual experiment started, 10 initial practice trials (5 words, 5 pseudowords) were run. The whole experiment contained 1056 trials and was split into four blocks which lasted about 10 to 12 minutes each. In between these blocks participants were allowed to rest as long as they wished.

Table 2.1: Means and standard deviations for manipulated variables (Lex. Val. = lexical valence ratings, Lex. Aro. = lexical arousal ratings, Sublex. Val. = computed sublexical valence values, Sublex. Aro. = computed sublexical arousal values) and some of the control variables (LogFreq+1 = word frequency in terms of dec. logarithms + 1, Letters = number of letters, Syllables = number of syllables, Orth. Neighbors = number of orthographic neighbors, LogBigramFreq+1 = mean positional bigram frequencies in terms of dec. logarithms + 1, CVC complexity = combined consonant complexity patterns of each syllable) plus p values for tests of significant mean differences between the respective two conditions.

	Lex. Val.		**Lex. Aro.**		**Sublex. Val.**		**Sublex. Aro.**		**LogFreq+1**		**Letters**		**Syllables**		**Orth. neighbors**		**LogBigram Freq+1**		**CVC complexity**	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Maximally manipulated stimulus set																				
Lexical affective content																				
Negative valence high Arousal	-1.58	0.50	3.62	0.42	-0.08	0.10	2.90	0.06	2.18	0.73	6.08	1.35	1.91	0.69	1.89	2.16	3.42	0.39	1.39	0.52
Neutral valence low Arousal	0.11	0.38	2.33	0.29	-0.08	0.10	2.90	0.06	2.22	0.86	6.04	1.30	1.92	0.65	2.19	2.50	3.42	0.42	1.33	0.51
p	0.00		0.00		0.85		0.72		0.69		0.80		0.93		0.28		0.88		0.38	
Sublexical affective potential																				
Negative valence high Arousal	-0.72	0.96	3.00	0.75	-0.15	0.09	2.94	0.05	2.24	0.80	6.17	1.32	1.68	0.63	1.88	2.05	3.36	0.39	1.63	0.57
Neutral valence low Arousal	-0.74	0.96	2.95	0.73	-0.01	0.03	2.85	0.03	2.17	0.80	5.96	1.33	2.15	0.63	2.21	2.59	3.48	0.41	1.09	0.24
p	0.86		0.53		0.00		0.00		0.47		0.15		0.00		0.22		0.01		0.00	
Maximally controlled stimulus set																				
Lexical affective content																				
Negative valence high Arousal	-1.60	0.46	3.59	0.40	-0.06	0.10	2.90	0.06	2.06	0.79	6.22	1.48	1.93	0.71	1.58	1.87	3.39	0.35	1.38	0.44
Neutral valence low Arousal	0.11	0.38	2.32	0.28	-0.05	0.10	2.89	0.06	2.06	0.80	6.15	1.45	1.94	0.69	1.76	2.30	3.37	0.39	1.39	0.44
p	0.00		0.00		0.44		0.88		0.95		0.67		0.87		0.46		0.67		0.88	
Sublexical affective potential																				
Negative valence high Arousal	-0.76	0.94	2.98	0.71	-0.12	0.07	2.94	0.04	2.01	0.77	6.28	1.43	1.92	0.70	1.61	1.96	3.37	0.39	1.39	0.45
Neutral valence low Arousal	-0.74	0.97	2.94	0.74	0.02	0.08	2.85	0.05	2.11	0.81	6.08	1.49	1.95	0.71	1.73	2.23	3.39	0.34	1.38	0.42
p	0.89		0.68		0.00		0.00		0.28		0.23		0.75		0.60		0.66		0.72	

Table 2.2: Phoneme (segments) distribution (in DISC Phonetic Encoding Convention; Burnage, 1990) across the conditions of *sublexical affective potential* in both stimuli sets (neg-high = combination of negative valence and high arousal, neut-low = combination of neutral valence and low arousal).

	Maximally controlled set		***Maximally manipulated set***	
Phonemes	**neg-high**	**neut-low**	**neg-high**	**neut-low**
Onsets				
=	16	6	10	1
=v	1	0	3	0
b	6	18	5	25
bl	2	5	1	0
br	6	0	10	0
d	14	14	10	31
dr	5	0	8	0
f	16	7	9	7
fl	0	6	0	2
fr	0	13	0	1
g	11	19	2	30
gl	0	2	0	2
gn	1	1	0	2
gr	5	0	4	0
h	8	5	6	9
k	4	20	9	18
kl	0	9	0	3
kn	2	0	6	0
kr	11	0	17	0
ks	3	0	1	0
l	6	26	1	34
m	12	19	8	33
n	15	8	8	21
N	2	2	2	2
p	8	18	7	9
pr	4	1	5	0
r	33	6	25	5
s	7	0	8	0
S	14	0	13	0
Sl	1	0	3	0
Sp	0	0	2	0

Sr	1	0	4	0
st	2	0	3	0
St	4	11	3	1
Str	2	0	10	0
Sv	0	0	3	0
t	27	17	16	21
tr	4	0	12	0
v	7	16	5	19
x	3	4	4	5
z	6	17	3	31
Nuclei				
&	37	35	30	22
)	3	3	1	3
/	1	0	2	0
@	61	79	71	106
|	6	0	3	0
a	8	21	6	29
B	5	5	3	4
e	11	18	2	18
E	38	18	35	12
i	18	21	10	25
I	43	16	29	17
o	13	13	7	29
O	7	22	15	15
u	2	12	0	10
U	15	13	19	9
W	12	15	13	27
X	13	1	8	1
y	2	3	3	3
Y	5	8	5	6
Codas				
+	1	0	4	0
=	4	0	7	0
b	0	1	0	2
d	2	2	3	0
f	0	5	1	5
ft	1	0	1	1
g	8	0	1	0
k	10	12	15	4

l	6	26	6	35
ln	6	0	1	0
lt	2	4	2	0
lx	0	2	0	0
m	2	14	3	13
n	44	38	41	47
N	3	7	3	11
Nk	1	0	3	0
nt	0	22	0	5
p	12	0	11	0
r	36	29	26	31
r=	2	0	3	0
rk	0	2	0	2
rn	1	4	0	2
rS	2	0	2	0
rt	0	2	0	0
s	19	4	12	2
S	8	0	3	1
st	11	0	8	0
t	12	12	9	14
v	0	0	2	0
x	14	8	13	7
xt	6	0	7	0

2.3.4. EEG recording and (pre-)processing

The EEG was recorded from 61 AgCl-electrodes (Fp1, Fpz, Fp2, AF3, AF4, F5, F3, F1, Fz, F2, F4, F6, FT7, FC3, FC1, FCz, FC2, FC4, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P9, P7, P5, P3, P1, Pz, P2, P4, P6, P8, P10, PO9, PO7, PO3, POz, PO4, PO8, PO10, O1, Oz, O2, Iz, M1, M2) fixed to the scalp via an elastic cap using two 32-channel amplifiers (BrainAmp, Brain Products, Germany). Electrodes were arranged according to the International 10–20 system (American Electroencephalographic Society, 1991; Jasper, 1958) and average impedances were kept below 2 kΩ. The electrooculogram (EOG) was monitored by two electrodes at the outer canthi of the participant's eyes and two electrodes above and below the right eye. EEG and EOG signals were recorded with a sampling rate of 500 Hz, referenced to the right mastoid, but re-referenced offline to linked mastoids. The AFz electrode was used as ground electrode. Later offline filtering included a bandpass filter of 0.1-20 Hz and a notch filter of 50 Hz. Independent component analysis (ICA; Makeig, Bell, Jung, & Sejnowski, 1996; Jung, Makeig, Bell, & Sejnowski, 1998) was carried out to identify and remove eye movement artifacts. The continuous EEG signal was cut into segments of 950 ms total length, consisting of a 150 ms pre-stimulus baseline and an 800 ms post-stimulus interval. After baseline correction, trials containing artifacts were excluded from further analysis using an automatic artifact rejection: differences > 80 µV in intervals of 70 ms or amplitudes > 50 or < -50 µV were considered artifacts. Segments containing correctly answered word trials got averaged per condition, participant and electrode, before grand averages were computed across all participants. To visually compare the ERP signals of different conditions the (sublexically) neutral low-arousing words were always subtracted from the (sublexically) negative high-arousing words.

2.3.5. Data analysis

Behavioral data

Mean correct response latencies and error rates of the word stimuli were submitted to separate ANOVAs – testing whether a potentially given effect generalizes over subjects (F1 analysis) and over items (F2 analysis) – for the factors *lexical affective content* (2) and *sublexical affective potential* (2).

EEG data

Time windows for the expected ERP components of the *lexical affective content* of words were defined based on the literature (see Citron, 2012) and visual inspection of the grand averages: 200 – 300 ms for the EPN, and 400 – 700 ms for the LPC.

For potential effects of the *sublexical affective potential* of the word stimuli, there are no prior studies to base hypotheses on. We thus used an exploratory approach where a time-line analysis with 20 ms time windows (starting from each data point) was carried out. To reduce the chances of false positives potentially arising through consecutive testing, only total time windows of at least 50 ms length – consisting of consecutively significant single time windows revealed by the time-line analysis – were used for further analysis (based on the approach suggested by Guthrie & Buchwald, 1991).

Repeated-measures ANOVAs were conducted with the mean activity [µV] values of the selected time windows using the software IBM SPSS Statistics. The ANOVAs involved the within-subject factors *lexical affective content* (2) or *sublexical affective potential* (2). In order to assess topographical potential distributions of relevant effects over the scalp through an *a priori* designed, hypothesis-independent approach using data from a maximum of electrodes, the ANOVAs further included the topographic factors left-mid-right (3) and anterior-central-posterior (3). For these topographic analyses the scalp electrodes were subdivided into the following 9 clusters of 6 electrodes each: right anterior (FP2, AF4, F4, F6, FC4, FT8), mid anterior (F1, Fz, F2, FC1, FC2, FCz), left anterior (FP1, AF3, F3, F5, FC3, FT7), right central (C4, C6, T8, CP4, CP6, TP8), mid central (C1, Cz, C2, CP1, CPz, CP2), left central (C3, C5, T7, CP3, CP5, TP7), right posterior (P4, P6, P8, PO4, PO8, O2), mid posterior (P1, Pz, P2, POz, Oz, Iz), and left posterior (P3, P5, P7, PO3, PO7, O1).

Furthermore, a region of interest (ROI) for the early posterior negativity (EPN) was defined using a cluster of the 11 most posterior electrodes (PO9, PO7, PO3, POz, PO10, PO8, PO4, O1, Oz, O2, Iz), based on earlier topographic data regarding EPN effects in our research group (Conrad et al, 2011; Recio et al., 2014). If the visual topography patterns suggested so, data of EPN ROI were submitted to paired t-tests between the affective conditions. The combination of these two approaches toward topographic analysis, one unbiased and one guided by hypotheses, should offer a most comprehensive insight in this novel research topic. All topographic clusters and the ROI are displayed in Figure 2.1.

Greenhouse-Geisser corrected p-values (Greenhouse & Geisser, 1959) are reported for all ANOVA results. Significant interactions with topographic factors were followed up by paired t-tests within the respective topographic clusters. The p values of multiple post-hoc t-tests got Bonferroni-Holm adjusted (Holm, 1979) and are marked as p_{adj}. As measure of effect size η_p^2 is reported for the ANOVAs (Keppel, 1991; Tabachnick & Fidell, 2001) and Pearson's *r* for the t-tests (Clark-Carter, 2003; Field, 2009).

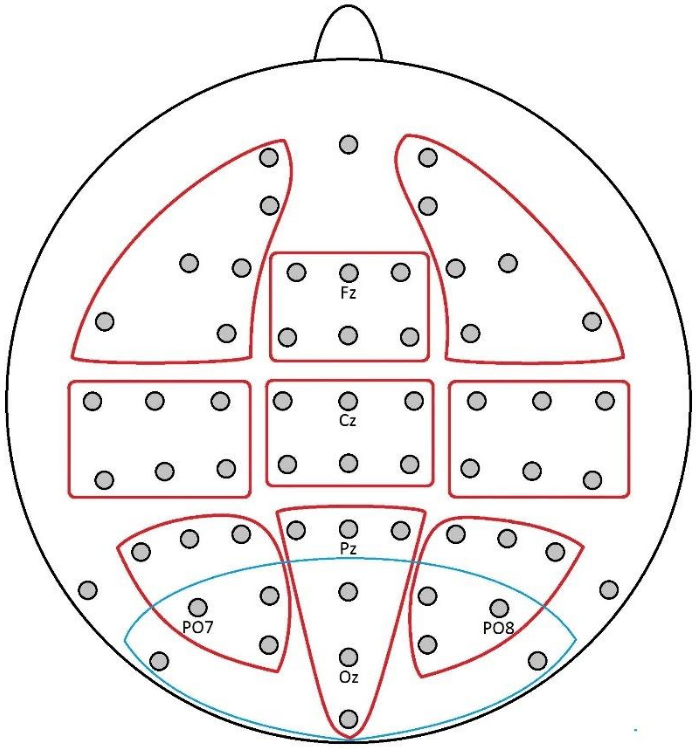

Figure 2.1: Electrode positions of the applied 10–20 system with marked topographic clusters (ROIs) as used in the analyses: red = exploratory topographic clusters, blue = EPN ROI.

2.4. Results

2.4.1. Behavioral results

Maximally manipulated stimulus set

The analysis of reaction times (RTs) for the *sublexical affective potential* yielded no significant differences between the RTs to sublexically negative high-arousing words and to sublexically neutral low-arousing words (*F1(1,40)* = 3.66, p = 0.06, η_p^2 = 0.08; *F2(1,306)* = 1.45, *p* = 0.23, η_p^2 = 0.01). For the *lexical affective content,* we found a significant F1 effect (with slower responses to negative high-arousing words than to neutral low-arousing words), but the F2 analysis remained non-significant (*F1(1,40)* = 6.35, *p* = 0.02, η_p^2 = 0.14; *F2(1,306)* = 1.01, p = 0.32, η_p^2 = 0.003). Regarding error rates, again we do not find a significant effect for *sublexical* *affective potential* (*F1(1,40)* = 3.43, *p* = 0.07, η_p^2 = 0.08; *F2(1,306)* = 1.28, *p* = 0.26, η_p^2 = 0.004). There is also no effect for error rates regarding the *lexical affective content* (*F1(1,40)* = 0.01, *p* = 0.91, η_p^2 = 0.00; *F2(1,306)* = 0.00, *p* = 0.99, η_p^2 = 0.00).

<u>Maximally controlled stimulus set</u>

Although the F1 analysis of RTs renders a significant effect for the *sublexical affective potential* with faster responses to the sublexically neutral low-arousing words, the F2 analysis is non-significant (*F1(1,40)* = 5.56, *p* = 0.02, η_p^2 = 0.12; *F2(1,305)* =1.29, *p* = 0.26, η_p^2 = 0.004). Also for the *lexical affective content*, there is no significant effect in RTs' analysis (*F1(1,40)* = 2.96, *p* = 0.09, η_p^2 = 0.07; *F2(1,305)* = 1.93, *p* = 0.17, η_p^2 = 0.01). Looking at the error rates, a significant F1 difference between lexical affective conditions (more errors on negative high-arousing words compared to neutral low-arousing ones) is not accompanied by a significant F2 analysis (*F1(1,40)* = 5.52, *p* = 0.02, η_p^2 = 0.12; *F2(1,305)* = 1.03, *p* = 0.31, η_p^2 = 0.003). Further, there is no significant effect for the *sublexical affective potential* (*F1(1,40)* = 1.67, *p* = 0.2, η_p^2 = 0.04; *F2(1,305)* = 0.27, *p* = 0.6, η_p^2 = 0.001).

2.4.2. ERP results

<u>Maximally manipulated stimulus set</u>

Lexical affective content

An early effect of the *lexical affective content* was found in the time window of the early posterior negativity (EPN) between 200 and 300 ms in interaction with the topographic factor left-mid-right (*F(2,68)* = 3.7, *p* = 0.03, η_p^2 = 0.1). T-tests within each of the three laterality clusters only showed a trend toward a difference between neutral low-arousing and negative high-arousing words in the left cluster (*t(34)* = -2.17, p_{adj} = 0.12, *r* = 0.35) with a larger negativity for negative high-arousing words. Yet, the topographic map (Figure 2.2) reveals that this negativity is of a shape that cannot be caught well by the cluster formation of the exploratory topographic analysis. Rather, most distinct negativity shows in a left posterior area, as would be hypothesized for the expected EPN. Results of EPN ROI analysis were: $t(34) = -1.87$, $p = 0.07$, $r = 0.31$. Although, here again, we can only find a trend toward significance, in both analyses the postulated effect is of a medium size, which cannot be neglected (see discussion for why the effect might not be as strong as in previous literature).

A late positive complex (LPC) can be found between 400 and 700 ms as a significant main effect for the *lexical affective content* ($F(1,34) = 8$, $p = 0.01$, $\eta_p^2 = 0.19$) with more positive values for the negative high-arousing words compared to the neutral low-arousing words. Furthermore, we find a significant interaction of this lexical effect with the topographic cluster division anterior-central-posterior ($F(2,68) = 7.88$, $p = 0.004$, $\eta_p^2 = 0.19$): t-tests within each of these clusters revealed significant differences between the two *lexical affective content* conditions in the anterior ($t(34) = 4.12$, $p_{adj} < 0.003$, $r = 0.58$) and the central cluster ($t(34) = 2.74$, $p_{adj} = 0.02$, $r = 0.43$). This fronto-central positivity is also reflected in the topographic map as shown in Figure 2.2.

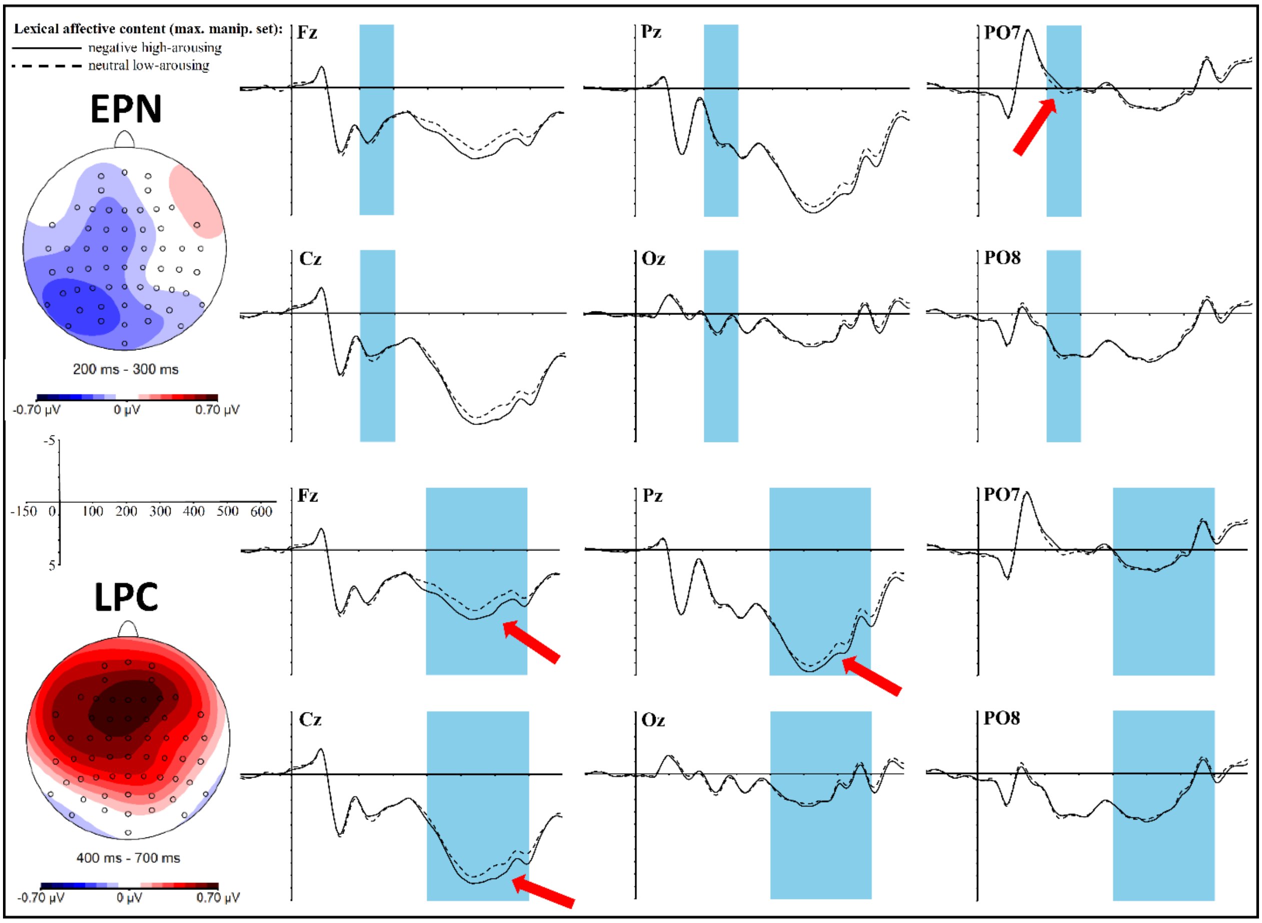

Figure 2.2: ERP effects of the *lexical affective content* (top: EPN, bottom: LPC) in the *maximally manipulated* stimulus set at selected electrode sites. For the topographic maps neutral low-arousing words were subtracted from negative high-arousing words.

Sublexical affective potential

Visual inspection already suggested a robust and long-lasting negativity between 250 and 650 ms that proved to be a significant main effect of the *sublexical affective potential* ($F(1,34)$ = 7.77, p = 0.01, η_p^2 = 0.19) with sublexically negative high-arousing words eliciting a larger negativity over this whole time interval than sublexically neutral low-arousing words. Also the 3-fold interaction of *sublexical affective potential* x topographic factor anterior-central-posterior x topographic factor left-mid-right turns out significant ($F(4,136)$ = 4.76, p = 0.003, η_p^2 = 0.12). After correction for multiple testing, one of the t-tests in each of the nine topographic clusters turned out significant (in the right central cluster with $t(34)$ = -3.06, p_{adj}

= 0.036, *r* = 0,46), one marginally significant (in the right posterior cluster with *t(34)* = -2.86, p_{adj} = 0.056, *r* = 0,44), and four more neighboring clusters still showed trends (left anterior cluster with *t(34)* = -2.42 , p_{adj} = 0.11, *r* = 0.38, mid anterior cluster with *t(34)* = -2.57, p_{adj} = 0.09, *r* = 0.4, right anterior cluster with *t(34)* = -2.32, p_{adj} = 0.11, *r* = 0.37, and mid central cluster with *t(34)* = -2.74, p_{adj} = 0.07, *r* = 0.43), always with a larger negativity for sublexically negative high-arousing words. Figure 2.3 displays the topography of this right-central negativity and the ERP graphs at selected electrodes.

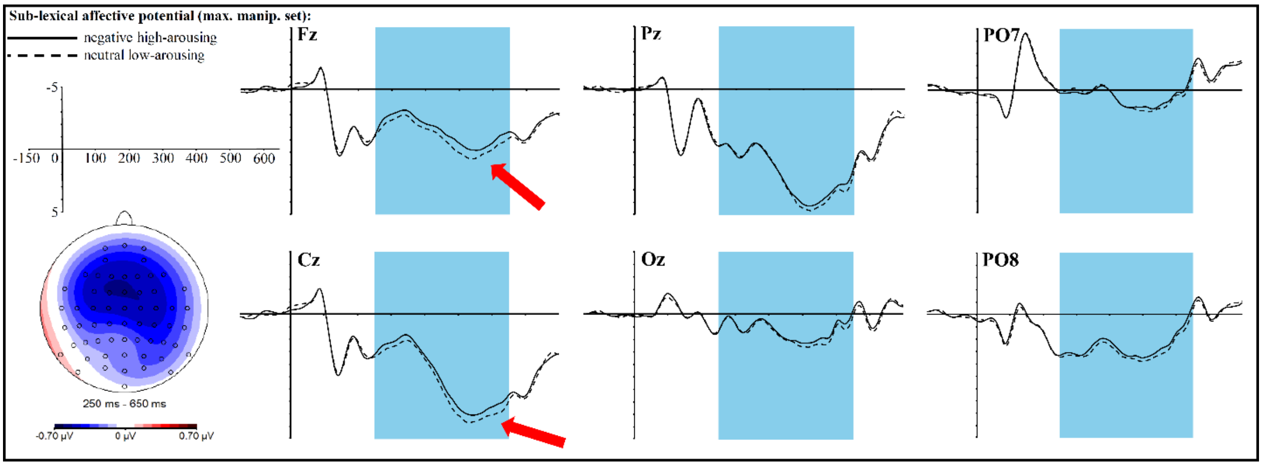

Figure 2.3: ERP effect of the *sublexical affective potential* in the *maximally controlled* stimulus set at selected electrode sites. For the topographic map sublexically neutral low-arousing words were subtracted from sublexically negative high-arousing words.

Maximally controlled stimulus set

Lexical affective content

In the EPN time window between 200 and 300 ms an early effect of *lexical affective content* exists in interaction with the topographic factor anterior-central-posterior ($F(2,68)$ = 8.23, p = 0.003, η_p^2 = 0.2) as well as in a 3-fold interaction also including the left-mid-right factor ($F(4,136)$ = 3.68, p = 0.01, η_p^2 = 0.1). T-tests within the respective topographic clusters reveal a significant difference between neutral low-arousing and negative high-arousing words in the whole posterior cluster ($t(34)$ = -2.71, p_{adj} = 0.03, r = 0.42) as well as trends in the single posterior clusters: left posterior ($t(34)$ = -2.88, p_{adj} = 0.06, r = 0.44), mid posterior ($t(34)$ = -2.46, p_{adj} = 0.13, r = 0.39), and right posterior ($t(34)$ = -2.5, p_{adj} = 0. 14, r = 0.39), always showing a higher negativity for the lexically negative and high-arousing words. A t-test within the EPN ROI shows a significant difference between the two lexical affective conditions ($t(34)$ = -3.17, p = 0.003, r = 0.48) going in the same direction. The topography of this effect does well reflect the EPN pattern as expected. It is shown together with the EEG graphs at selected electrodes in Figure 2.4 (upper part).

A late positive complex (LPC) shows between 400 and 700 ms as a significant main effect for *lexical affective content* ($F(1,34)$ = 6.16, p = 0.02, η_p^2 = 0.15) with more positive values for the negative high-arousing words compared to neutral low-arousing words. Also the interaction of *lexical affective content* with the topographic division anterior-central-posterior is significant ($F(2,68)$ = 8.04, p = 0.01, η_p^2 = 0.19), with a significant t-test result in the anterior cluster ($t(34)$ = 3.71, p_{adj} = 0.003, r = 0.54) as well as a trend showing within the central cluster ($t(34)$ = 2.22, p_{adj} = 0.07, r = 0.36). This fronto-central positivity with negative high-arousing words displaying a higher positivity than neutral low-arousing words is displayed in the lower topographic map of Figure 2.4.

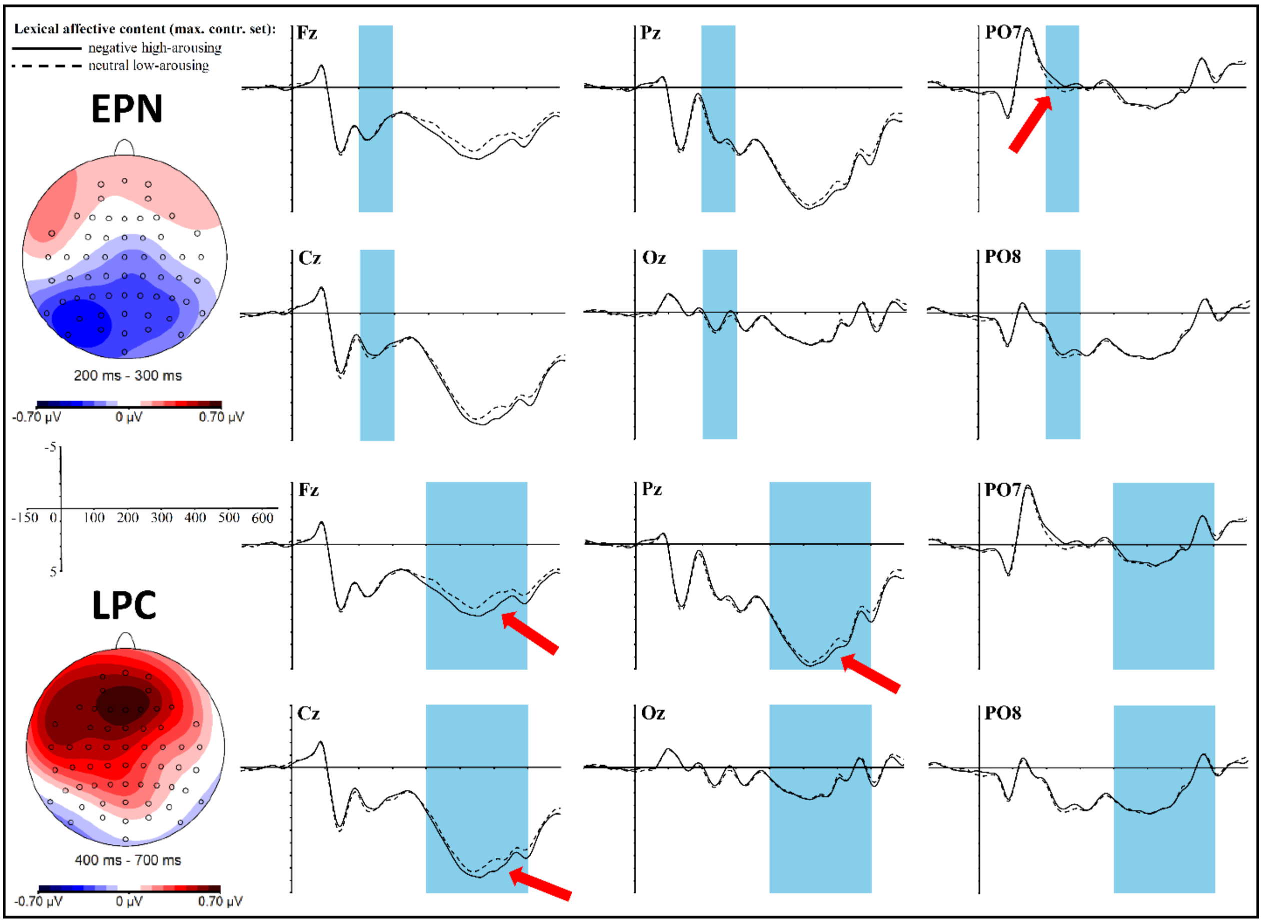

Figure 2.4: ERP effects of the *lexical affective content* (top: EPN, bottom: LPC) in the *maximally controlled* stimulus set at selected electrode sites. For the topographic maps neutral low-arousing words were subtracted from negative high-arousing words.

Sublexical affective potential

The exploratory time-line analysis revealed contiguous significant time windows between 226 and 276 ms for the interaction of the *sublexical affective potential* with the topographic factors anterior-central-posterior. Thus, we analyzed this time window as a whole, which yields a significant interaction of *sublexical affective potential* with the anterior-central-posterior clustering ($F(2,68) = 6.67$, $p = 0.01$, $\eta_p^2 = 0.16$). Solving this interaction only leads to a rough trend within the whole posterior cluster ($t(34) = -1.9$, $p_{adj} = 0.2$, $r = 0.31$) with a more negative amplitude for the sublexically negative high-arousing words, yet of medium effect size. Visual inspection of the topographic map (see Figure 2.5) reveals that this posterior negativity looks quite similar to the lexical EPN. Hence, we also tested for significance within the EPN ROI: the t-test shows a significant difference between sublexically negative high-arousing words and sublexically neutral low-arousing words ($t(34)$ = -2.68, p = 0.01, r = 0.42). The topography and ERP graphs at selected electrodes are displayed in Figure 2.5.

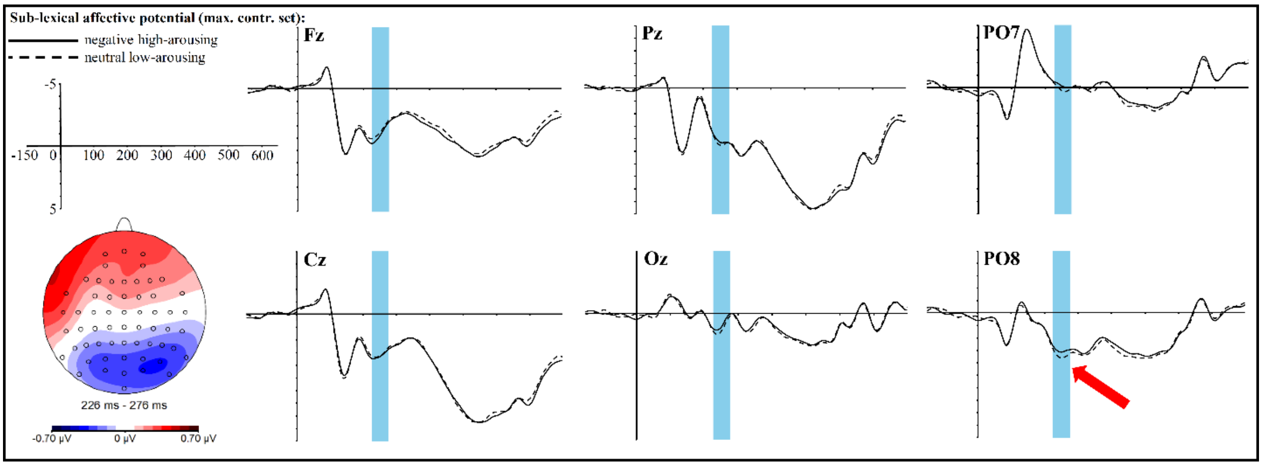

Figure 2.5: ERP effect of the *sublexical affective potential* in the *maximally controlled* stimulus set at selected electrode sites. For the topographic map sublexically neutral low-arousing words were subtracted from sublexically negative high-arousing words.

2.5. Discussion

The present study investigates whether systematic sound-to-meaning correspondences that we had detected in the German language influence the neural processes of language perception – assessed by EEG recordings during the most standard task used in psycholinguistic research: visual lexical decision.

There is a longstanding debate in theoretical linguistics oscillating between the well-known axiom of arbitrary relations between the signifier and the signified on the one hand, and numerous studies on phenomena of *sound symbolism* and *phonological iconicity* on the other hand (for reviews see Dingemanse et al., 2015; Perniss et al., 2010; Schmidtke et al., 2014).

Here, we focused on sound-to-meaning correspondences assumed to represent *phonological iconicity* with regard to a sublexical encoding of affect: Certain phonological segments – syllabic onsets, nuclei or codas – were found to occur particularly often in words of negative and/or high-arousing semantic meaning. As these findings proved statistically reliable across a large-scale database of over 6,000 German words, we assume they might represent a certain degree of iconic organization of language rather than merely idiosyncratic “Gestalt” features of single words (Conrad et al., in preparation).

Based on this assumption, we calculated:

- first, *sublexical affective values* (SAVs) for single phonological segments as a function (average) of the affective values of all words they occur in
- second, an estimate of the *sublexical affective potential* of whole words as a function (average) of the SAVs of all phonological segments forming this word

We then tested – using EEG measurements – whether apparent sound-to-meaning correspondences represent anything more than a hard-to-interpret “intriguing finding” arising from statistical analyses of large-scale lexical databases. We used these measures of *sublexical* *affective potential* – derived directly from the large-scale database – as an experimental factor distinguishing between words that "should" sound – according to these sound-to-meaning correspondences in the database – highly arousing and negative vs. words with rather neutral phonological affective qualities.

Our data suggest that these sound-to-meaning correspondences or statistical regularities of German with regard to sublexical phonology and affective content of words are rooted in phenomena that crucially influence basic online reading processes: Regardless of the actual *lexical affective content* of stimuli, words that were composed of phonological segments typically occurring in words of negative high-arousing meaning caused a very robust and long-lasting negativity in the ERP signal when participants simply tried to lexically access these words – compared to words consisting of affectively "neutral" phonological segments. As the most important finding of our study, this effect is strong evidence for the psychological relevance of affective sound-to-meaning correspondences in the German language at the level of sublexical units.

However, it is more difficult to attribute this effect to a specific type of processing. This is because those phonological segments typically occurring in words with threatening affective content (high arousal and negative valence) tend to be of formal salience as well: their frequency of occurrence is considerably low and/or they are phonologically rather complex, i.e., combining several consonants in syllabic onsets or codas. Note that this makes perfectly sense from an evolutionary perspective: If language would choose a specific phonological segment as a sublexical sign of threatening affective content, it should use this sign not too often to avoid inflation or decay of the alerting sign character. Further, the alerting character of the sign would clearly benefit from salient perceptive characteristics such as, for instance, complex phonological structure requiring increasing effort for articulation processes for several consonants combined in one syllabic onset or coda. In a strict sense, this confound with structural saliency makes it difficult to interpret our robust effect for the manipulation of *sublexical affective potential* in the maximally manipulated set as anything else than an effect of general sublexical encoding processes during silent reading – arising from the complexity and/or low frequency of the sublexical units (see, e.g., Nuerk et al., 2000, for phonological/subsyllabic component frequency; Goslin, Grainger, & Holcomb, 2006, for syllabic structure; Barber, Vergara, Carreiras, 2004 and Hutzler et al., 2004, for syllable frequency; Hauk et al., 2006, and Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006, for bigram frequency). According to a two-fold representation of phonological units comprising an auditory as well as motor template (Hickok, 2012), also articulatory activations – especially with regard to the complex phonological clusters – are possibly involved. Neuroimaging studies, indeed, show motor circuits responsible for articulatory movements to be activated in response to visually presented word stimuli (Burton, LoCasto, Krebs-Noble, & Gullapalli, 2005; Hagoort, Indefrey, Brown, Herzog, Steinmetz, & Seitz, 1999).

To control for the influence of these potential intervenient factors we had prepared and presented an additional, *maximally controlled* stimulus set involving the same manipulations but controlling for the confounds of *sublexical affective potential* with formal complexity and frequency. In this set – though massively deteriorating the natural variance of the manipulated variable and respectively the strength of the manipulation – the *sublexical affective potential* of stimulus words still produced a small but significant effect in the ERP signal of non-neglectable medium effect size. More interestingly, the distribution of this effect across the scalp and the moment it appears during the reading process closely resemble what is typically reported – and also present in our data – for manipulations of affective content at the lexical level: an increased negativity at posterior electrode sites arising at around 200 ms after stimulus onset (EPN). Yet, although this topographic and temporal coincidence with the lexically driven EPN appears somewhat striking, this novel finding – obtained through explorative time-line analysis – certainly calls for corroboration in future studies that should also explore which brain regions may be involved in these processes.

Note also that both EPN and LPC effects of *lexical affective content* manipulations appear somewhat diminished in our data when compared to previous experimental reports focusing on general emotion effects during visual word recognition (e.g., Conrad et al., 2011; Recio et al., 2014; just to quote two from the same lab). In our study, these manipulations of *lexical affective content* only served as control measures allowing us to relate both the moment when effects of the *sublexical affective potential* would arise and how their morphology would look like in comparison to more classical effects of *lexical affective content* within one and the same experimental context. Such simultaneous manipulations of different factors that have to be kept independent from each other clearly have the consequence that the strength for each manipulation gets attenuated as compared to when manipulated alone. In consequence, resulting empirical effects may have got attenuated too.

Further, our specific manipulations of affective content combining negative valence with high arousal may not have favored lexical affective effects to show up in most robust ways, as these effects have been shown to be stronger for positive as compared to negative valence (Recio et al., 2014). We assume that this restriction to negative affective content may be responsible for the lack of effects in our behavioral data. Whereas a processing advantage for positive stimuli is consistently being reported in the literature, the picture is more heterogeneous for negative contents: One the one hand, the automatic evaluation hypothesis predicts faster processing of positive or negative words compared to neutral words, supported by several lexical decision studies (Kousta, Vinson, & Vigliocco, 2009; Hofmann, Kuchinke, Tamm, Võ, &

Jacobs, 2009). However, also opposite findings, where reaction times for negative words are not different from neutral words (Briesemeister, Kuchinke, & Jacobs, 2012; Recio et al., 2014) or even longer compared to neutral or positive words (Carretié et al., 2008; Estes & Adelman, 2008) have been reported. Such findings are explained by the automatic vigilance hypothesis (Pratto & John, 1991), according to which fast and automatic evaluation of especially negative stimuli directs attention away from the actual task, e.g., lexical decision, causing prolonged response times and higher error rates due to a deeper processing of the negative word content or even because of a tendency to withdraw from negative stimuli.

The same may, of course, explain the absence of *sublexical affective potential* effects in our behavioral data. But note also that even though our ERP data show that this *sublexical affective potential* together with its formal salience do play a role for automatic reading processes, we do not see why this should necessarily bias – speed or delay – the tendency to decide that a given stimulus is a word or not. We do clearly not posit that these phenomena should – besides potentially attracting attention at some point of the reading process – trigger a fundamental general cognitive bias, and sublexical and lexical affective content are, further, unrelated in our stimuli. Taken together, the contrast between significant ERP effects and the lack of such effects at the behavioral level in our study may best serve as a good example of how RT effects only represent the final point of a decision process, whereas ERPs may better reveal fine-grained and potentially contradicting processes that precede a final response – concerning the latency of which their contradictory effects may have cancelled each other out.

Whereas the topographical potential distribution of our early ERP effects aligns well with homogenous reports on classical EPN effects, the topography of the LPC effects deserves a bit more discussion, as in some studies, the LPC has been found to be more posterior (Herbert et al., 2008; Kissler et al., 2009). Yet in general, the amplitude, latency, and topographic dispersion of the LPC have been found to be task-dependent (Fischler & Bradley, 2006; Schacht & Sommer, 2009). Whereas a word counting task yielded a posterior LPC (Kissler et al., 2009), it showed a bit further central when subjects just had to passively listen to words (Herbert et al., 2008). With lexical decision tasks, the LPC usually is found in a fronto-central position (Conrad et al., 2011; Recio et al., 2014; Schacht & Sommer, 2009), and even further frontal when asking the participants to rate the words on affective dimensions (Dillon et al., 2006) – all latter reports being compatible to our findings *for lexical affective content.* On the other hand, we found no such typical LPC-like component for the contrast of *sublexical affective potential.* The reason therefore is probably that this component generally appears linked to higher-cognitive elaborative processing, whereas our sublexical manipulation taps into more basic processing stages.

What our data – obtained with highly controlled experimental manipulations and providing an excellent signal-to-noise ratio involving more than 150 stimuli per condition and 35 participants – suggest is that already specific phonological segments can trigger at the sublexical level what is classically observed and reported as (lexical) emotion effects during the reading process: an early posterior negativity (EPN) at around 200 ms after stimulus onset. In combination with the finding of the long-lasting negativity in the less controlled stimulus set, our data thus represent novel neurophysiological evidence for *phonological iconicity* as a principle systematically influencing the organization of the vocabulary AND the online processing of a language like German. The reading system appears to be sensitive to the transport of affective information via sublexical signs of affective meaning. The EPN is usually interpreted as evidence for an early automatic attention shift towards emotionally relevant stimuli. So far, this emotional relevance was determined by the *lexical affective meaning* (or *content*) of word stimuli in a number of previous ERP studies (see Citron, 2012, for a review). In the case of our study, the same effect might already be elicited by sublexical phonological segments alone. One possibility of how this effect might arise can be seen in statistical learning: the sound-to-meaning correspondences our experimental manipulations are based upon could represent such well learned regularities, that presentation of certain phonological segments is sufficient to elicit the same emotional attention processes as whole word forms representing emotion-laden concepts. Phonological segments, in that case, would have acquired symbolic affective values via associative links across the lexicon. However, an alternative explanation would refer more directly to an internal relation between acoustic or phonological properties of specific phonological segments and affective meaning at the conceptual level: As we outline in Conrad et al. (in preparation), phonemes occurring more frequently in words of high arousal (and negative valence) tend to possess phonemic features – e.g., sibilants or unvoiced stops – that go along with an increasing arousal at the level of acoustic impressions, according to the *distinct features theory* by Jakobson, Fant, & Halle (1952). Therefore, it might have been the increasing arousal at the level of phonemic features typically occurring in words of high arousal and negative valence that has triggered the EPN in our data. This interpretation aligns with the general assumption of *phonological iconicity* to represent an internal relation between the conceptual and the sublexical level: Certain phonological segments – iconic for high arousal – could provoke the same pattern of electrophysiological activity – reflected by the EPN – as emotion-laden words, because the phonemic features of these segments are of similar affective salience. The fact that respective ERP effects of the *sublexical affective potential* appear as clearly diminished in the *maximally controlled* stimulus set compared to the *maximally manipulated* stimulus set is probably mainly due to the constraint of controlling for the major co-variation of *sublexical affective* *potential* with formal salience. But it has to be kept in mind that already this empirical confound per se sheds light on the *phonological iconicity* effects, as the German language apparently made use of phonological segments that leave most impressive "footmarks" in neural correlates of the language processing – as evident from the robust effects of our *maximal manipulation* of *sublexical affective potential* – to encode threatening affective meaning. Taken together, this pattern of findings strongly points toward an internal relation between sublexical signs and affective meaning at the conceptual level and is in clear opposition to the arbitrariness axiom of linguistic theory concerning the relation between a signifier and the signified.

Finally, note that also processes of production or articulation preparation may have influenced our ERP data for *sublexical affective potential* – even though the task was visual lexical decision. Phonological iconicity may well be rooted in articulation processes determining an internal relation between the conceptual and the sublexical level. This appears even more plausible considering the relation between SAVs and structural complexity of consonant syllabic segments (increasing complexity of negative/high arousal segments). As the motor theory of speech perception (Liberman & Mattingly, 1985) states, perception and articulation aspects are highly entangled during neural processing of language (D'Ausilio et al, 2009; Pulvermüller et al., 2006), and our design does not allow to clearly distinguish between either perception or articulation preparation as potential sources of effects – which, in turn, appears a most fruitful field for future research.

Language comparisons could provide interesting insights concerning potentially "universal" vs. language-dependent features of *phonological iconicity*. In particular, as our data involve "phonological" iconicity effects after visual presentation using orthographic codes from a shallow orthography, it might be interesting to see whether similar effects could be obtained in languages with less transparent orthographies, e.g., using English words. If effects persisted for both consistent and inconsistent grapheme-to-phoneme mappings, this would suggest that iconicity with regard to affective content might have already generalized from the phonological to the orthographic domain.

2.6. Acknowledgements

This research was funded by two grants (410 "Sound physiognomy in language organization, processing and production" und 201 "Bilingualism and affectivity in reading") to Markus Conrad, Sonja Kotz, and Arthur Jacobs from the German Research Foundation (DFG) via the Cluster of Excellence “Languages of Emotion” at the Freie Universität Berlin. We thank Luisa Egle, Maren Luitjens, Mariam Murusidze, Kathrin Schreiter, Susanne Löhne, Johannes Ecker, Hauke Blume, Chun-Ting Hsu, and Gesche Schauenburg for their help in conducting the EEG experiments. This study got approved by the ethics committee of the Freie Universität Berlin and was conducted in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.7. References

Abramson, M., and Goldinger, S.D. (1997). What the reader’s eye tells the mind’s ear: Silent reading activates inner speech. *Perception & Psychophysics, 59,* 1059-1068.

Allott, R. (1995). Sound symbolism. In U.L. Figge (Ed.), *Language in the Würm Glaciation* (pp. 15-38). Bochum: Brockmeyer.

American Electroencephalographic Society (1991). Guidelines for standard electrode position nomenclature. *J. Clin. Neurophysiol.*, *8*(2), 200-202.

Aryani, A., Jacobs, A.M., and Conrad, M. (2013). Extracting salient sublexical units from written texts: “Emophon,” a corpus-based approach to phonological iconicity. *Frontiers in psychology, 4,* 654. doi:10.3389/fpsyg.2013.00654

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A.M., and Conrad, M. (2015). Measuring the Basic Affective Tone of Poems via Phonological Saliency and Iconicity. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(2), 191-204. doi:10.1037/aca0000033

Barber, H., Vergara, M., and Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. *Neuroreport*, *15*(3), 545-548.

Belin, P., Bestelmeyer, P. E., Latinus, M., and Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, *102*(4), 711-725.

Bergen, B.K. (2004). "The psychological reality of phonaesthemes." *Language,* 80(2): 290-311.

Bloomfield, L. (1933). *Language*. New York: Holt.

Bradley, M.M., Greenwald, M.K., Petry, M.C., and Lang, P.J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 379-390.

Brand, M., Giroux, I., Puijalon, C., and Rey, A. (2007). Syllable onsets are perceptual reading units. *Memory & Cognition*, *35*(5), 966-973.

Braun, M., Hutzler, F., Ziegler, J.C., Dambacher, M., and Jacobs, A.M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mappings, 30*(7)*,* 1977-1989.

Briesemeister, B.B., Kuchinke, L., and Jacobs, A.M. (2012). Emotional valence – a bipolar continuum or two independent dimensions? *SAGE Open*, *2*(4). doi: 10.1177/2158244012466558

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A.M., Bölte, J., and Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Exp. Psychol.*, *58*, 412-424. doi: 10.1027/1618-3169/a000123

Burnage, G. (1990). *CELEX: A guide for users.* Nijmegen, CELEX Centre for Lexical Information. Appendix 2: "Computer Phonetic Character Codes"

Burton, M.W., LoCasto, P.C., Krebs-Noble, D., and Gullapalli, R.P. (2005). A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *NeuroImage*, *26*(3), 647-661.

Carreiras, M., Álvarez, C.J., and de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory & Language*, *32,* 766-780.

Carretié, L., Hinojosa, J.A., Albert, J., López-Martín, S., de la Gándara, B., Igoa, J.M., and Sotillo, M. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology*, *45*, 188-196.

Citron, F.M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*(3), 211-226.

Clark-Carter, D. (2003). Effect size - The missing piece in the jigsaw. *Psychologist*, *16*(12), 636-638.

Coltheart, M., Davelaar, E., Jonasson, J.T., and Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). New York: Academic Press.

Conrad, M., and Jacobs, A.M. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language & Cognitive Processes*, *19*, 369-390.

Conrad, M., Grainger, J., and Jacobs, A.M. (2007). Phonology as the source of syllable frequency effects in visual word recognition: Evidence from French. *Memory & Cognition*, *35*(5), 974-983.

Conrad, M., Carreiras, M., Tamm, S., and Jacobs, A.M. (2009). Syllables and bigrams: orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 461.

Conrad, M., Recio, G., and Jacobs, A.M. (2011). The time course of emotion effects in first and second language processing: a cross cultural ERP study with German-Spanish bilinguals. *Frontiers in Psychology*, *2*, 351.

Conrad, M., Aryani, A., Schmidtke, D.S., and Jacobs, A.M. (in preparation). The sound of dictators: Encoding of affective meaning via phonological iconicity.

Darwin, C. (1871). *The Descent of Man.* London: John Murray.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, *19*(5), 381-385.

Davis, R. (1961). The fitness of names to drawings: a cross-cultural study in Tanganyika. *British Journal of Psychology*, *52*, 259-268.

Davis, S. (1982). On the internal organization of syllable constituents. In T. Larson (ed.), *Coyote Papers: Working Papers in Linguistics from A-Z*, Department of Linguistics, University of Arizona, *3*, 3-22.

Dillon, D.G., Cooper, J.J., Grent, T., Woldorff, M.G., and LaBar, K.S. (2006). Dissociation of event-related potentials indexing arousal and semantic cohesion during emotional word encoding. *Brain and cognition*, *62*(1), 43-57.

Dingemanse, M., Blasi, D.E., Lupyan, G., Christiansen, M.H., and Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in cognitive sciences*, *19*(10), 603-615.

Estes, Z., and Adelman, J.S. (2008). Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006). *Emotion*, *8*(4), 441-444.

Field, A.P. (2009). *Discovering statistics using SPSS*. London, England: SAGE.

Firth, J. (1930). *Speech*. London, England: Oxford University Press.

Fischler, I., and Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in Brain Research*, *156*, 185-203.

Folkins, C., and Lenrow, P.B. (1966). An investigation of the expressive values of graphemes. *The Psychological Record*, *16*, 193-200.

Goslin, J., Grainger, J., and Holcomb, P.J. (2006). Syllable frequency effects in French visual word recognition: An ERP study. *Brain Research*, *1115*(1), 121-134.

Graf, R., Nagler, M., and Jacobs, A.M. (2005). Faktorenanalyse von 57 Variablen der visuellen Worterkennung [Factor analysis of 57 variables in visual word recognition]. *Zeitschrift für Psychologie, 213,* 205-218.

Greenhouse, S.W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112.

Guthrie, D., and Buchwald, J.S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*(2), 240-244.

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., and Seitz, R.J. (1999). The neural circuitry involved in the reading of German words and pseudowords: a PET study. *Journal of Cognitive Neuroscience*, *11*(4), 383-398.

Hall, T.A. (1992). *Syllable structure and syllable-related processes in German*, Tübingen: Niemeyer.

Hauk, O., Davis, M.H., Ford, M., Pulvermüller, F., and Marslen-Wilson, W.D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, *30*(4), 1383-1400.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermüller, F., and Rogers, T. (2006). [Q:] When would you prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Cognitive Neuroscience*, Journal of, *18*(5), 818-832.

Heise, D.R. (1966). Sound-meaning correlation among 1,000 English words. *Language and Speech, 9*(1), 14-27.

Hellbernd, N., and Sammler, D. (2016). Prosody conveys speaker's intentions: Acoustic cues for speech act perception. *Journal of Memory and Language*, *88*, 70-86.

Herbert, C., Junghöfer, M., and Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology, 45*, 487-498.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135-145.

Hofmann, M.J., Kuchinke, L., Tamm, S., Võ, M. L.-H, and Jacobs, A.M. (2009). Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective, & Behavioral Neurscience, 9*(4)*,* 389-397.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Statist.*, 6, 65-70.

Hutzler, F., Bergmann, J., Conrad, M., Kronbichler, M., Stenneken, P., and Jacobs, A.M. (2004). Inhibitory effects of first syllable-frequency in lexical decision: An event-related potential study. *Neuroscience Letters, 372*(3), 179-184.

Jackendoff, R. (2002). *Foundations of language.* Oxford: Oxford University Press.

Jacobs, A.M., Grainger, J., Rey, A., and Ziegler, J.C. (1998). MROM-p: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In J. Grainger & A.M. Jacobs (Eds.), *Localist connectionist approaches to human cognition*, 147-188, Hillsdale, NJ: Lawrence Erlbaum Associates.

Jakobson, R. (1937). *Six lectures on sound and meaning*. Cambridge, Mass.: MIT Press

Jakobson, R., Fant, G., and Halle, M. (1952). *Preliminaries to speech analysis: The distinctive features and their correlates. Acoustics Laboratory, 13*, MIT, Technical Report.

Jasper, H. (1958). The ten–twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.*, *10*, 371-375.

Jung, T-P., Makeig, S., Bell, A.J., and Sejnowski, T. J. (1998). Independent component analysis of electroencephalographic and event-related potential data, In: Poon & Brugge (Eds.), *Auditory Processing and Neural Modeling* (pp. 160-188), Plenum Press.

Junghöfer, M., Bradley, M.M., Elbert, T.R., and Lang, P.J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology, 38,* 175-178.

Keppel, G. (1991). *Design and analysis: A researcher's handbook (3rd ed.)*. Englewood Cliffs, NJ: Prentice-Hall.

Keuper, K., Zwanzger, P., Nordt, M., Eden, A., Laeger, I., Zwitserlood, P., ... and Dobel, C. (2014). How 'love' and 'hate' differ from 'sleep': Using combined electro/magnetoencephalographic data to reveal the sources of early cortical responses to emotional words. *Human brain mapping, 35*(3), 875-888.

Kissler, J., Herbert, C., Peyk, P., and Junghöfer, M. (2007). Buzzwords: Early cortical responses to emotional words during reading. *Psychological Science, 18*, 475-482.

Kissler, J., Herbert, C., Winkler, I., and Junghöfer, M. (2009). Emotion and attention in visual word processing – An ERP study. *Biological psychology, 80*(1)*,* 75-83.

Kita, S. (2008). World-view of protolanguage speakers as inferred from semantics of sound symbolic words: a case of Japanese mimetics. In: N. Masataka (Ed.), *The origins of language*, pp. 25–38, Tokyo: Springer.

Köhler, W. (1929). *Gestalt Psychology*. New York, NY: Liveright.

Kotz, S.A., and Paulmann, S. (2011). Emotion, language, and the brain. *Language and Linguistics Compass*, *5*(3), 108-125.

Kousta, S.T., Vinson, D.P., and Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473-481.

Liberman, A.M., and Mattingly, I.G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36.

Lockwood, G., Dingemanse, M., and Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. [Epub ahead of print]

Makeig, S., Bell, A.J., Jung, T-P., and Sejnowski, T.J. (1996). Independent component analysis of Electroencephalographic data. *Advances in Neural Information Processing Systems*, *8*, 145-151.

Maurer, D., Pathman, T., and Mondloch, C.J. (2006). The shape of boubas: sound-shape correspondences in toddlers and adults. *Developmental Science*, *9*, 316-322.

Morton, E.W. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *American Naturalist*, *111*, 855-869.

Newman, S.S. (1933). Further Experiments in Phonetic Symbolism. *American Journal of Psychology, 45,* 53-75.

Nuerk, H.C., Rey, A., Graf, R., and Jacobs, A.M. (2000). Phonographic sublexical units in visual word recognition. *Current Psychology Letters*, *2,* 25.

Nygaard, L.C., Cook, A.E., and Namy, L.L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition, 112*(1), 181-186.

Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Olofsson, J.K., Nordin, S., Sequeira, H., and Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*(3), 247-265.

Osgood, C.E., and Suci, G.J. (1955). Factor analysis of meaning. *Journal of experimental psychology, 50*(5), 325-338.

Peirce, C.S. (1931). *The Collected Papers.* Eds. C. Hartshorne, P. Weiss, and A. Burks, Cambridge: Harvard University Press.

Perniss, P., Thompson, R.L., and Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 227. doi:10.3389/fpsyg.2010.00227

Perniss, P., and Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Phil. Trans. R. Soc. B*, *369*(1651), 20130300.

Plato (transl. 1892), “Cratylus,” in *The Dialogues of Plato*, Vol. 1, transl. by Benjamin Jowett, Oxford: Clarendon, 253-289.

Pratto, F., and John, O.P. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380-391.

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F.M., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. Proceedings of the National Academy of Sciences, *103*(20), 7865-7870.

Ramachandran, V.S., and Hubbard, E. M. (2001). Synaesthesia – A window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34.

Recio, G., Conrad, M., Hansen, L.B., and Jacobs, A.M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and language*, *134*, 34-43.

Russell, J.A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of Personality and Social Psychology*, *36*(10), 1152-1168.

Russell, J.A. (2003). *Core affect and the psychological construction of emotion.* Psychological Review, *110*(1), 145-172. doi: 10.1037/0033-295X.110.1.145

Sapir, E. (1929). A Study in Phonetic Symbolism. *Journal of Experimental Psychology, 12*(3)*,* 225-39.

Saussure, F. de (1916). *Cours de linguistique general.* Publie par Charles Bally et Albert Schehaye avec la collaboration de Albert Riedlinge, Lausanne et Paris, Payot

Schacht, A., and Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience, 9*(1)*,* 28-43.

Schmidtke, D.S., Conrad, M., and Jacobs, A.M. (2014). Phonological iconicity. *Frontiers in Psychology, 5,* 80*.* doi:10.3389/fpsyg.2014.00080

Schmidtke, D.S., Schröder, T., Jacobs, A.M., and Conrad, M. (2014). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior research methods*, *46*(4), 1108-1118.

Schupp, H.T., Hamm, A.O., and Weike, A.I. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science, 14*, 7-13.

Schupp, H.T., Öhman, A., Junghöfer, M., Weike, A.I., Stockburger, J., and Hamm, A.O. (2004). The facilitated processing of threatening faces: an ERP analysis. *Emotion, 4,* 189-200.

Tabachnick, B.G., and Fidell, L.S. (2001). *Using multivariate statistics (4th ed.).* Boston: Allyn and Bacon.

Van Orden, G.C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*(3), 181-198.

Võ, M. L.-H., Conrad, M., Kuchinke, L., Hartfeld, K., Hofmann, M.J., and Jacobs, A.M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods, 41*(2)*,* 534-539.

Wallis, J. (1699). *Grammar of the English language.* 5th edn. Oxford: L. Lichfield.

Watson, D., and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219-235.

Werner, H. (1934). 'L'unité des sens' [The unity of the senses]. *Journal de Psychologie Normale et Pathologique*, *31*, 190-205.

Werner, H. (1957). *Comparative Psychology of Mental Development.* (Rev. ed.) New York: International Universities Press.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and language*, *93*(1), 10-19.

Whissell, C.M. (1999). Phonosymbolism and the emotional nature of sounds: evidence of the preferential use of particular phonemes in texts of differing emotional tone. *Perceptual and Motor Skills*, *89*, 19-48. doi: 10.2466/pms.1999.89.1.19

Whissell, C. M. (2000). Phonoemotional profiling: a description of the emotional flavour of English texts on the basis of the phonemes employed in them. *Perceptual and Motor Skills*, *91*, 617–648.

Wiese, R. (1996). *The phonology of German.* New York: Oxford University Press.

Wundt, W. (1896). *Grundriss der Psychologie*. Leipzig: Engelmann.

Ziegler, J.C., Jacobs, A.M., and Klüppel, D. (2001). Pseudohomophone effects in lexical decision: still a challenge for current word recognition models. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 547.

3. EEG study part 2 – ERPs reveal an iconic relation between sublexical phonology and affective meaning

This chapter is published as:

Conrad, M.*, Ullrich, S.*, Schmidtke, D., Kotz, S. A. (2022). ERPs reveal an iconic relation between sublexical phonology and affective meaning. *Cognition, 226,* Article 105182, doi: 10.1016/j.cognition.2022.105182, licensed under CC BY-NC-ND 4.0.

* The authors contributed equally.

3.1. Abstract

Classical linguistic theory assumes that formal aspects, like sound, are not internally related to the meaning of words. However, recent research suggests language might code affective meaning such as threat and alert sublexically. Positing affective phonological iconicity as a systematic organization principle of the German lexicon, we calculated sublexical affective values for sub-syllabic phonological word segments from a large-scale affective lexical German database by averaging valence and arousal ratings of all words any phonological segment appears in. We tested word stimuli with either consistent or inconsistent mappings between lexical affective meaning and sublexical affective values (negative-valence/high-arousal vs. neutral-valence/low-arousal) in an EEG visual-lexical-decision task. A mismatch between sublexical and lexical affective values elicited an increased N400 response. These results reveal that systematic affective phonological iconicity – extracted from the lexicon – impacts the extraction of lexical word meaning during reading.

Keywords: sound symbolism, visual word recognition, phonological iconicity, affective meaning, N400, ERPs

3.2. Introduction

According to Saussure (1959) the arbitrary relation between the signifier and the signified is a fundamental feature of language. Nevertheless, there is also a long tradition stating that some semantic residue echoes in the mere sound of words (Bühler, 1934; Jakobson & Waugh, 1979; Jespersen, 1922; Tsur, 1992). Potential form-meaning mappings through structural resemblance, i.e., iconicity (Peirce, 1931), have strong implications for the evolution, development, and processing of language linking linguistic form to human experience (see Perniss & Vigliocco, 2014; Lockwood & Dingemanse, 2015).

Empirical support for non-arbitrary sound-meaning mappings is continuously growing (see Dingemanse et al., 2015, 2020; Perniss et al., 2010, for reviews). For instance, the influential *kiki-bouba* effect links phonology to the spatial dimension of shape as participants agree in their labeling of spiky or curvy shapes as either *kiki* or *bouba* (or *takete/maluma*; see Köhler, 1929; Ramachandran & Hubbard, 2001), replicable across languages (Ćwiek et al., 2021; Styles & Gawne, 2017) and age groups (Kawahara et al., 2019; Maurer et al., 2006; Ozturk et al., 2012; Peña et al., 2011). Asano et al. (2015) showed that presenting 11-month-old infants with spiky or round shapes and congruent or incongruent pseudowords (*kipi* or *moma*) yielded a larger N400 response to incongruent stimuli. Kovic et al. (2010) presented very similar results for adults with event-related-potential (ERP) effects arising at 200ms. Also, the link between size and phonology, which shows in labeling small versus large objects dependent on phonemic contrasts (Sapir, 1929), later refined as frequency-code-hypothesis (Ohala, 1983), was replicated by Thompson & Estes (2011). Going beyond artificial pseudoword material, Winter and Perlman (2021) described the mimesis of acoustics of small objects or animals for English size adjectives.

Cross-linguistic studies used (Japanese) ideophones (marked words that depict sensory imagery, Dingemanse, 2012) as natural sound-symbolic stimuli (Dingemanse et al., 2016; Dingemanse, 2018). Lockwood et al. (2016) suggested principles of phonological iconicity to be effective across language boundaries: learning the real compared to an incorrect meaning of foreign ideophones correlated with more correct memories and an increased P3 and late positive ERP complex. Iconicity also seems to affect the efficiency of language processing. Building upon seminal studies on signed languages (Thompson et al., 2010), Sidhu at al. (2019) using iconicity ratings for English words (Perry et al., 2015; Winter et al., 2017), reported lexical-decision advantages for more iconic words. Recent corpus-linguistic studies explored phonological systematicity of existing vocabularies, indicating non-arbitrary sound-meaning relations to permeate the lexicon for shape (Sidhu et al., 2021, demonstrating the *kiki-bouba* effect in English words), size (Winter & Perlman, 2021), color (Johansson et al., 2020), spatial relations (Johansson & Zlatev, 2013), or grammar (Kelly, 1992; Pimentel et al., 2019; Shih, 2020). All in all, research on phonological iconicity has involved various meaning dimensions (see Dingemanse et al., 2020; Schmidtke et al., 2014, for reviews) and recent studies emphasized its importance in language acquisition (e.g., Kantartzis et al., 2019; Nielsen & Dingemanse, 2020) and diachronic language change (Dellert et al., 2021; Monaghan & Roberts, 2021; Vinson et al., 2021).

Here, we propose that emotional relevance is a driving factor of phonological iconicity. Osgood and Suci (1955) showed with a semantic differential that most variety in lexico-semantic meaning can be accounted for by valence and arousal, defining the most widely used dimensional conceptualization of emotion (Barrett, 2006; Russell, 1980; Wundt, 1904). In general, communication of affect, e.g., interjections, may have been crucial for language development (Darwin, 1871; Jespersen, 1922; Panksepp, 2008). Preceding and modulating the emergence of vocabulary, vocal emotion expressions, likely encoded by different sounds, might feature an iconic, internal relation with specific emotions. Consequently, we suggest that iconic phonological coding of affect is still part of modern vocabularies, providing affective cues at the sublexical level of phonemes.

Partly in line with this, cognitive poetics provided heterogeneous results on the use of specific phonemes varying with the emotional content of literature (Auracher et al., 2010, 2020; Aryani et al., 2016; Fónagy, 1961; Miall, 2001; Whissell, 1999, 2000). Also, due to their articulatory overlap with facial emotional expressions, single phonemes, e.g., /i:/ vs. /o:/, could represent positive vs. negative features (Rummer & Schweppe, 2019). Furthermore, valence ratings in five Indo-European languages were associated with word initial phonemes, with decreasing naming latencies for negative valence suggesting an iconic source of rapid alert (Adelman et al., 2018). Similarly, Aryani et al. (2018) related affective impressions to specific phonetic features in pseudowords (but see Monaghan & Fletcher, 2019).

To investigate whether affective iconicity systematically permeates the vocabulary of a language – beyond single “iconic words” or single phonemes – we used a large-scale approach for over 6,000 German words (Schmidtke & Conrad, 2018). We posit that, for instance, words with negative and arousing semantic meaning typically contain specific sublexical phonological units that serve as sublexical markers of alert (see Adelman et al., 2018). The wider such phenomena spread across the vocabulary, the more likely meaningful “sublexical affective values”, henceforth SAV, can be calculated for phonological units, based on their distribution across the overall lexical affective space of valence and arousal.

We have previously shown how these SAV differ between poems in accordance to affective labels from their author (Aryani et al., 2016; Ullrich et al., 2017), and correlate with ERPs during prelexical processing in a lexical-decision-task (Ullrich et al., 2016). Schmidtke and

Conrad (2018) showed that “high-arousal” sublexical units are detected faster in a visual search task, suggesting an iconic link between SAV and alert.

Importantly, these prior SAV results were obtained regardless of lexical affective meaning of target words (Schmidtke & Conrad, 2018; Ullrich et al. 2016). They can, therefore, only be attributed to prelexical processing (see also Sučević, et al. 2015). However, phonological units with a bias to occur more often in words with specific affective meaning might systematically carry saliency across levels of language processing. What remains unclear, is, thus, whether affective phonological iconicity plays a role for higher, cognitive, levels of language processing beyond prelexical perception, when a word’s formal aspects must be integrated with semantic meaning.

We hypothesize that extracting meaning from printed words is sensitive to systematic affective sound-meaning-correspondences across the lexicon of a language, and words with a consistent mapping between SAVs and lexical semantics are easier to process.

The present study tested this twofold hypothesis in a lexical-decision task, manipulating both the lexical affective meaning of German words (see Citron, 2012; Kotz & Paulmann, 2011, for reviews) and the potential affective iconicity or their sublexical phonology. We assessed the ease of lexical access by means of the N400 (Kutas & Federmeier, 2011), an event-related brain response in the EEG. An N400 amplitude indicates how difficult it is to integrate a word into a given context (see Barber & Kutas, 2007, for a review). We hypothesized that the N400 response would decrease when *sublexical* and *lexical affective values* match in affectively iconic words.

3.3. Materials and methods

3.3.1. Participants

41 right-handed native German speakers without neurological or vision problems, students of the Freie-Universität-Berlin, participated in the study. Data of 35 participants − with minimum 50 segments surviving artifact-rejection in every experimental condition − were further analyzed (21 women; age: M = 26.7, SD = 4.2 years).

3.3.2. Stimuli and design

312 words were selected from the Schmidtke and Conrad list (2018) in a 2x2 design, involving an orthogonal manipulation of the factors *lexical (LAV)* and *sublexical-affective-values (SAV)* (each time contrasting "alert": negative-valence/high-arousal vs. "neutral": neutral-valence/low-arousal. Lexically negative/high-arousal stimuli had at least moderately negative valence (< -.82 on a -3 to +3 scale) and at least moderately elevated arousal ratings (> 2.83; 1−5 scale). Neutral lexical values ranged between -.76 and +.75 (valence) and 1.67 to 2.79 (arousal).

SAV calculation: All words in the database were transcribed phonemically and segmented into syllabic onsets, nuclei, and codas. For each phonological segment, normative rating values of all words comprising it in an identical syllabic position were averaged, e.g., averaging arousal/valence ratings for all words containing a syllabic onset /kr/ for the SAV for /kr/. Then, for every word, SAVs for all segments were averaged to assign words to the different cells of the factor *sublexical-affective-values*. A split half of the resulting scales at -.04 (valence), and at 2.90 (arousal) assigned stimuli to different *sublecixal-affective-value* conditions (see Table 1).

This made, e.g., *Krieg* (war) and *Zucht* (military discipline) **negative/high-arousal stimuli with consistent affective phonology**, but *Fluch* (curse) and *Mord* (murder) **iconically inconsistent**.

For **lexically-neutral words**, phonology was **affectively consistent** in, e.g., *Glas* (glass) and *Land* (land), but **inconsistent** in *Topf* (pot) or *Preis* (price).

Words contained maximum nine letters to avoid refixations. Overall word length and frequency, imageability, word class, orthographic/phonological neighborhood, syllabic structure and complexity, positional frequencies of all single graphemes/phonemes, bigrams/biphones, or syllabic segments were matched across the four cells of the design. Nonwords were pronounceable pseudowords matching word stimuli in number, length, and syllabic structure. Stimuli (and stim-characteristics) are available in Conrad et al., 2022, (see also Ullrich et al., 2016, using the same stimuli as "maximum controlled set" for prelexical SAV effects).

3.3.3. Procedure

Stimuli appeared in the center of a computer screen in white color on a black background using Presentation software (Version 0.70, NeuroBehavioral Systems, Inc., 2004). Participants were instructed to indicate as fast and accurately as possible whether the presented stimulus was a "word" or not via two buttons. Left- and right-hand responses to words and nonwords were counterbalanced across participants. Each trial began with a fixation cross (500 ms), followed by a blank screen (500 ms). Stimuli were presented in randomized order for 500ms, followed by a blank screen until participants responded, and a subsequent uniform-random-scattered inter-stimulus interval of 700 − 1500ms.

Table 1 listing stimulus example words for the four conditions of the 2x2 design contrasting lexical and sublexical (SAV) affective values (negative-valence/high-arousal vs. neutral-valence/low-arousal). Specific valence (V) and arousal (A) values are given for syllabic onsets (ON), nuclei (NU), and codas (CO) determining SAV for the entire stimuli. N words gives the number of words in the database of Schmidtke and Conrad (2018) used to calculate this SAV for phonological segments. Conditions with consistent matching between SAV and lexical affective values are shadowed.

Category	STIMULUS	Lexical valence	SAV V_STIM	SAV_V_ON (N words)	SAV_V_NU (N words)	SAV_V_CO (N words)	Lexical arousal	SAV A_STIM	SAV_A_ON (N words)	SAV_A_NU (N words)	SAV_A_CO (N words)
Lex. neg-high and Sublex. neg-high	**Krieg** krik (war)	-2.90	**-0.12**	-0.51 (78)	0.08 (970)	0.02 (265)	4.57	**2.97**	3.10 (78)	2.94 (970)	2.93 (265)
	Zucht =Uxt (discipline)	-1.23	**-0.13**	-0.03 (345)	-0.13 (742)	-0.29 (82)	3.41	**2.95**	2.9 (345)	2.95 (742)	3.09 (82)
	Raub rBp (robbery)	-1.8	**-0.14**	0.02 (742)	-0.09 (342)	-0.38 (190)	3.89	**2.93**	3.02 (742)	2.85 342	3.0 (190)
Lex. neg-high and Sublex. neut-low	**Fluch** flux (curse)	-2.1	**0.05**	-0.007 (56)	-0.23 (376)	-0.12 (543)	3.65	**2.86**	2.94 (56)	2.86 (376)	2.87 (543)
	Mord mOrt (murder)	-2.8	**0.13**	0.004 (598)	-0.12 (488)	0.15 (84)	4.44	**2.89**	2.84 (598)	2.91 (488)	2.98 (84)
	Leid lWt (woe)	-2	**-0.02**	-0.03 (799)	-0.03 (614)	-0.01 (387)	4.17	**2.82**	2.79 (799)	2.91 (614)	2.88 (387)
Lex. neut-low and Sublex. neg-high	**Preis** prWs (price)	0.1	**-0.1**	-0.11 (79)	-0.03 (614)	-0.14 (522)	2.61	**2.91**	3.01 (79)	2.91 (614)	2.92 (522)
	Topf tO+ (pot)	0.19	**-0.16**	-0.004 (987)	-0.10 (488)	-0.41 (13)	2.04	**2.92**	2.92 (987)	2.91 (488)	2.9 (13)
	Reis 'rWs (rice)	0.69	**-0.05**	0.02 (742)	-0.03 (614)	-0.14 (522)	2.03	**2.92**	3.02 (742)	2.91 (614)	2.92 (522)
Lex. neut-low and Sublex. neut-low	**Glas** glas (glass)	0.64	**-0.03**	0.13 (30)	-0.04 (778)	-0.14 (522)	1.77	**2.80**	2.77 (30)	2.81 (778)	2.92 (522)
	Land l&nt (land)	0.50	**0.04**	-0.03 (799)	0.12 (1579)	0.26 (249)	1.82	**2.82**	2.79 (799)	2.91 (1579)	2.78 (249)
	Moll mOl (minor)	-0.1	**1.02**	-0.04 (598)	-0.01 (488)	0.04 (568)	2.28	**2.84**	2.84 (598)	2.91 (488)	2.82 (568)

3.3.4. EEG recording and (pre-)processing

EEG recording used, fixed to the scalp in an elastic cap, 61 AgCl-electrodes (Fp1, Fpz, Fp2, AF3, AF4, F5, F3, F1, Fz, F2, F4, F6, FT7, FC3, FC1, FCz, FC2, FC4, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P9, P7, P5, P3, P1, Pz, P2, P4, P6, P8, P10, PO9, PO7, PO3, POz, PO4, PO8, PO10, O1, Oz, O2, Iz, M1, M2) using two 32-channel amplifiers (BrainAmp, Brain Products, Germany) according to the International 10–20 system (American Electroencephalographic Society, 1991; Jasper, 1958). Average impedances were kept below 2 kΩ. The electrooculogram (EOG) was monitored by two electrodes at the outer canthi of the participant's eyes and two electrodes above and below the right eye. EEG and EOG signals were recorded with a sampling rate of 500 Hz, referenced to the right mastoid, but re-referenced offline to linked mastoids. The AFz electrode was used as ground electrode. Later offline filtering included a bandpass filter of 0.1 – 20Hz and a 50 Hz notch filter. Independent component analysis (ICA; Makeig et al., 1998) served to identify and remove eye movement artifacts. The continuous EEG signal was cut into segments of 950 ms total length, 150 ms pre-stimulus baseline plus 800 ms post-stimulus interval. After baseline correction, automatic artifact-rejection excluded trials containing differences > 80 µV in intervals of 70 ms or amplitudes > 50 or < -50µV. An N400 effect was expected around 300 and 500 ms post-stimulus onset (Kutas & Federmeier, 2011; Lau et al., 2008). Within this time window, we ran a peak-detection for the grand averaged data. The observed negative deflection at CPz, most representative of the N400 (Dimigen et al., 2011; Kutas & Federmeier, 2011), peaked at 354 ms. Accordingly, we chose a time window of 150 ms (280 – 430 ms) with this peak in its center. The following 9 central electrodes entered a centro-posterior region of interest (ROI) analysis for the N400 (see Dimigen et al., 2011; Kutas & Federmeier, 2011, but see Šoškić et al., 2021, for extensive review): C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2. Segments corresponding to correctly answered word trials with a mean activity [µV] deviating less than 2 SD from means per participant and condition were subjected to further analysis (mean N segments/condition: 60.3 − 61.7, no significant differences).

3.3.5. Data analysis

Behavioral data

Correct response latencies (2 SD outlier-trimmed) and errors were analyzed using linear mixed-effects models and mixed-effects logistic regression.

EEG data

ROI analyses were conducted based on linear mixed-effects models with the mean activity [µV] values of the selected time window using R 4.1.1 (R Core Team, 2021) and *lme4* (v1.1-27.1; Bates et al., 2015). Visual inspection of plots of residuals against fitted values did not reveal deviations from normality or homoscedasticity. Within-subject factors *lexical* (LAV; 2 levels) and *sublexical-affective-values* (SAV; 2 levels) were used as fixed factors. Random effects were constructed using random intercepts for items and subjects as well as random slopes for subjects for the main effects, constituting the most complex structure of random effects before models failed to converge.

The final model reads:

µV ~ 1 + LAV * SAV (1 + LAV + SAV|subject) + (1|item)

P-values were obtained by testing the full model containing all fixed effects of interest against the null-model containing only random effects by likelihood-ratio-tests, followed by an ANOVA based on Satterthwaite approximation of degrees of freedom using the *afex*-package (Singmann et al., 2021). Significant interactions between the two factors were examined using Tukey-adjusted post-hoc tests using the *emmeans*-package (Lenth, 2021) looking for a) a main effect of inconsistency (non-matching vs. matching sublexical-lexical affective values), and

b) differential effects of SAV within the two cells of the factor *lexical-affective-values*. Measuring effect size, we calculated Cohen's d using *EMAtools*-package (Kleiman, 2021). Figures were created using *ggplot2*-package (Wickham, 2016).

3.4. Results

Dataset and R-analysis-script can be retrieved from https://osf.io/gam9y/ (Conrad et al., 2022).

3.4.1. Behavioral results

No effects for either analysis were observed in the RT or error data (all *p's* > .1).

3.4.2. N400 ROI analysis

Within the time window of 280 – 430ms, we found no main effects, but a highly-significant interaction between the factors *sublexical* and *lexical-affective-values*, $\chi^2(3, N = 12) = 13.66$, $p = 0.003$, $b = 1.34$, $SE = 0.37$, $d = 0.42$, ($F(1, 305.23) = 13.35$, $p < .001$). Collapsing the four cells into a contrast of *inconsistent* vs. *consistent* sound-to-meaning mapping of words, the interaction resulted in a significant consistency effect (Figure 1), $\chi^2(1, N = 7) = 10.78$, $p = 0.001$, $b = -0.67$, $SE = 0.20$, $d = -0.88$, ($F(1, 61.01) = 11.85$, $p = .001$), with increasing negativity for *inconsistent* ($M = 3.31$) compared to *consistent* words ($M = 3.98$). Resolving the interaction (see Figure 2), we found inverted SAV effects for the different conditions of *lexical-affective-values*: N400 amplitudes significantly increased with alert as compared to neutral SAVs for lexically-neutral words ($M = 3.19$ vs. $M = 3.90$; $z = -2.676$, $p = 0.037$, $d = -0.45$; Figure 3), but tended to decrease for the same contrast in negative/high-arousing words ($M = 4.05$ vs. $M = 3.42$; $z = 2.384$, $p = 0.0801$, $d = 0.41$; Figure 4).

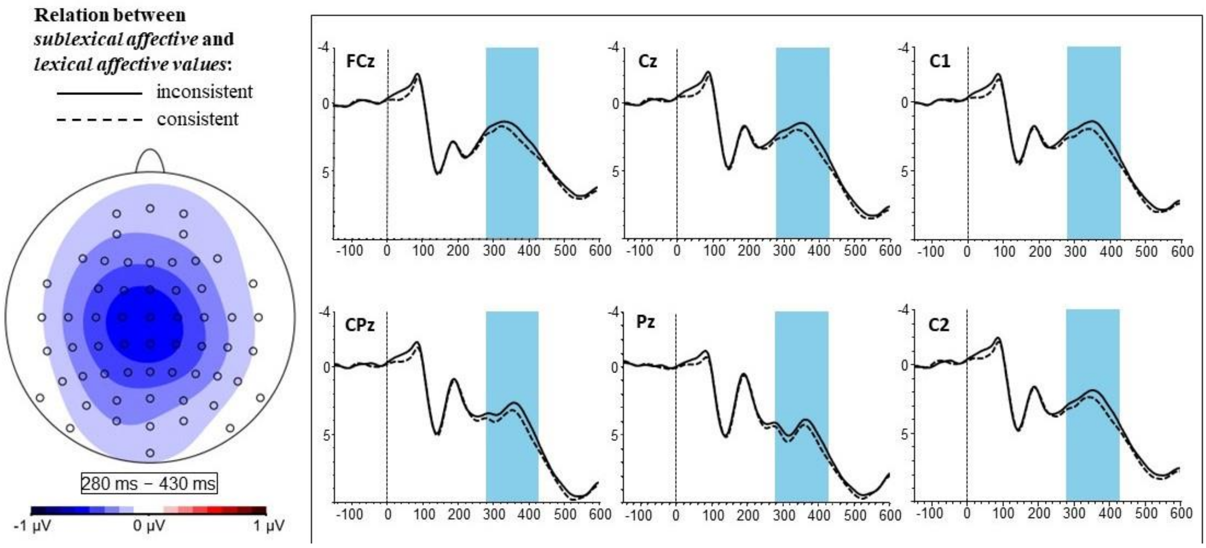

Figure 3.1: Overall ERP effect of inconsistency between lexical and sublexical affective values.

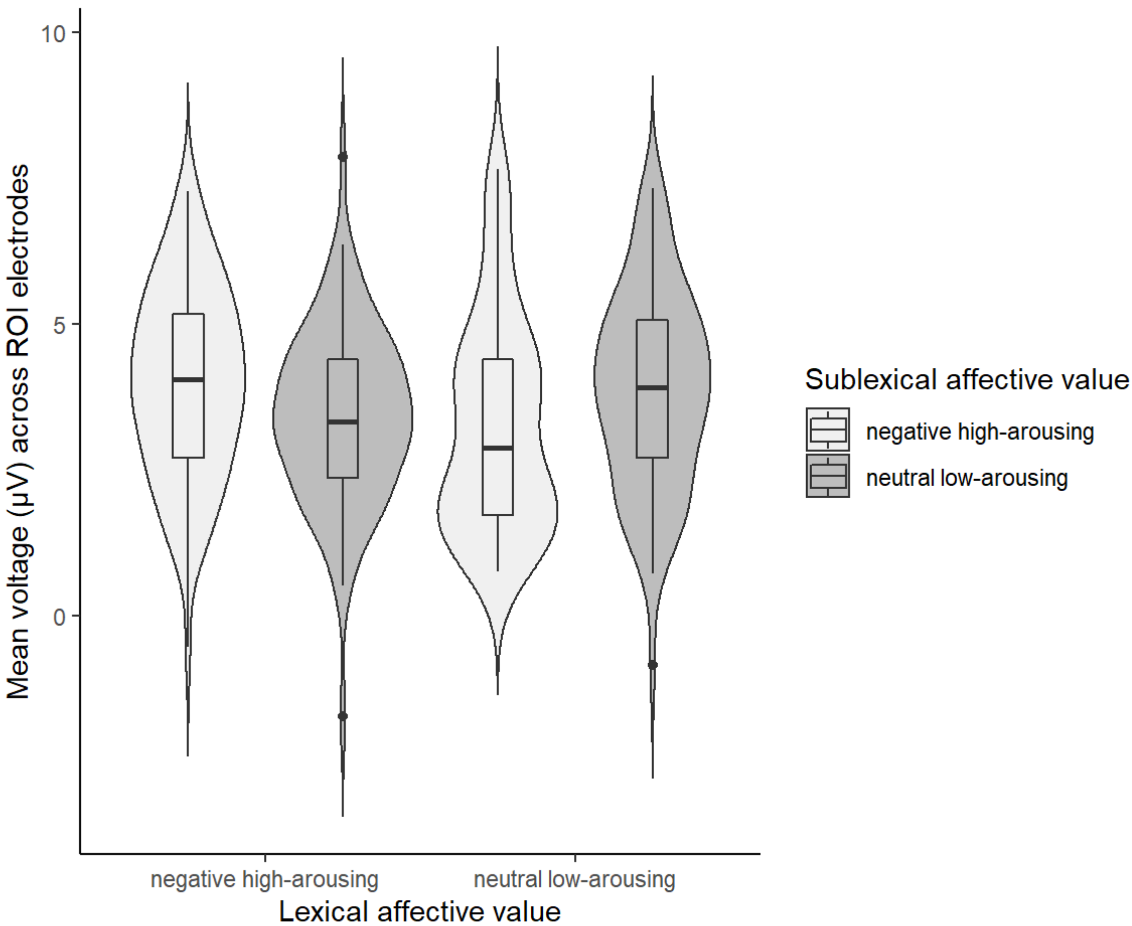

Figure 3.2: Mean N 400 amplitudes for stimulus words in all different conditions of lexical and sublexical affective values.

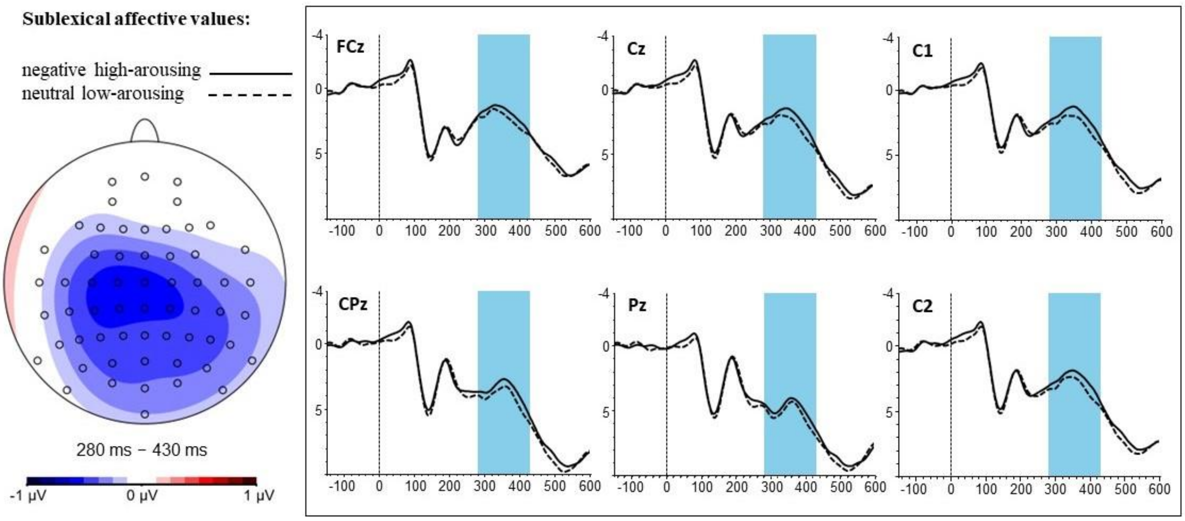

Figure 3.3: ERP effects of sublexical affective values in words of neutral-valence/low-arousal.

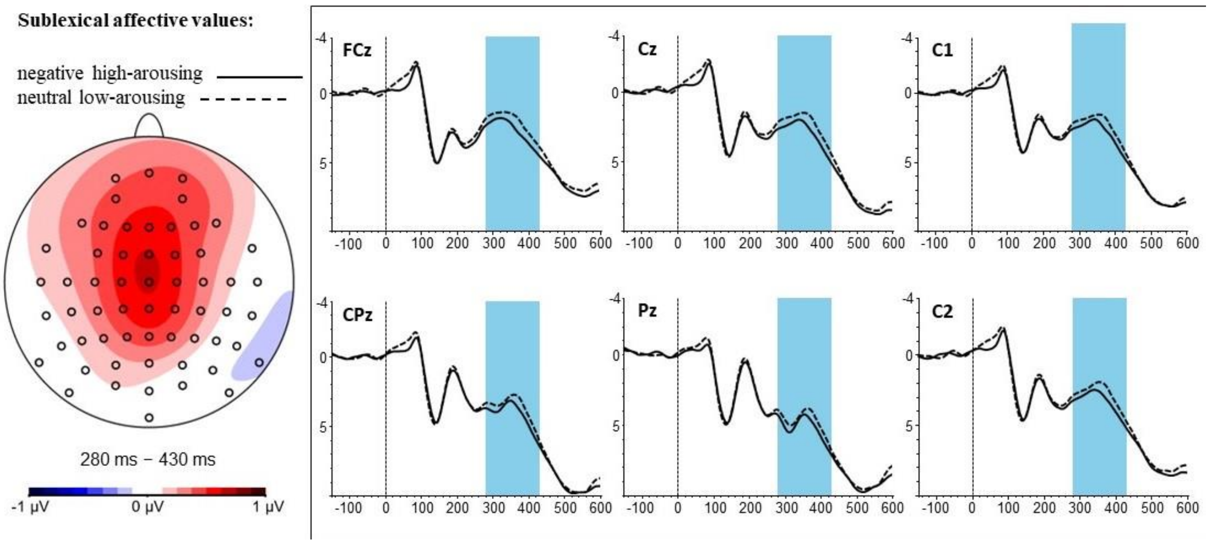

Figure 3.4: ERP effects of sublexical affective values in words of negative-valence/high-arousal.

3.5. Discussion

We confirm an increased N400 response for mismatching sublexical phonology and semantic affective meaning – in a lexical-decision-task where neither overt phonology nor explicit processing of affective dimensions was required.

This suggests that participants were sensitive to affective phonological iconicity using sublexical markers of affect – increasing semantic processing effort for respective mismatch. Our results extend previous findings of iconicity enhancing language processing (Schmidtke & Conrad, 2018; Sidhu at al., 2019; Thompson et al., 2010), connecting a large-scale approach on the systematic affective-iconic organization of language with neuroscientific evidence of resulting consequences:

Our data show that the German language systematically involves affective phonological iconicity in both a) the way phonological segments occur across the bi-dimensional lexical affective space of valence and arousal, and b) these sound-to-meaning correspondences in the lexicon determine automatic access to the meaning of words.

Our SAV operationalization implies language uses position-specific intrasyllabic phoneme clusters as markers of affect – increasing the signal power of words with high emotional relevance, in particular, involving threat (see Adelman et al., 2018). Iconic phonological patterns emerging from our purely numerical approach can already be seen at the phonemic feature level (Johansson & Zlatev, 2013), e.g., shorter vowel length, voiceless sibilants, and decreasing sonority of consonants apparently associate with "alert" SAV (see Table 1 and the Appendix). On the other hand, also complex combinations of consonants – typical for German language – offer a wide range of highly salient phonological units possibly carrying intrinsic relations with affect to a more complex level, e.g., syllabic onsets of our sublexical "alert" stimuli – frequently occurring in words of high arousal – kr (kr) and z (=), may appear like a mimesis of the menacing sounds produced by predators highly feared by our ancestors (wolves and snakes).

Our data suggest that language phonologically grounds symbols (words) representing semantic meaning in affective experience – presumably since its origin. Language serves a denotative and an appellative function (Bühler, 1934; Jakobson, 1960). As to the latter, using the entire span of arousal and valence, affective iconicity of phonemes might mark any message as alerting or reassuring, and intuitively trigger approach or avoidance – sounding, for instance, exciting or dull, smooth and mellow, or sharp and cynical – using preferentially sonorants vs. plosives, voiceless sibilants, long vs. short vowels, etc. – supplying words with an affective tonality – described by SAV.

Using only negative affective stimuli may have prevented behavioral effects to arise. Behavioral and ERP effects are not always associated or directly correlated (e.g., Barber et al., 2013). The N400 is considered a more sensitive measure of semantic activation than – and often appears without – RT effects (e.g., Heil et al., 2004; Kotz, 2001). These represent only the final (response) point of a complex process and are very sensitive to control (Neely, 1991). Unlike positive stimuli that consistently trigger speeded responses, negative words produce heterogenous behavioral effects where a general processing advantage for affective stimuli (Kousta et al., 2009, see Kauschke et al., 2019, for review) can be opposed to a tendency to avoid negative stimuli (Kuperman et al., 2014; Estes & Verges, 2008). Here, a cognitive processing advantage for iconic negative words (as reflected in the N400) might not have speeded responses because of behavioral avoidance tendencies (Estes & Verges, 2008, Brouillet & Syssau, 2005) that might increase when the affective load of negative stimuli is emphasized through phonological iconicity.

3.6. Funding

This work was supported by grants to M. C. (201 “Bilingualism and affectivity in reading”) and to M. C. and S. K. (410 “Sound physiognomy in language organization, processing, and production”) from the German Research Foundation (DFG) at the Cluster of Excellence “Languages of Emotion” at the Free University of Berlin.

3.7. References

Adelman, J., Estes, Z., & Cossu, M. (2018). Emotional sound symbolism: Languages rapidly signal valence via phonemes, *Cognition, 175*, 122–130.

American Electroencephalographic Society (1991). Guidelines for standard electrode position nomenclature. *J. Clin. Neurophysiol., 8*, 200–202. doi:10.1097/00004691-199104000-00007

Aryani, A., Conrad, M. Schmidtke, D., & Jacobs, & A. M. (2018). Why “piss” is ruder than “pee”? The role of sound in affective meaning making. *PloS one, 13*(6), e0198430. doi:10.1371/journal.pone.0198430

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. (2016). Measuring the Basic Affective Tone of Poems via Phonological Saliency and Iconicity. *Psychology of Aesthetics, Creativity, and the Arts, 10,* 191–204. doi:10.1037/aca0000033

Auracher, J., Albers, S., Zhai, Y., Gareeva, G., & Stavniychuk, T. (2010). P is for happiness, N is for sadness: universals in sound iconicity to detect emotions in poetry. *Discourse Processes*, *48*(1), 1–25. doi:10.1080/01638531003674894

Auracher, J., Menninghaus, W., Scharinger, M. (2020) Sound Predicts Meaning: Cross-Modal Associations Between Formant Frequency and Emotional Tone in Stanzas. *Cognitive Science, 44*(10): e12906. doi:10.1111/cogs.12906

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex, 63,* 196–205. doi:10.1016/j.cortex.2014.08.025

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews, 53*(1), 98–123. doi:10.1016/j.brainresrev.2006.07.002

Barber, H. A., Otten, L. J., Kousta, S. T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and language*, *125*(1), 47–53.

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and social psychology review*, *10*(1), 20–46. doi:10.1207/s15327957pspr1001_2

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Grothendieck, G. (2011). Package 'lme4'. *Linear mixed-effects models using S4 classes. R package version*, *1*(6).

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

BrainVision Analyzer (Version 2.2.0) [Software]. (2019). Gilching, Germany: Brain Products GmbH.

Brouillet, T., & Syssau, A. (2005). Connection between the evaluation of positive or negative valence and verbal responses to a lexical decision making task. *Canadian journal of experimental psychology = Revue canadienne de psychologie experimentale*, *59*(4), 255–261.

Bühler, K. (1934). *Sprachtheorie.* Stuttgart: G. Fischer.

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and language*, *122*(3), 211–226. doi:10.1016/j.bandl.2011.12.007

Conrad, M., Ullrich, S., Schmidtke, D., Kotz, S. A. (2022). Dataset: ERPs reveal an iconic relation between sublexical phonology and affective meaning. Retrieved from https://osf.io/gam9y.

Ćwiek, A., Fuchs, S., Draxler, C., Asu, E. L., Dediu, D., Hiovain, K., … & Winter, B. (2021). The bouba/kiki effect is robust across cultures and writing systems. *Phil. Trans. R. Soc. B*, *377*: 20200390. https://doi.org/10.1098/rstb.2020.0390

Darwin, C. (1871). The descent of man. *The Great Books of the Western World*, *49*, 320.

Dellert, J., Erben Johansson, N., Frid, J., & Carling, G. (2021). Preferred sound groups of vocal iconicity reflect evolutionary mechanisms of sound stability and first language acquisition: evidence from Eurasia. *Phil. Trans. R. Soc. B,* 376: 20200190. doi:10.1098/rstb.2020.0190

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Co-Registration of eye movements and EEG in natural reading: analyses and review. *Journal of Experimental Psychology: General*, *140*(4), 552–572. doi:10.1037/a0023885

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics compass*, *6*(10), 654–672.

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: a journal of general linguistics*, *3*(1).

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in cognitive sciences*, *19*(10), 603–615.

Dingemanse, M., Perlman, M., & Perniss, P. (2020). Construals of iconicity: experimental approaches to form-meaning resemblances in language. *Language and Cognition, 12*(1), 1–14. doi:10.1017/langcog.2019.48

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language*, *92*(2), e117–e133.

Estes, Z., & Verges, M. (2008). Freeze or flee? Negative stimuli elicit selective responding. *Cognition, 108(*2), 557–565.

Fónagy, I. (1961). Communication in poetry. *Word*, *17*, 194–218.

Heil, M., Rolke, B., & Pecchinenda, A. (2004). Automatic semantic activation is no myth: Semantic context effects on the N400 in the letter-search task in the absence of response time effects. *Psychological Science*, *15*(12), 852–857.

Holcomb, P. J., Grainger, J., and O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience, 14*(6), 938–950. doi:10.1162/089892902760191153

Johansson, N. & Zlatev, J. (2013). Motivations for sound symbolism in spatial deixis: a typological study of 101 languages. *Public J. Semiotics, 5*, 3–20.

Jakobson, R. (1960). Linguistics and poetics. In *Style in language* (pp. 350–377). MA: MIT Press.

Jakobson, R., & Waugh, L. R. (1979). *The Sound Shape of Language.* Bloomington and London: Indiana University Press.

Jasper, H. (1958). The ten–twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol., 10*, 371–375.

Jespersen, O. (1922). *Language – Its Nature, Development and Origin.* London: George Allen and Unwin Ltd.

Johansson, N., Anikin, A., & Aseyev, N. (2020). Color sound symbolism in natural languages. *Language and Cognition*, *12*(1), 56–83.

Johansson, N., & Zlatev, J. (2013). Motivations for sound symbolism in spatial deixis: A typological study of 101 languages. *Public Journal of Semiotics*, *5*(1), 3–20.

Kantartzis, K., Imai, M., Evans, D., & Kita, S. (2019). Sound Symbolism Facilitates Long-Term Retention of the Semantic Representation of Novel Verbs in Three-Year-Olds. *Languages*, *4*(2), 21. doi:10.3390/languages4020021

Kauschke, C., Bahn, D., Vesker, M., & Schwarzer, G. (2019). The role of emotional valence for the processing of facial and verbal stimuli—positivity or negativity bias? *Frontiers in psychology*, *10*, 1654.

Kawahara, S., Isobe, M., Kobayashi, Y., Monou, T., Okabe, R., & Minagawa, Y. (2019). Acquisition of the takete-maluma effect by Japanese speakers. *Reports of the Keio Institute of Cultural and Linguistic Studies*, *50*(63), 78.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological review*, *99*(2), 349.

Kleiman, E. (2021). EMAtools: Data Management Tools for Real-Time Monitoring/Ecological Momentary Assessment Data. R package version 0.1.4. https://CRAN.R-project.org/package=EMAtools

Köhler, W. (1929). *Gestalt psychology*. New York, USA: Liveright.

Kotz, S.A. (2001). Neurolinguistic evidence for bilingual memory representation: A comparison of reaction times and event-related brain potentials. *Bilingualism: Language and Cognition, 4*(2), 143–154.

Kotz, S. A., & Paulmann, S. (2011). Emotion, language, and the brain. *Lang. Linguist. Compass*, *5*, 108–125. doi:10.1111/j.1749-818X.2010.00267.x

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition, 114*(1), 19–28

Kousta, S. T., Vinson, D. & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*, 473–81.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, *143*(3), 1065.

Kutas, M. & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.*, *62*, 621–47.

Lau, E., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci., 9*, 920–933. doi:10.1038/nrn2532

Lenth, R. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.0. https://CRAN.R-project.org/package=emmeans

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in psychology*, *6*: 1246. doi:10.3389/fpsyg.2015.01246

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: Neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra: Psychology*, *2*(1), 7.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of Electroencephalographic data. *Adv. Neural Inf. Process. Syst., 8*, 145–151.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound-shape correspondances in toddlers and adults. *Dev. Sci., 9*, 316–322. doi: 10.1111/j.1467-7687.2006.00495.x

Miall, D. S. (2001). Sounds of contrast: An empirical approach to phonemic iconicity. *Poetics, 29*, 55–70. doi:10.1016/S0304-422X(00)00025-5

Monaghan, P., & Fletcher, M. (2019). Do sound symbolism effects for written words relate to individual phonemes or to phoneme features? *Language and Cognition*, *11*(2), 235–255. doi:10.1017/langcog.2019.20

Monaghan, P., & Roberts, S. G. (2021). Iconicity and diachronic language change. *Cognitive Science*, *45*(4), e12968.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.

NeuroBehavioral Systems, Inc. (2004). Presentation (Version 0.70) [Software]. Albany, CA: Retrieved from www.neurobs.com

Nielsen, A. K., & Dingemanse, M. (2020). Iconicity in word learning and beyond: A critical review. *Language and speech*, *64*(1), 52–72. doi:10.1177/0023830920914339.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In *The production of speech* (pp. 189–216). Springer, New York, NY.

Osgood, C. E., & Suci, G. J. (1955). Factor analysis of meaning. *Journal of experimental psychology*, *50*(5), 325.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: evidence for sound–shape cross-modal correspondences in 4-month-olds. *J. Exp. Child Psychol., 114*, 173–186. doi:10.1016/j.jecp.2012. 05.004

Panksepp, J. (2008). The power of the word may reside in the power of affect. *Integrative Psychological and Behavioral Science*, *42*(1), 47–55. doi:10.1007/s12124-007-9036-5

Peirce, C. S. (1931). *The Collected Papers of Charles Sanders Peirce* (C. Hartshorne, P. Weiss, & A. Burks, eds.). Cambridge: Harvard University Press.

Peña, M., Mehler, J., & Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psychol. Sci., 22*, 1419–1421. doi:10.1177/0956797611421791

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Front. Psychol., 1,* 227. doi:10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Phil. Trans. R. Soc. B.*, *369*(1651), 20130300.

Perry, L. K., Perlman, M. & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLoS ONE, 10*(9), e0137147. https://doi.org/10.1371/journal.pone.0137147

Pimentel, T., McCarthy, A. D., Blasi, D. E., Roark, B., & Cotterell, R. (2019). Meaning to form: measuring systematicity as information. *arXiv preprint arXiv:1906.05906*.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia – A window into perception, thought and language. *Journal of Consciousness Studies, 8,* 3–34.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and language*, *134*, 34–43. doi:10.1016/j.bandl.2014.03.009

Rummer, R., & Schweppe, J. (2019). Talking emotions: vowel selection in fictional names depends on the emotional valence of the to-be-named faces and objects. *Cognition and Emotion, 33*(3), 404–416. doi:0.1080/02699931.2018.1456406

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology, 12*, 225–239.

Saussure, F. de (1959). *Course in General Linguistics*. New York, NY: Philosophical Library. (Original work published 1922)

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in psychology*, *5,* Article 80. doi:10.3389/fpsyg.2014.00080

Schmidtke, D., & Conrad, M. (2018). Effects of affective phonological iconicity in online language processing: Evidence from a letter search task. *Journal of Experimental Psychology: General, 147*(10), 1544–1552. doi:10.1037/xge0000499

Shih, S. S. (2020). Gradient categories in lexically-conditioned phonology: An example from sound symbolism. *Proceedings of the Annual Meetings on Phonology, 8*. doi:10.3765/amp.v8i0.4689

Sidhu, D. M., Vigliocco, G., & Pexman, P. M. (2019). Effects of iconicity in lexical decision. *Language and Cognition, 12*(1), 164–181.

Sidhu, D. M., Westbury, C., Hollis, G., & Pexman, P. M. (2021). Sound symbolism shapes the English language: The maluma/takete effect in English nouns. *Psychonomic Bulletin & Review, 28*(4), 1390–1398.

Singmann, H., Bolker, B., Westfall, J., Aust, F. & Ben-Shachar, M. S. (2021). afex: Analysis of Factorial Experiments. R package version 1.0-1.

Sučević, J., Savić, A. M., Popović, M. B., Styles, S. J., & Ković, V. (2015). Balloons and bavoons versus spikes and shikes: ERPs reveal shared neural processes for shape–sound–meaning congruence in words, and shape–sound congruence in pseudowords. *Brain and language, 145*, 11–22.

Šoškić, A., Jovanović, V., Styles, S. J., Kappenman, E. S., & Ković, V. (2021). How to do better N400 studies: Reproducibility, consistency and adherence to research standards in the existing literature. *Neuropsychology Review*, 1–24.

Styles, S. J., & Gawne, L. (2017). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. *i-Perception, 8*, 1–17. doi:10.1177/2041669517724807

Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Quart. J. Exp. Psychol., 64*, 2392–2404. doi:10.1080/17470218.2011.605898

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British Sign Language: Effects of iconicity for phonological decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(4), 1017.

Tsur, R. (1992). *What makes sound patterns expressive? The poetic mode of speech perception*. Durham, NC: Duke University Press.

Ullrich, S., Aryani, A., Kraxenberger, M., Jacobs, A. M., & Conrad, M. (2017). On the Relation between the General Affective Meaning and the Basic Sublexical, Lexical, and Inter-lexical Features of Poetic Texts—A Case Study Using 57 Poems of H. M. Enzensberger. *Frontiers in Psychology, 7*, Article 2073. doi:10.3389/fpsyg.2016.02073

Ullrich, S., Kotz, S. A., Schmidtke, D. S., Aryani, A., & Conrad, M. (2016). Phonological iconicity electrifies: An ERP study on affective sound-to-meaning correspondences in German. *Frontiers in Psychology, 7*, Article 1200. doi:10.3389/fpsyg.2016.01200

Vinson, D., Jones, M., Sidhu, D. M., Lau-Zhu, A., Santiago, J., & Vigliocco, G. (2021). Iconicity emerges and is maintained in spoken language. *Journal of Experimental Psychology: General, 150*(11), 2293–2308. doi:10.1037/xge0001024

Whissell, C. (1999). Phonosymbolism and the emotional nature of sounds: Evidence of the preferential use of particular phonemes in texts of differing emotional tone. *Perceptual and Motor Skills, 89*, 19–48.

Whissell, C. (2000). Phonoemotional profiling: A description of the emotional flavour of English texts on the basis of the phonemes employed in them. *Perceptual and Motor Skills, 91*, 617–648.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer.

Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic?: Iconicity in English sensory words. *Interaction Studies*, *18*(3), 443–464.

Winter, B., & Perlman, M. (2021). Size sound symbolism in the English lexicon. *Glossa: a journal of general linguistics, 6*(1): 79. 1–13. doi:10.5334/gjgl.1646

Wundt, W. (1904). *Völkerpsychologie: Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythus und Sitte* [Social Psychology: Language] (Vols. 1–2). Leipzig, Germany: W. Engelmann.

4. Poetry study – On the relation between the general affective meaning and the basic sublexical, lexical, and inter-lexical features of poetic texts – A case study using 57 Poems of H. M. Enzensberger

This chapter is published as:

Ullrich, S., Aryani, A., Kraxenberger, M., Jacobs, A. M., & Conrad, M. (2016). On the relation between the general affective meaning and the basic sublexical, lexical, and inter-lexical features of poetic texts – A case study using 57 Poems of H. M. Enzensberger. *Frontiers in Psychology, 7*, Article 2073. doi:10.3389/fpsyg.2016.02073, licensed under CC BY 4.0.

4.1. Abstract

The literary genre of poetry is inherently related to the expression and elicitation of emotion via both content and form. To explore the nature of this affective impact at an extremely basic textual level, we collected ratings on eight different *general affective meaning* scales – valence, arousal, friendliness, sadness, spitefulness, poeticity, onomatopoeia, and liking – for 57 German poems ("*die verteidigung der wölfe*") which the contemporary author H. M. Enzensberger had labeled as either "friendly", "sad", or "spiteful". Following Jakobson's (1960) view on the vivid interplay of hierarchical text levels, we used multiple regression analyses to explore the specific influences of affective features from three different text levels (sublexical, lexical, and inter-lexical) on the perceived *general affective meaning* of the poems using three types of predictors: 1) Lexical predictor variables capturing the mean valence and arousal potential of words; 2) Inter-lexical predictors quantifying peaks, ranges and dynamic changes within the lexical affective content; 3) Sublexical measures of *basic affective tone* according to sound-meaning correspondences at the sublexical level (see Aryani, Kraxenberger, Ullrich, Jacobs, & Conrad, 2016). We find the lexical predictors to account for a major amount of up to 50 % of the variance in affective ratings. Moreover, inter-lexical and sublexical predictors account for a large portion of additional variance in the perceived *general affective meaning*. Together, the affective properties of all used textual features account for 43 to 70 % of the variance in the affective ratings and still for 23 to 48 % of the variance in the more abstract aesthetic ratings. In sum, our approach represents a novel method that successfully relates a prominent part of variance in perceived *general affective meaning* in this corpus of German poems to quantitative estimates of affective properties of textual components at the sublexical, lexical, and inter-lexical level.

Keywords: general affective meaning, inter-lexical measures, phonological iconicity, EMOPHON, basic affective tone, neurocognitive poetics, Enzensberger.

4.2. Introduction

Emotional impact constitutes an important aspect of poetry (Cupchik, 1994; Schrott & Jacobs, 2011; Turner & Poeppel, 1983; van Peer, Hakemulder, & Zyngier, 2007) – people read poems to be amused, pleased, or emotionally and aesthetically moved (Jacobs, 2015a). The underlying affective and aesthetic processes of reading are just beginning to be tackled by research on literature reception. On the one hand, there is a tradition of explaining aesthetic sensation to literature and other works of art by foregrounding effects as deviations from a normative background (Miall & Kuiken, 1994; van Peer, 1986) – focusing mainly on structural and stylistic properties of poetry. On the other hand, the Neurocognitive Poetics Model (NCPM) of Literary Reading (Jacobs, 2011, 2014, 2015a, 2015b) postulates that background elements facilitate emotional involvement in general, for example via mood empathy (Lüdtke, Meyer-Sickendieck, & Jacobs 2014), while foregrounding features promote aesthetic evaluation (Jacobs, Lüdtke, & Meyer-Sickendiek, 2013).

In this study, we will focus on the role of basic textual features in affective poetry reception – investigating how sublexical, lexical, and inter-lexical affective features determine the perception of the *general affective meaning* of a poem.

Narrowing, thus, our research focus on these basic textual elements, our approach deliberately leaves aside – or aims beyond – the influence of higher level variables as, e.g., context information, rhetorical features, or familiarity and comprehensibility of literary texts, as well as the interaction of these higher level variables, on the overall affective perception of art (e.g., Bohrn, Altmann, Lubrich, Menninghaus, & Jacobs, 2013; Leder, Gerger, Dressler, & Schabmann, 2012, Menninghaus et al., 2015).

The *general affective meaning* of a text probably closely relates to global affective appraisals of the reader concerning the overall theme and impression of a text (see also Aryani et al., 2016). In their most basic form, such appraisals should involve the core dimensions of affect: valence and arousal (Bradley, Greenwald, Petry, & Lang, 1992; Russell, 1978, 2003; Watson & Tellegen, 1985; Wundt, 1896). But they can also be captured on more discrete affective/emotional scales, or even for higher-order cognitive-aesthetic concepts, using respective rating scales (see *Methods* section for details). Especially in poetry, aesthetic emotions triggered by textual features and style might crucially add to the overall affective impact – besides immersive emotions arising from the plot.

A theoretical guideline for our approach to explain variance in the perceived *general affective meaning* of literary works by basic textual elements from different hierarchical text levels is provided by Jakobson's postulations about the "Framework of language" (1980a): "Each level above [that of language sounds] brings new particularities of meaning: they change substantially by climbing the ladder which leads from the phoneme to the morpheme and from there to words (with all their grammatical and lexical hierarchy), then go through various levels of syntactic structures to the sentence [...]. Each one of these successive stages is characterized by its clear and specific properties and by its degree of submission to the rules of the code and to the requirements of the context. At the same time, each part participates, to the extent possible, in the meaning of the whole" (1980a, p. 20). Recent brain imaging research reveals close matches between this hierarchy of linguistic structures and the respective hierarchies of brain processes during language processing (Ding et al., 2016).

Although the hierarchical processing of language applies to everyday speech or prose as well, this study focuses on poetry because this genre intertwines content and form in most intimate ways – or, like Jakobson put describing the general "poetic function" of language: "The message focuses on the message for its own sake" (Jakobson, 1960, 1980a).

In the following paragraphs we will introduce empirical evidence for how the affective impact of texts can depend on specific lexical, inter-lexical, and sublexical levels of processing. In the empirical part of this study we will then operationalize affective properties at these three different levels and statistically examine their relation with the perceived *general affective meaning* of poems from a corpus of the German author Hans Magnus Enzensberger.

4.2.1. Lexical effects on general affective meaning

Lexical affective meaning has been shown to be of reliable predictive potential for the affective perception of different types of texts (Anderson & McMaster, 1982; Bestgen, 1994; Hsu, Jacobs, Citron, & Conrad, 2015; Whissell, Fournier, Pelland, Weir, & Makarec, 1986; Whissel, 1994). The importance of lexical affective meaning is increasingly stressed in sentiment analyses of online social media texts (Paltoglou, 2014; Thelwall, Buckley, Paltoglou, Cai, & Kappas, 2010). Valence and arousal ratings of words are most often employed for lexical affective analyses, as most of the variance in a word's meaning on a variety of scales can be accounted for by these two largely independent factors, as has been shown via semantic differential techniques (Osgood & Suci, 1955; Osgood, Suci, & Tannenbaum, 1957; Osgood, May, & Miron, 1975). Furthermore, valence and arousal are also the core dimensions around which several well-established two-dimensional emotion and affect theories are built (e.g., Bradley *et al.*, 1992; Wundt, 1896). Hence, large-scale affective word databases have been gathered to provide normative affective ratings for several thousand words from a given language (English: e.g., ANEW: Bradley & Lang, 1999; DAL: Sweeney & Whissell, 1984, and Whissell, 2009; German: e.g., BAWL: Jacobs et al., 2015; Võ, Jacobs, & Conrad, 2006; Võ et al., 2009; ANGST: Schmidtke, Schröder, Jacobs, & Conrad, 2014; also see Schauenburg, Ambrasat, Schröder, von Scheve, & Conrad, 2014). For examples of usages, see, for instance, Conrad, Recio, & Jacobs, 2011; Hofmann, Kuchinke, Tamm, Võ, & Jacobs, 2009; Hsu, Conrad, & Jacobs, 2014; Hsu, Jacobs, & Conrad, 2015; Hsu, Jacobs, Citron, & Conrad, 2015; Hsu, Jacobs, Altmann, & Conrad, 2015; Kuchinke et al., 2005; Palazova, Mantwill, Sommer, & Schacht, 2011; Recio, Conrad, Hansen, & Jacobs, 2014; Scott, O'Donnell, Leuthold, & Sereno, 2009. However, the *general affective meaning* is, most probably, more than just a direct function of lexical affective values in the text since the processing of affective words is expected to interact with the surrounding sentence context.

4.2.2. Inter-lexical effects on general affective meaning

Lüdtke & Jacobs (2015) show that the succession of two words of similar valence in a sentence can lead to priming effects, with shorter sentence verification times in the case of affectively compatible words – specifically for positive words. In this vein, one might ask what effect a continuous rise or fall of affective lexical values throughout a poem could have on the affective perception by the reader. Furthermore, affective connotations of a single word can dominate the *general affective meaning* of a whole sentence – especially in the case of negative adjectives, which have been shown to exert a negativity bias (Liu, Hu, & Peng, 2013; Lüdtke & Jacobs, 2015). Yet, it remains an open question whether one single word with an extreme affective value could even dominate the affective perception of a whole text paragraph or poem. Moreover, the span width between the two most extreme lexical affective values might also be of relevance for the *general affective meaning*: For example, the arousal span has been shown to account for about 25 % of the variance in suspense ratings in the story "The Sandman" by E. T. A. Hoffmann (Jacobs, 2015a; Lehne, 2014). Furthermore, the arousal span strongly contributes to the perceived arousal level of text passages as well as to the activation of emotion-related brain areas when reading passages from Harry Potter books (Hsu, Jacobs, Citron, & Conrad, 2015).

4.2.3. Sublexical effects on general affective meaning

Poetry inherently involves the structuring of sound, which is why it is important to consider the phonological composition at the sublexical level – also and especially when investigating the emotional impact of poetry. Our study draws on the general theoretical assumption of *phonological iconicity*: Sublexical language sounds have been found to evoke highly consistent assessments of meaning dimensions – potentially relevant for affect – such as size, shape, or pleasantness (Köhler, 1929; Sapir, 1929; Taylor & Taylor, 1965; for a review on the phenomenon of *phonological iconicity* see Schmidtke, Conrad, & Jacobs, 2014). Such findings inspired literature scientists and psychologists to compare the phonetic content of poems of opposite *general affective meaning*. While some of these studies indicated that, for example, plosives appear more often in positive or happy poems, whereas nasals appear rather in sad contexts (Albers, 2008; Auracher, Albers, Zhai, Gareeva, & Stayniychuk, 2011; Wiseman & Peer, 2003), other studies found contradictory evidence, for example, that plosives reflect negative characteristics (Fónagy, 1961; Whissel, 1999, 2000), or that nasal vowels represent beauty (Tsur, 1992). A general problem of these studies is that they were merely investigating the frequency of occurrence of the phonemes of interest, which could be misleading due to specifics of phoneme distributions in the poetic language mode. An alternative approach is to calculate the deviation of existing phonological patterns in a poem from expected standard patterns (Aryani, Jacobs, & Conrad, 2013). This reflects the idea of foregrounding, which is held responsible for the interruption of the automated reading process, thus leading to deeper cognitive processing and potentially aesthetic sensations (Jacobs, 2011, 2015a, 2015b; Miall & Kuiken, 1994; Mukařovský, 1964; van Peer, 1986; van Peer et al., 2007). Aryani et al. (2013) compared the use of phonological units in a poem to the standard distribution of phonological units in prosaic language. This is based on proposed differences between poetic and prosaic language use (see Jakobson's "poetic function" as mentioned above). The resulting deviant phonological units may be responsible – by the foregrounding effects of their salience – for a specific impact of the poem's sound onto the reader. The *basic affective tone* approach of Aryani et al. (2016) further involves intrinsic affective values of the salient phonological segments. This is inspired from the finding that certain phonological clusters tend to occur particularly often in words of specific affective meaning (e.g., high arousal and negative valence). Sublexical affective values were computed averaging the valence and arousal values respectively of all words in which a particular phonological segment occurs in a normative database containing valence and arousal ratings for over 6,000 German words (an extension of Võ et al., 2006 & 2009, by Conrad et al., in preparation) – assuming an internal relation between the signifier and the signified. For the corpus "verteidigung der wölfe", a compendium of 57 German poems by Hans Magnus Enzensberger (1957), Aryani et al. (2016) have investigated the match of the author-given affective chapter labels "friendly", "sad", and “spiteful” with the readers’ affective appraisals of the poems, and connected these comparisons to an analysis of *basic affective tone* at the sublexical phonological level – connecting thus all three parts of Jakobson’s extension of Bühler’s organon model: sender (the author), message (the text), and receiver (the readers) (Bühler, 1934; Jakobson, 1960). They could show how a close match between author labels and readers’ affective appraisals appears to be mediated through a specific use of phonology: the *basic affective tone* (term introduced by Aryani et al., 2016) alone accounted for up to 20% of the variance in readers’ ratings of the *general affective meaning*.

Here, we will extend the analyses of Aryani et al. (2016) on the relation between text and reader for the same corpus of poems to the lexical (referring to the words in a text) and inter-lexical (concerning the relations between words) text levels in order to achieve a more comprehensive understanding of how basic textual elements may determine the affective impact of poetry.

In general, research described above has shown that affective features of different text levels can contribute to the *general affective meaning* of a text. It remains unclear, though, whether such effects of different text levels are independent of each other, and how much of the general affective perception of poetry could be explained via these relatively basic textual dimensions altogether. In the following, we will try to quantify these influences on readers’ affective perception of poetry via multiple regression analyses.

We hypothesize, in particular, that (i) lexical variables will generally be the best predictors of *general affective meanings* as assessed by ratings. Nonetheless, we assume that (ii) affective features at all text levels significantly contribute to the perceived *general affective meaning* of poems. Partialling out the influence of lexical variables via multiple regression should reveal independent sublexical and inter-lexical contributions to the affective impact of poetry.

Consciously leaving aside important supra-lexical features such as familiarity with a literary genre (Bohrn et al., 2013), comprehensibility (Leder et al., 2012), experience with literary work in general (Winston & Cupchik, 1992), and many other personality variables (Bleich, 1978), as well as syntactic and structural characteristics of the poems, we search to estimate how much affective potential resides already within more basic constituents of the text itself: single phonemes, words, or basic inter-relations between words.

4.3. Materials and methods

4.3.1. Ratings

Poem corpus

"*die verteidigung der wölfe*" ("the defense of the wolves") was written in 1957 by the contemporary German author Hans Magnus Enzensberger (*1929, see Astley, 2006, for an English introduction to Enzensberger's poetry). These 57 poems are partitioned by the author into three chapters of 21 "sad" (*"traurig"*), 19 "friendly" (*"freundlich"*), and 17 "spiteful" (*"böse"*) poems. This assures a sufficient variety of affectivity across all poems, paving the way for a differentiated prediction of the variance in the perception of their *general affective meaning* via affective information at subjacent text levels. An advantage of this contemporary poem volume is the employed free verse poetry which should prevent our operationalization of phonological salience to be confounded with features of a strong metrical ordering and rhyme that also exert a specific influence on aesthetic judgments of poems (Menninghaus, Bohrn, Altmann, Lubrich, & Jacobs, 2014, Menninghaus et al., 2015; Obermeier et al., 2013).

Participants

German native speakers were recruited through a post on the institute's website and a diversity of Facebook webpages. More than 300 people participated, 252 of which left evaluable data (173 females; age range from 17 to 76, $M = 35.9$, $SD = 12.1$).

Procedure and Variables

General affective meaning ratings were acquired via an online survey using the QuestBack Unipark software. After being welcomed, instructed, and asked to enter a few personal data (age, sex, native language, profession), people were free to read and rate as many poems as they liked ($M = 4.3$, $SD = 5.4$). The poems were presented in a pseudo-randomized order. People were asked whether they already knew each poem – only unknown poems were used in the analyses. A minimum of 15 complete ratings for each poem were acquired on each of the following eight dimensions – presented to participants in randomized order:

Ratings of valence and arousal – linking our approach to psychological emotion models:

- *Valence* (1) on a 7-point scale ranging from -3 (very negative) via 0 (neutral) to 3 (positive)
- *Arousal* (2) on a 5-point scale ranging from 1 (very calming) to 5 (very arousing), also using the SAM scale mannequins of Bradley & Lang (1999)

Ratings on discrete affective categories to directly assess the labels the author suggested for his poems:

- *Friendliness* (3) on a 5-point scale ranging from 1 (not friendly at all) to 5 (very friendly)
- *Sadness* (4) on a 5-point scale ranging from 1 (not sad at all) to 5 (very sad)
- *Spitefulness* (5) on a 5-point scale ranging from 1 (not spiteful at all) to 5 (very spiteful)

The basic two levels of our approach toward capturing the perception of general affective meaning in poetic texts – dimensional and discrete aspects of emotion – are derived from the dual-process model of emotional responses to art of Cupchik & Winston (1992; also see Cupchik, 1994). While arousal and valence ratings form the reactive part of their model, the discrete affective dimensions – which require more context information (i.e., appraisals) – form the reflective part of the model. In poetry, though, immersive emotions – arising from the plot – may be less dominant than in narrative fiction (Jacobs, 2011, 2015a; Mar, Oatley, Djikic, & Mullin, 2011; Oatley, 1994), whereas aesthetic emotions – characterized mainly by the affective evaluation and appreciation of artistic style, beauty, etc. – play a more dominant role (Frijda, 1989; Leder, Belke, Oeberst, & Augustin, 2004; Marković, 2012). Hence, we extend the model of emotional responses to art by a meta-reflective layer comprising aesthetic concepts: A *liking* rating is assumed to assess the affective part of aesthetic judgment, referring to personal emotional experiences during poetry reception, whereas we assume *poeticity* ratings to capture the more cognitive aspects of aesthetic judgment, as they have been shown to be influenced by linguistic competence in general (Hoffstaedter, 1987). Such aesthetic preferences represent a much more abstract level of the perception of general affective meaning, as they strongly depend on context and personality factors as well (Bleich, 1978; Jacobs, 2011, 2015a). As our study also refers to the phenomena of phonological salience and *basic affective tone* at the sublexical level (see Aryani et al., 2016, and further below), we additionally collected *onomatopoeia* ratings. Onomatopoeia represents the use of words whose sound is suggestive of their meaning. Hence, this rating is supposed to assess how well the (imagined) sound of a poem is perceived by the reader to fit the overall meaning of a poem – as „poetry is a province where the internal nexus between sound and meaning changes from latent into patent and manifests itself most palpably and intensely" (Jakobson, 1960).

Ratings on aesthetic evaluations as well as the fit of sound and meaning:

- *Liking* (6) on a 5-point scale ranging from 1 (not at all) to 5 (very much)
- *Poeticity* (7) on a 5-point scale ranging from 1 (not poetic at all) to 5 (very poetic)
- *Onomatopoeia* (8) on a 5-point scale ranging from 1 (not onomatopoetic at all) to 5 (very onomatopoetic)

4.3.2. Multiple Regression

The rating variables (including the absolute value of valence) were used as dependent variables in a multiple regression approach. To provide a most extensive screening for potential effects of different phenomena at different text levels we included a considerable number of predictor variables (55 in total, listed in Table S1 in the supplementary materials) from the three basic levels of text processing into the forward stepwise multiple regression models. As stop criterion we used the rather conservative Bayesian information criterion (BIC), which seems an appropriate way to constrain the number of significant results – putting specific effort in avoiding false positives – for an a priori high number of predictors.

We assume deviation to be an important precursor of foregrounding, which is supposed to be responsible for many affective and aesthetic effects while reading literary works. Hence, we tried to operationalize the degree of deviation from expected values at each text level: instead of using raw mean values of, e.g., valence or arousal values of words or subsyllabic segments (see Aryani et al., 2016) in a poem, we rather used the degree of deviation of these values from neutral global means within a representative database of everyday German language (Brysbaert et al., 2011) to predict readers' affective perception of the poems (see sections below for calculations).

To capture potential quadratic effects of variables with potentially bipolar character we used both standard values – including positive/negative algebraic signs – and their absolute values as predictors of ratings. This should enable us to capture effects such as, e.g., arousal ratings increasing with both more negative and more positive lexical content of poems (as compared to neutral valence) – or negative or positive deviations from neutral at the sublexical level, respectively.

Lexical predictors

All poem texts were PoS (part-of-speech) tagged to identify the word forms and infinitives of each word. Function words were omitted for their use is determined mainly by grammatical requirements (Anderson & McMaster, 1982; Miller, 1954). For the remaining words lexical valence and arousal values were looked up in an extended version of the BAWL database, containing more than 6,000 German words with affective ratings (Võ et al., 2006, 2009; publication of the extended version *in prep.*). For poem words that did not appear in the database but were standard German words, we collected additional valence and arousal ratings. Finally, this yielded a 90 % matching rate. The missing 10 % mainly consist of names of people and places, and a few foreign or obsolete words.We determined the lexical affective predictors – in terms of valence and arousal – for a poem by their deviations from the respective affective mean of a null model. That is, we calculated the extent to which the mean of valence and arousal values of the words in a poem deviates from the affective mean of the same number of words randomly pulled from a linguistic corpus. For this, the valence and arousal values of the words in the BAWL database were first z-standardized according to their lexical frequency in the SUBTLEX-DE corpus (Brysbaert et al., 2011) resulting in a normal distribution with a mean of zero and a standard deviation of one. In order to calculate the standard deviation of a randomly pulled sample – given statistical independence of the values in each sample – the standard deviation of the whole words in the database (i.e., one) was divided by the square root of the size of the random sample (i.e., the total number of words appearing in the corresponding poem). This value represents the standard deviation of the null model. By dividing the mean of valence or arousal values of all words in a poem by the standard deviation of its corresponding null model, we obtained a "sigma factor" which indicates the extent to which the valence or arousal mean deviates from an expected value (i.e., the null model). As example, the formula for the sigma factor of lexical arousal looks as follows:

$$Sigma(arousal) = \frac{M(arousal)}{1\sqrt{\mathrm{N}}}$$

Where *M(arousal)* is the mean of arousal values of all words in a poem, and *N* is the total number of words appearing in the corresponding poem.

- These sigma factors (together with their resulting absolute values) for lexical valence and arousal of each poem served as predictors in the multiple regression models.

<u>Inter-lexical predictors</u>

The inter-lexical variables that we included in the analysis are thought to reflect tensions and dynamics within a text. Here as well, deviation matters: Standard deviations and spans of all words' valence and arousal values may serve as a proxy for the general affective spread of a poem:

- Standard deviations of lexical valence and arousal in each poem
- Valence and arousal spans (difference between lexical maximum and minimum value of valence/arousal in a poem)

As already one single affectively deviant word can dominate the *general affective meaning* of a whole poem, also affective minima and maxima are included in the range of inter-lexical predictor variables:

– Minimum and maximum values of lexical valence and arousal per poem

Correlations between words' positions in the text and their affective values might reflect the development of affectivity throughout the course of a poem:

– Correlation coefficients (together with their absolute values) between words' positions (beginning to end) within a poem and arousal, valence, and the absolute valence values

In addition, the number of words per poem was also included as a predictor variable.

<u>Sublexical predictors</u>

The EMOPHON tool (Aryani et al., 2013) translates an input text into its phonemic notation and then analyses the text for salient phonological units based on a probabilistic model: a reference linguistic corpus for the German language (SUBTLEX-DE; Brysbaert et al., 2011) determines confidence intervals for the frequency of occurrence of all sublexical units in a text depending on its length. If the actual frequency of occurrence of a specific sublexical unit in the text exceeds its confidence interval, it is regarded as salient. Here, we chose the tool's option to segment the texts into the subsyllabic units onsets, nuclei, and codas (instead of single phonemes) – which are used for all following analyses. For the number of salient phonological segments that exceed their confidence intervals, we used the

– *N*s of salient onsets, *N*s of salient nuclei, *N*s of salient codas as well as the *N*s of all salient subsyllabic segments altogether (in each case weighted by the length of the respective poem)

The recent update of EMOPHON (Aryani et al., 2016) further provides a quantitative measure for the *basic affective tone* by integrating the detected salient phonological segments with affective values assigned to each of them. These sublexical affective values for subsyllabic onsets, nuclei and codas were computed by averaging lexical valence and arousal values of all words in a lexical database containing rating values for over 6,000 words a specific phonological segment appears in (see Conrad et al, in preparation, for details). Again, we used the sigma factors – the extent to which the respective mean affective value of salient phonological segments in a poem deviates from the mean of the random distribution used in the model – as predictor variables. The sigma factor reflects how strongly the affective sublexical value of the poem deviates from an expected value for a text of the same length – and into which direction. As predictors, we used:

- The sigma factors (together with their absolute values) for valence and arousal of *salient* onsets, nuclei and codas, as well as for all salient syllabic segments altogether
- And for control reasons, the sigma factors for valence and arousal of *all* phonological segments (or one type of segments) in the text, no matter if they are salient or not (see Aryani et al., 2016)

Note that the present analyses go beyond the ones presented in Aryani et al. (2016) by also addressing:

- The role of the general degree of phonological salience, i.e., the number of salient segments (ignoring their affective values) in a poem
- Potential specific effects of specific types of subsyllabic units, i.e., onsets, nuclei or codas

Furthermore, by letting sublexical and lexical values compete in multiple regression analyses we want to provide answers to the following research questions:

- Can the *basic affective tone* be seen as completely independent from the lexical inventory of the poems, and therefore be purely attributed to *phonological iconicity*?
- How strong will effects of the *basic affective tone* still appear besides – presumably dominant – general lexical effects?

4.4. Results

4.4.1. Descriptive Statistics of the Rating Variables

Means and spreads of rating variables are summarized in Table 4.1. The mean ratings for the discrete affective concepts friendliness, sadness, and spitefulness are highest in the respective poem categories. This shows that readers' perceptions of the poems' *general affective meaning* generally correspond well with the author-given affective categorization. Valence and arousal ratings further support the character of these discrete affective categories. For example, the author-defined spiteful poems show the strongest negative valence and highest arousal ratings (for statistical comparisons see Aryani et al., 2016).

Table 4.1: Means (*M*) and standard deviations (*SD*) of the rating variables for each author-given affective category (*N* being the number of poems rated).

	Discrete affective concept categories					
	Friendly (*N*=351)		**Spiteful** (*N*=325)		**Sad** (*N*=400)	
Ratings	*M*	*SD*	*M*	*SD*	*M*	*SD*
Liking	2.98	1.14	2.68	1.21	2.77	1.22
Poeticity	2.91	1.13	2.56	1.15	2.93	1.17
Onomatopoeia	2.52	1.14	2.37	1.10	2.51	1.12
Valence	-0.32	1.45	-1.40	1.13	-1.07	1.23
Arousal	3.21	0.74	3.65	0.59	3.42	0.68
Friendliness	2.09	1.10	1.39	0.67	1.56	0.84
Spitefulness	1.88	1.05	2.72	1.19	2.13	1.13
Sadness	1.98	1.06	2.49	1.16	2.63	1.11

Bivariate correlation coefficients between all dependent variables are shown in Table 4.2. Especially valence is highly correlated with friendliness in a positive way, and with spitefulness in a negative way. An opposite pattern is found for the correlation between arousal and spitefulness (positive) compared to friendliness (negative). Whereas liking and poeticity correlate moderately with each other as well as with valence and friendliness, onomatopoeia and sadness are the two ratings correlating least with the other ones.

Table 4.2: Bivariate correlation coefficients between all rating variables.

	Liking	Poeticity	Onoma-topoeia	Valence	|Valence|	Arousal	Friendli-ness	Spiteful-ness	Sadness
Liking	1	0.67***	0.29*	0.59***	-0.56***	-0.17	0.52***	-0.46***	-0.12
Poeticity		1	0.53***	0.61***	-0.41**	-0.34*	0.56***	-0.47***	-0.2
Onomatopoeia			1	0.39**	-0.35**	-0.16	0.32*	-0.17	-0.38**
Valence				1	-0.8***	-0.7***	0.89***	-0.77***	-0.62***
|Valence|					1	0.58***	-0.63***	0.71***	0.46***
Arousal						1	-0.73***	0.73***	0.38**
Friendliness							1	-0.75***	-0.45***
Spitefulness								1	0.33*
Sadness									1

Indicators of significance: *** $p \leq 0.001$, ** $0.01 \geq p > 0.001$, * $0.05 \geq p > 0.01$

4.4.1. Multiple regression results

Results for forward stepwise regression models using all predictors on different rating dimensions of the *general affective meaning* as dependent variables are summarized in Tables 4.3 to 4.5. Figures 4.1 to 4.9 display bivariate correlations between all rating dimensions and up to four significant predictors emerging from the Multiple Regression Models.

Significant predictors of ratings on the basic affective dimensions valence (including also the absolute values of valence) and arousal are shown in Table 4.3: Around half of the variance

Table 4.3: Predictors of the basic affective dimensions' ratings.

Step	Predictor variable	bivariate correlation coefficient	partial correlation coefficient	t value	cumulated R^2 corrected
"Valence" as dependent variable					
1	Lexical valence	0.65^{***}	0.47^{***}	3.77	0.41
2	Lexical arousal	-0.58^{***}	-0.39^{**}	-3.02	0.47
3	Sublexical arousal all salient segments	-0.44^{***}	-0.47^{***}	-3.76	0.5
4	|Sublexical arousal all salient segments|	-0.27^{*}	0.39^{**}	3.05	0.55
5	Total number salient segments	0.09	0.32^{*}	2.42	**0.59**
"|Valence|" as dependent variable					
1	Lexical arousal	0.63^{***}	0.6^{***}	5.46	0.39
2	Sublexical arousal all salient segments	0.51^{***}	0.46^{***}	3.85	**0.51**
"Arousal" as dependent variable					
1	Maximum lexical arousal	0.57^{***}	0.46^{***}	3.73	0.32
2	Lexical arousal	0.54^{***}	0.32^{*}	2.43	0.38
3	Correlation valence with word position	-0.13	-0.31^{*}	-2.34	**0.43**

Abbreviations: |...| = absolute value;

Indicators of significance: *** $p \leq 0.001$, ** $0.01 \geq p > 0.001$, * $0.05 \geq p > 0.01$;

Color coding: Red: lexical variables, Blue: inter-lexical variables, Yellow: sublexical variables.

The bold number indicates the respective overall cumulative R^2 corrected for each regression model.

(43 % – 59 %) of the ratings of the basic affective dimensions valence and arousal can be explained solely by the employed lexical, sublexical, and inter-lexical affective measures.

The variance in the *valence* ratings can be predominantly accounted for by the lexical valence and arousal patterns in the poems – 41 % of the ratings' variance can already be explained by lexical valence alone. Higher valence ratings go along with increasing lexical valence values but decreasing lexical arousal values (see Fig. 4.1A and 4.1B). This pattern is in line with the general negative correlation between the two affective dimensions in, for example, words from German affective word databases (BAWL: Võ et al., 2009; ANGST: Schmidtke et al., 2014a). At the sublexical level, more salient segments in general also lead to higher valence ratings (Fig. 4.1D). Regarding the sublexical arousal level of all these salient segments, the inclusion of the absolute arousal values allows a more detailed characterization of the underlying mechanisms: Generally speaking, low sublexical arousal leads to higher valence ratings (Fig. 4.1C). This is confirmed by the absolute values of sublexical arousal of all salient segments showing a positive partial correlation. For high-arousing sublexical segments however, the two variables would predict opposite patterns which cancel out each other. Hence, although very low/calming sublexical arousal values lead to higher valence ratings, very high sublexical arousal values do not coincide with more negative supra-lexical valence ratings.

The *absolute values of valence* ratings, representing the intensity of valence ratings irrespective of their direction, can best be predicted by lexical arousal and the sublexical arousal values of all salient segments. At both text levels, higher arousal leads to a higher intensity of valence ratings (see Fig. 4.2). This clearly reflects the U-shaped distribution of lexical valence and arousal values in affective word databases (BAWL: Võ et al., 2009, ANEW: Bradley & Lang, 1999), where both the arousal levels of positive and negative words are higher than for neutral words, even if the arousal for positive words does not reach the same height as for negative words in the German language (Schmidtke et al., 2014a). The fact that sublexical arousal adds another 12 % explanation of variance, hints toward a similar distribution of valence and arousal at the sublexical level.

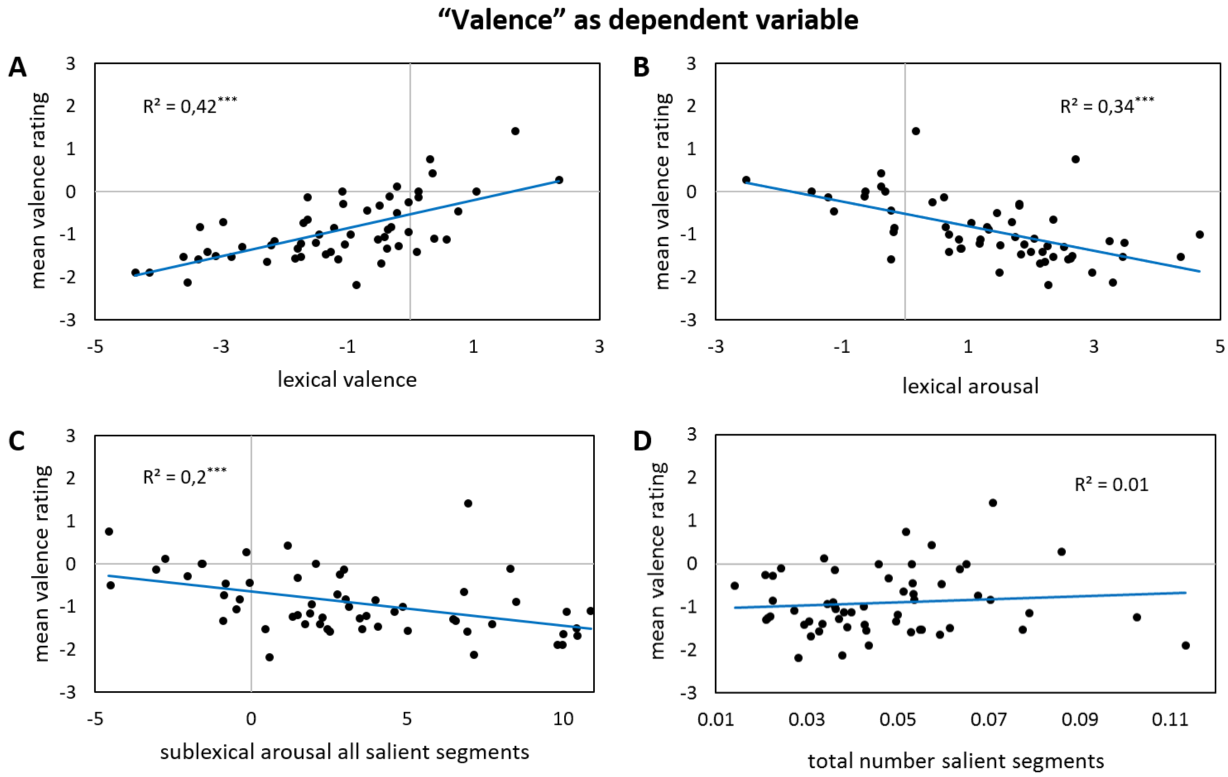

Figure 4.1: Bivariate correlations between valence ratings and four predictors: the sigma factors for lexical valence (A), the sigma factors for lexical arousal (B), the sigma factors for sublexical arousal of all salient segments (C), and the total number of salient segments per poem weighted by its length – note that the correlation gets significant after partialling out the influence of the other predictors (D).

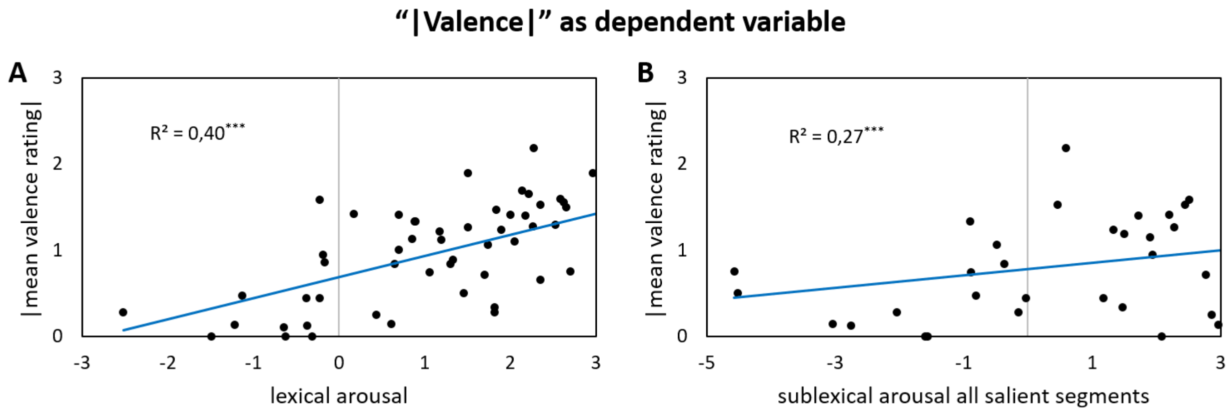

Figure 4.2: Bivariate correlations between the absolute values of valence ratings and two predictors: the sigma factors for lexical arousal (A) and the sigma factors for sublexical arousal of all salient segments (B).

For the *arousal* ratings, no sublexical affective values appear in the regression model. The variance in these ratings is mainly accounted for by the word with the highest arousal level in the poem, but also by the overall lexical arousal level: the higher the lexical arousal and its maximum value, the higher the arousal ratings (Fig. 4.3A and 4.3B). But also a changing level of lexical valence throughout the poem has an influence on the perceived arousal: a rise of valence values toward the end of the poem leads to diminished arousal ratings – marked by a negative partial correlation of the supra-lexical arousal and the correlation values of lexical valence with the word order – whereas a decline of the words' valence throughout the poem leads to a higher perceived arousal level (Fig. 4.3C).

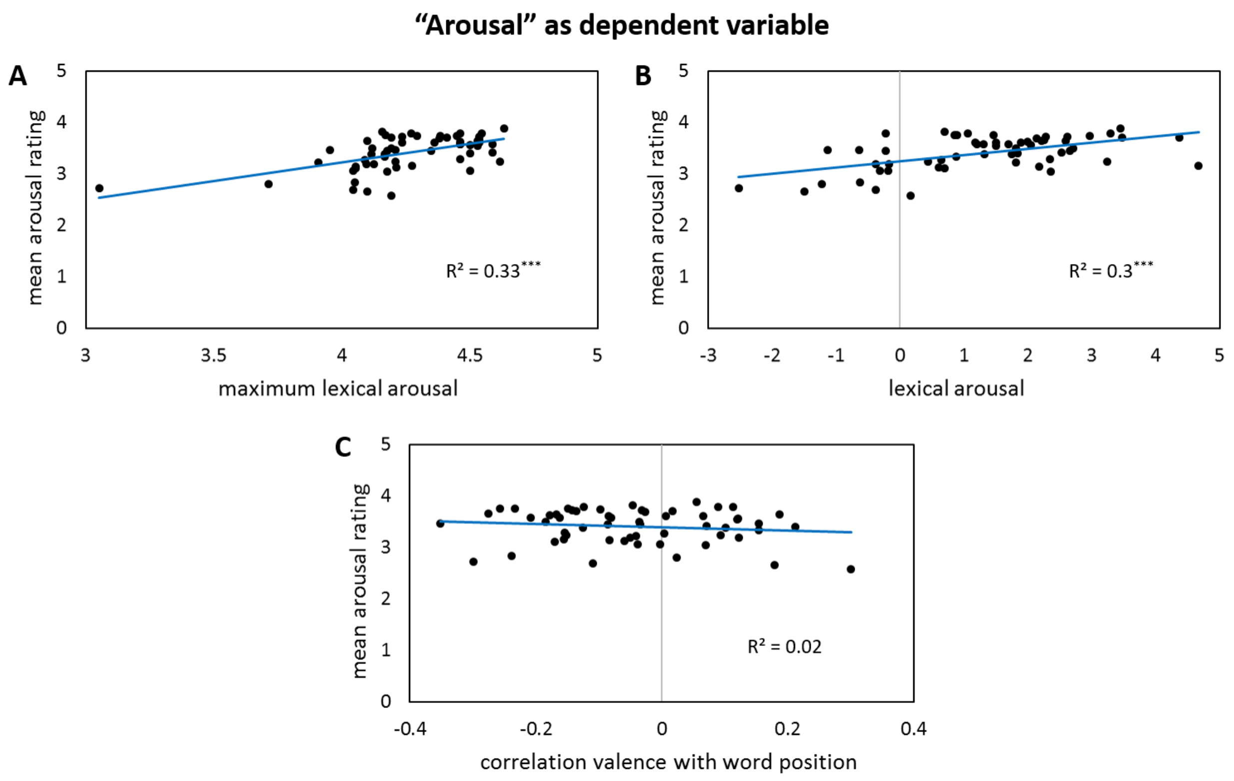

Figure 4.3: Bivariate correlations between arousal ratings and three predictors: the maximum values of lexical arousal per poem (A), the sigma factors for lexical arousal (B), and the correlation coefficients between lexical valence values and word positions in a poem – note that the correlation gets significant after partialling out the influence of the other predictors (C).

Table 4.4: Predictors of the discrete affective concepts' ratings.

Step	Predictor variable	bivariate correlation coefficient	partial correlation coefficient	t value	cumulated R^2 corrected
"Friendliness" as dependent variable					
1	Lexical valence	0.66***	0.57***	4.87	0.43
2	Lexical arousal	-0.61***	-0.43**	-3.41	0.51
3	Sublexical arousal salient nuclei	-0.39**	-0.57***	-4.93	0.55
4	|Sublexical arousal salient nuclei|	-0.16	0.45***	3.53	0.62
5	|Sublexical valence all nuclei|	0.13	0.41**	3.14	0.67
6	Correlation valence with word position	0.15	0.32*	2.37	**0.7**
"Spitefulness" as dependent variable					
1	Lexical valence	-0.64***	-0.46***	-3.73	0.39
2	|Sublexical arousal all salient segments|	0.45***	0.31*	2.33	0.46
3	Total number words	0.31*	0.30*	2.28	0.50
4	Lexical arousal	0.55***	0.29*	2.14	**0.53**
"Sadness" as dependent variable					
1	Minimum lexical valence	-0.47***	-0.56***	-4.63	0.21
2	|Sublexical arousal all codas|	0.25⁺	0.40**	3.05	0.26
3	|Sublexical valence all salient segments|	0.17	0.35*	2.6	0.31
4	Correlation arousal with word position	-0.13	-0.44**	-3.36	0.34
5	Correlation valence with word position	-0.04	-0.36*	-2.64	0.37
6	Total number salient nuclei	-0.19	-0.35*	-2.61	0.40
7	Sublexical arousal all salient segments	0.31*	0.28*	2.02	0.44
8	|Correlation arousal with word position|	0.07	0.27⁺	1.96	**0.47**

Abbreviations: |...| = absolute value;

Indicators of significance: *** $p \leq 0.001$, ** $0.01 \geq p > 0.001$, * $0.05 \geq p > 0.01$, ⁺ $0.1 \geq p > 0.05$;

Color coding: Red: lexical variables, Blue: inter-lexical variables, Yellow: sublexical variables.

The bold number indicates the respective overall cumulative R^2 corrected for each regression model.

Table 4.4 lists significant predictors of discrete affective concepts' ratings: Very similar to the valence ratings, also *friendliness* is mainly driven by the positive influence of lexical valence (Fig. 4.4A) and the negative influence of lexical arousal (Fig. 4.4B), together accounting already for 51 % of the variance in the friendliness ratings. Additionally, if the valence values of words rise with their position in the poem, higher friendliness ratings occur – and vice versa (Fig. 4.4D).

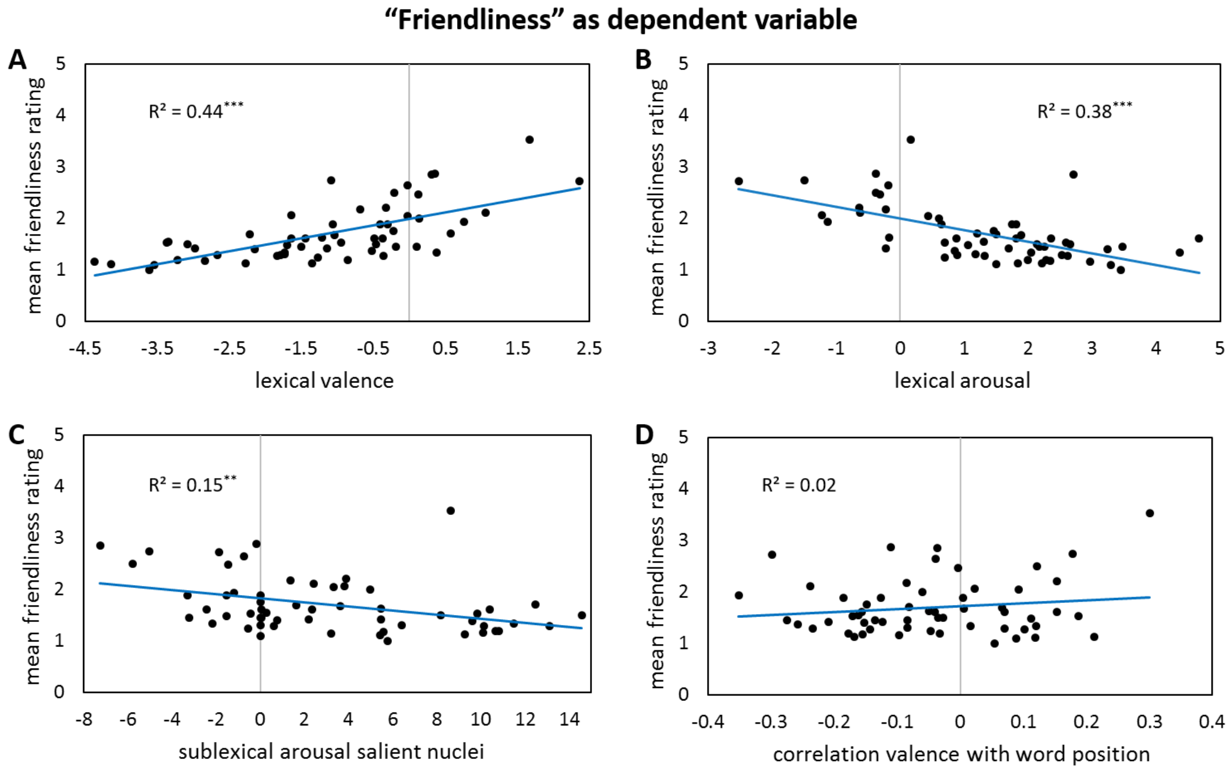

Figure 4.4: Bivariate correlations between friendliness ratings and four predictors: the sigma factors for lexical valence (A), the sigma factors for lexical arousal (B), the sigma factors for sublexical arousal of all salient nuclei (C), and the correlation coefficients between lexical valence values and word positions in a poem – note that the correlation gets significant after partialling out the influence of the other predictors (D).

At the sublexical arousal level, we find the same pattern as for the valence ratings, just that here the one-sided effect of very low arousal values leading to higher friendliness ratings – whereas high arousal does not lead to diminished friendliness – stems from salient nuclei only (Fig. 4.4C), not from all types of salient segments. Furthermore, a higher intensity of the sublexical valence values of all nuclei in the text – being represented by the absolute value – seems to lead to higher friendliness ratings.

Contrary to friendliness, in the *spitefulness* model, lower lexical valence and higher lexical arousal lead to higher spitefulness ratings (Fig. 4.5A and 4.5D), as could be expected for high arousing negative poems such as spiteful ones. At the sublexical arousal level, the more

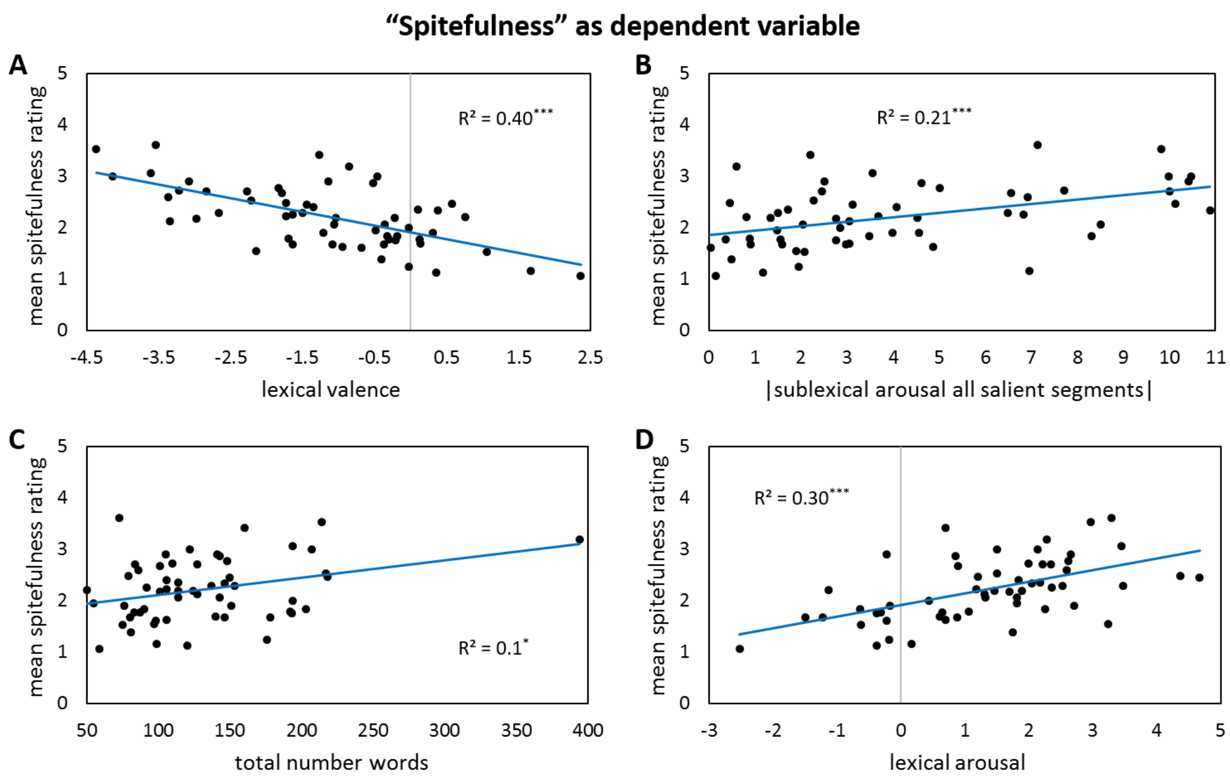

Figure 4.5: Bivariate correlations between spitefulness ratings and four predictors: the sigma factors for lexical valence (A), the absolute values of the sigma factors for sublexical arousal of all salient segments (B), the total numbers of words in a poem (C), and the sigma factors for lexical arousal (D).

extreme the arousal of all salient segments is, regardless in which direction, the higher are the spitefulness ratings (Fig. 4.5B). Also the total number of words seems to play a role here, with longer poems being rated as slightly more spiteful than shorter ones (Fig. 4.5C). This, however, might be a specific quality of this particular poem corpus, not being transferable into general.

A first glance at the regression model for *sadness* shows that, unlike in every other of the analyzed models, neither lexical valence nor arousal per se is included. However, the word with the smallest valence value in a poem is the most influential predictor of the sadness ratings – the smaller its valence is, the sadder is the overall impression of the poem (Fig. 4.6A). Another important inter-lexical aspect in the case of sadness is the correlation of word affectivity with the word order. For lexical valence, more negative word values toward the end of a poem raise the sadness rating, and more positive values toward the poem's end make it less sad (Fig. 4.6D). In the case of lexical arousal, declining word arousal values throughout the poem account for a sad poem (Fig. 4.6C), but rising arousal levels to the end of a poem do not necessarily lead to smaller sadness ratings. This is reflected by the absolute value of the correlation of lexical arousal with the words' positions entering the regression model as well, which neutralizes the potential influence of higher lexical arousal values. At the sublexical level, the absolute value of the arousal level of all codas in the text seems to be the strongest predictor. Thus, any coda's arousal value being significantly different from the distribution's mean – no matter whether it is especially low or high-arousing – leads to higher sadness ratings. The same holds for the valence values of all types of salient segments in the poems. Regarding their arousal level, the higher it is, the sadder the poem is perceived (Fig. 4.6B). Furthermore, the occurrence of many salient nuclei in a text goes along with lower sadness rating.

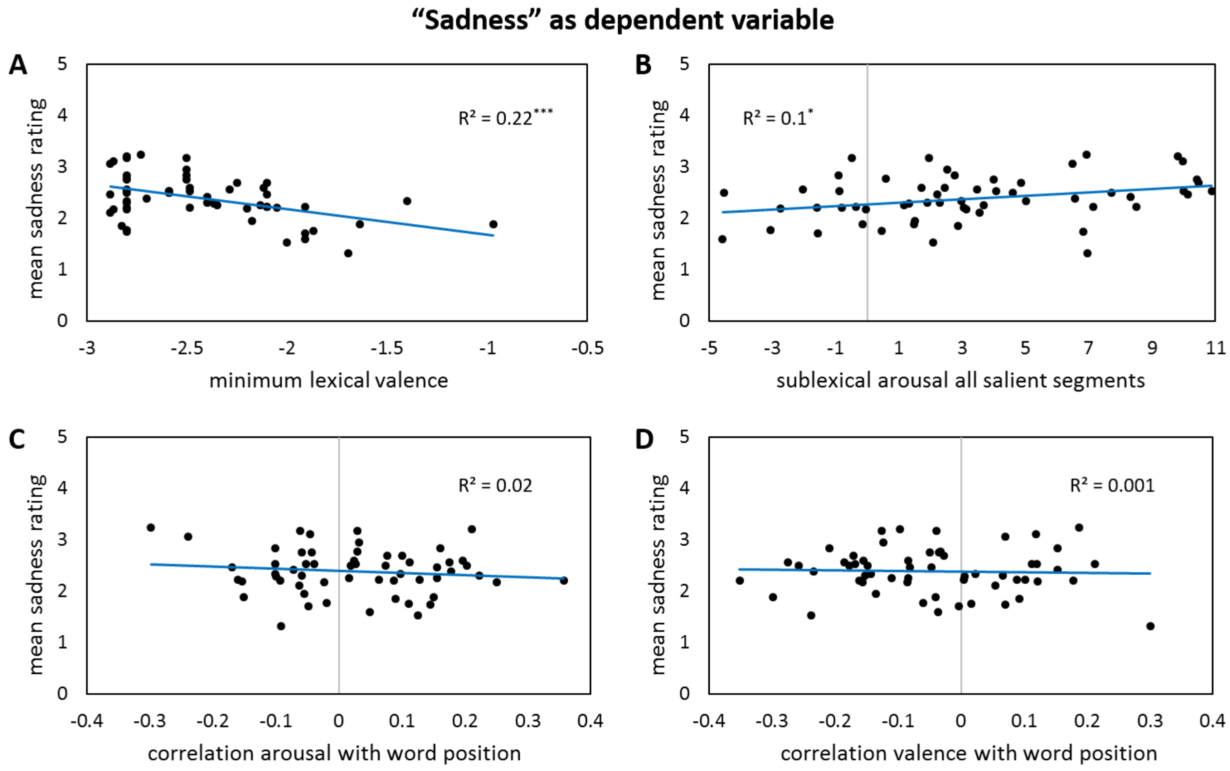

Figure 4.6: Bivariate correlations between sadness ratings and four predictors: the minimum values of lexical valence per poem (A), the sigma factors for sublexical arousal of all salient segments (B), the correlation coefficients between lexical arousal values and word positions in a poem – note that the correlation gets significant after partialling out the influence of the other predictors (C), and the correlation coefficients between lexical valence values and word positions in a poem – again, note that the correlation gets significant after partialling out the influence of the other predictors (D).

Table 4.5 lists significant predictors of the two aesthetic as well as the onomatopoeia ratings: The only two variables that significantly predict part of the *liking* ratings (23 %) are lexical arousal and the sublexical arousal of all salient segments. Both types of arousal show a negative partial correlation with the dependent variable: poems appear to be “liked” less when containing words of relatively high arousal, but more when they are low-arousing (Fig. 4.7A). The same holds for the arousal potential of salient phonological segments (Fig. 4.7B).

Table 4.5: Predictors of the two aesthetic and the onomatopoeia ratings.

Step	Predictor variable	bivariate correlation coefficient	partial correlation coefficient	t value	cumulated R^2 corrected
"Liking" as dependent variable					
1	Lexical arousal	-0.43***	-0.37**	-2.97	0.17
2	Sublexical arousal all salient segments	-0.37**	-0.29*	-2.23	**0.23**
"Poeticitiy" as dependent variable					
1	Lexical arousal	-0.42**	0.17	1.2	0.16
2	Total number salient segments	0.29*	0.26⁺	1.93	0.23
3	|Lexical arousal|	-0.4**	-0.46***	-3.67	0.27
4	Sublexical arousal salient nuclei	-0.24⁺	-0.39**	-3.03	0.32
5	|Sublexical arousal salient nuclei|	-0.12	0.3*	2.21	0.35
6	Total number salient codas	0.28*	0.29*	2.11	**0.39**
"Onomatopoeia" as dependent variable					
1	SD lexical valence	-0.43***	-0.33*	-2.47	0.17
2	Total number salient nuclei	0.34*	0.38**	2.88	0.22
3	|Sublexical arousal all nuclei|	-0.25⁺	-0.49***	-3.95	0.28
4	Lexical valence	0.01	-0.52***	-4.26	0.32
5	SD lexical arousal	-0.39**	-0.43**	-3.35	0.40
6	Sublexical valence salient codas	-0.04	0.31*	2.27	0.45
7	Maximum lexical valence	0.07	0.28*	2.02	**0.48**

Abbreviations: SD = standard deviation, |…| = absolute value;

Indicators of significance: *** $p \leq 0.001$, ** $0.01 \geq p > 0.001$, * $0.05 \geq p > 0.01$, ⁺ $0.1 \geq p > 0.05$;

Color coding: Red: lexical variables, Blue: inter-lexical variables, Yellow: sublexical variables.

The bold number indicates the respective overall cumulative R^2 corrected for each regression model.

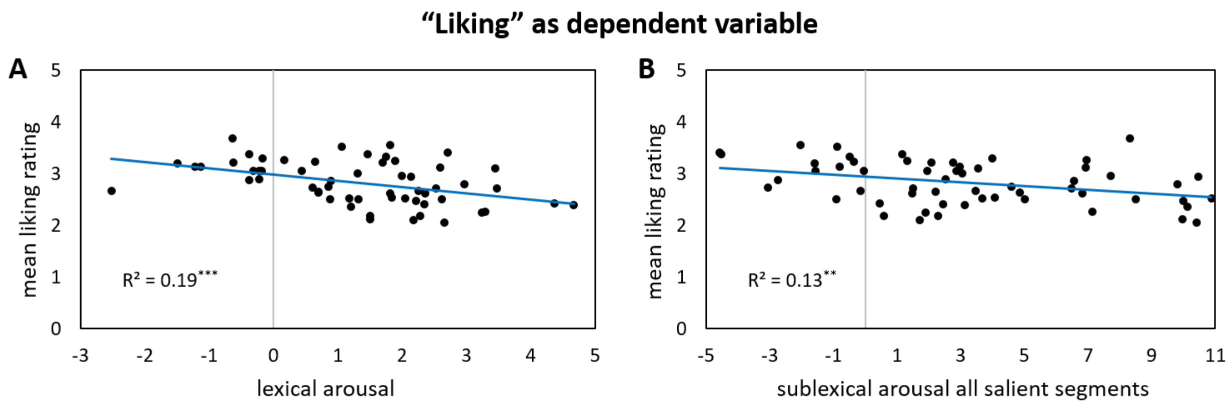

Figure 4.7: Bivariate correlations between liking ratings and two predictors: the sigma factors for lexical arousal (A) and the sigma factors for sublexical arousal of all salient segments (B).

For the dependent variable *poeticity*, lexical arousal appears as a highly significant predictor variable if its absolute values are considered: The more deviant the lexical arousal values are from zero, no matter whether into a higher arousing or more calming direction, the less poetic the poem is rated (Fig. 4.8A and 4.8B). Thus, poems that contain predominantly words of a rather unremarkable arousal – not significantly high- or low-arousing – are perceived as more poetic than poems with salient lexical arousal features. Moreover, the poeticity ratings are also strongly influenced by sublexical affective values. The number of salient segments, in particular of the salient codas, accounts for a reasonable part (> 10 %) of the ratings’ variance: poems that use phonological segments more often than expected from everyday language are perceived as more poetic (Fig. 4.8C). Furthermore, the arousal level of respective salient nuclei seems to play a differentiated role, as specifically the low-arousing salient nuclei lead to a higher perceived poeticity (Fig. 4.8D). This results from the finding that the continuous arousal values of the nuclei are negatively correlated with the poeticity ratings, while the absolute arousal values correlate in a positive manner. Thus, for the negative range – namely the low-arousing part – the inferred statement is the same, whereas in the positive – high-arousing –

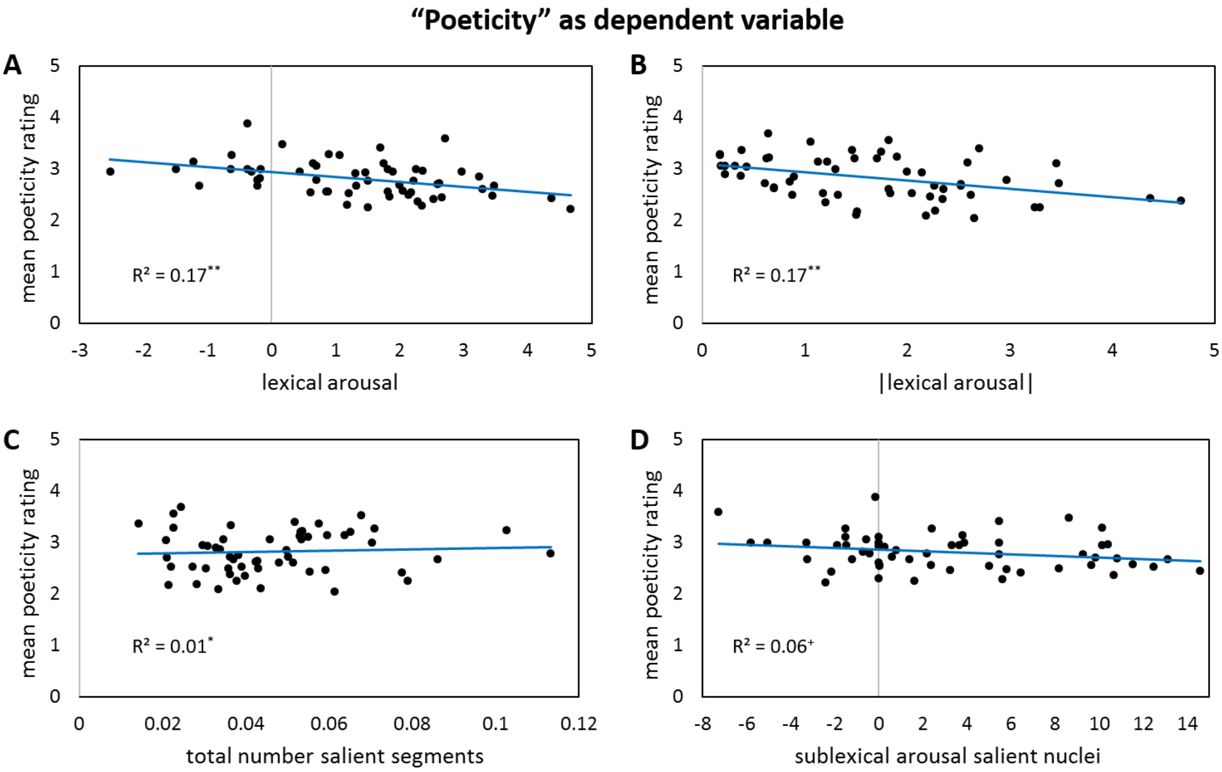

Figure 4.8: Bivariate correlations between poeticity and four predictors: the sigma factors for lexical arousal (A), the absolute values of the sigma factors for lexical arousal (B), the number of all salient segments per poem weighted by its lengths (C), and the sigma factors for sublexical arousal of all salient nuclei – note that the correlation gets significant after partialling out the influence of the other predictors (D).

range the correlation patterns oppose and hence zero out each other. Consequently, more arousing nuclei values do not necessarily lead to diminished poeticity ratings.

The *onomatopoetic* perception is significantly influenced by variables from all three text levels. At the lexical level, a higher occurrence of negatively valenced words in a poem leads to increased onomatopoeia ratings. In contrast, with a higher maximum value of lexical valence in a poem, the ratings for onomatopoeia become slightly higher as well. However, this partial correlation is not a very strong one. Regarding the spread of lexical valence and arousal in each poem – depicted by their standard deviations – higher deviations involve lower onomatopoeia ratings (Fig. 4.9A and 4.9B). At the sublexical level, the nuclei seem to play an important role: On the one hand, a high number of salient nuclei in a poem predict higher onomatopoeia ratings (Fig. 4.9C). On the other hand, if the arousal level of all nuclei in a poem taken together is getting very high or very low, the poem is perceived less onomatopoetic (Fig. 4.9D). The overall picture receives further complexity by the fact that a more positive valence specifically of salient codas augments the onomatopoeia ratings.

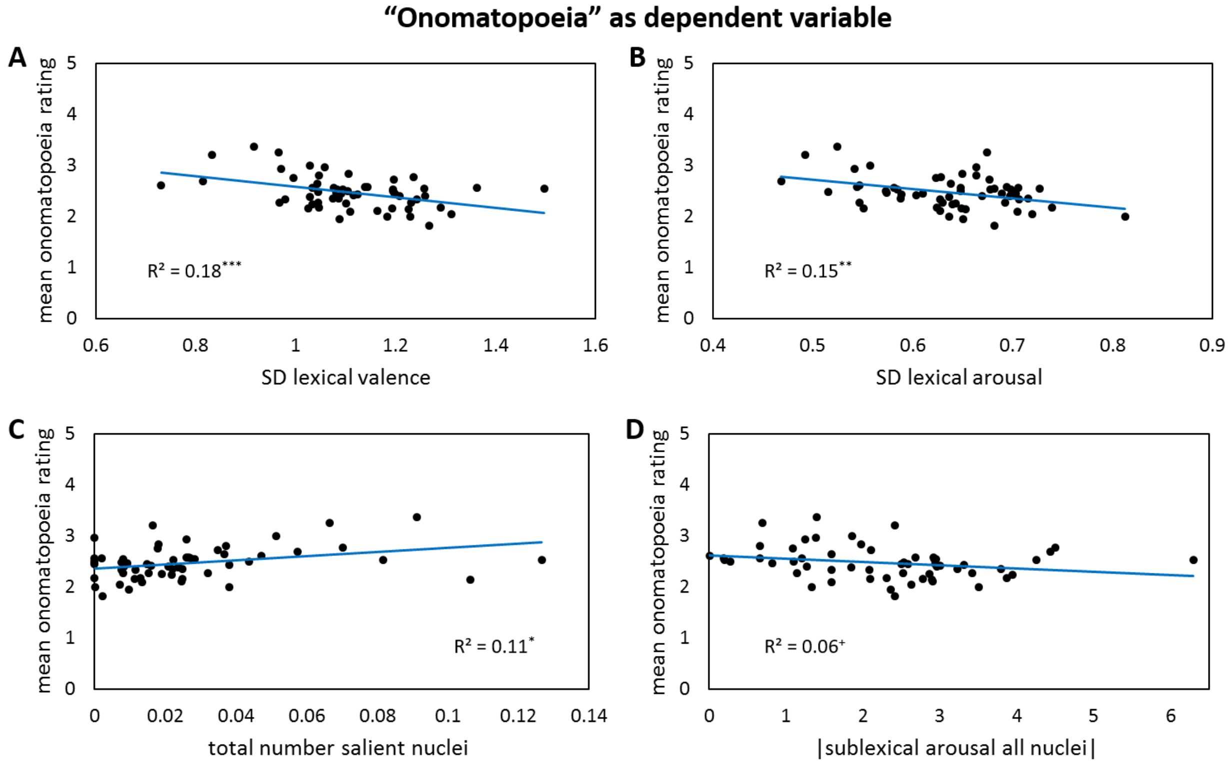

Figure 4.9: Bivariate correlations between onomatopoeia ratings and four predictors: the standard deviations of lexical valence (A) and lexical arousal in each poem (B), the number of all salient nuclei per poem weighted by its lengths (C), and the absolute values of the sigma factors for sublexical arousal of all nuclei – note that the correlation gets significant after partialling out the influence of the other predictors (D).

In summary, it can be stated that in all of the regression models at least two out of the three examined levels of affective text analysis contribute significantly but differently to the variance in the respective dependent rating variable. In eight out of nine cases, at least one of the lexical variables valence or arousal is contained in the regression model, in six cases it enters the model first. Especially the inclusion of lexical arousal in seven models increases the amount of explained variance to a noticeable extent. Lexical valence supports four models significantly. The newly defined inter-lexical variables, whose task it is to represent dynamic shifts and spreads of affective lexical content, find their way into the regression equations in five out of nine models. From the huge number of sublexical predictor variables, prominently the arousal level of salient segments consistently explains variance in eight out of nine models. In addition, the pure number of salient segments in a poem, disregarding their affective values, plays a role in four of the nine regression models.

Regarding the different abstraction levels of the rating variables, the best goodness of fit is achieved for the discrete affective concepts ratings (47 – 70 % of variance accounted for), closely followed by the dimensional affective ratings (43 – 59 % variance accounted for). Even for ratings at the most abstract level of *general affective meaning* – including aesthetic as well as onomatopoetic ratings – still 23 – 48 % of the variance are accounted for by basic textual predictor variables.

4.5. Discussion

This study investigates to which extent affective connotations at the rather basic textual dimensions of phonological units and single words (or the relative positions of the latter) influence the overall affective perception of poetry. For this purpose, we used the volume *“verteidigung der wölfe”* by the author Hans Magnus Enzensberger that is categorically divided into friendly, sad, and spiteful poems. To estimate their affective perception by the reader, we collected ratings of the poems on several affective scales, ranging from the basic dimensions valence and arousal to the author-based discrete affective dimensions friendliness, spitefulness, and sadness, to aesthetic evaluations of poeticity and liking, as well as the concept of onomatopoeia. To identify basic textual sources potentially determining these ratings we quantified affective properties of the texts (using valence and arousal values from large-scale normative lexical databases) at three different basic text levels: sublexical, lexical and inter-lexical. We then used these measures as predictor variables in a stepwise multiple regression approach to test how much of the variance in the perceived *general affective meaning* can be accounted for by these textual variables, and how these influences may vary across different rating dimensions.

Overall, our results from the different regression models show that a prominent portion of the variance in affective and further aesthetic and onomatopoetic ratings of our poems can be accounted for by affective features at the sublexical, lexical and inter-lexical level. These findings suggest that very basic affective processes play a crucial role in poetry perception. Note that we do not argue that higher-level processes would not matter, they are just not studied in our approach.

The best predictors of the perception of the *general affective meaning* of the poems – assessed via ratings – were the average lexical valence and arousal values of words – in terms of their deviation from an expected average value – contained in the poems. Pragmatically speaking, this would mean that it is sufficient to put words with specific affective connotations together to create half of the affective impact a poem is able to provoke in the reader. Again, while this view may appear extremely minimalistic, it is well in line with other findings from reading studies using normal sentences or passages from novels (Anderson & McMaster, 1982; Bestgen, 1994; Hsu, Jacobs, Citron, & Conrad, 2015; Whissell et al., 1986; Whissel, 1994). Beyond the single word level, our study provides a number of novel results for inter-lexical phenomena and how they contribute to the affective reading experience. From a neuroscientific perspective, Hsu, Jacobs, Citron, & Conrad (2015) and Jacobs (2014, 2015a) have already shown how inter-lexical affective features such as the span of lexical arousal values across a text passage can account for variance of arousal (Hsu, Jacobs, Citron, & Conrad, 2015) and suspense ratings (Lehne, 2014) as well as elicit increased activation of brain areas associated with affective processing (Hsu, Jacobs, Citron, & Conrad, 2015). In our data, for instance, the overall ratings of arousal induced by a poem were best predicted not by the average lexical arousal values but rather by specific maxima of lexical arousal. The maximum lexical arousal value in a text is a mathematical constituent of the arousal span (max – min) and probably the most relevant one, as it represents salient peaks or particularly exciting moments in a text – which well fits the general view on this emotion dimension as an alert system reacting immediately to salient affective input. Such findings underline the importance of deviation from expected patterns – here the outstanding arousal level of one single word in a text – for foregrounding effects (compare with the Neurocognitive Poetics model, NCPM, Jacobs, 2014, 2015a).

Furthermore, our novel operationalization of the evolution of affective content throughout a text – correlating lexical affective values with word position – yields a number of interesting results: The respective measures for lexical valence and arousal evolution significantly contribute to predicting affective evaluations of poems' general "sadness", "friendliness", or "arousal". For instance, poems were perceived as sadder when affective values of words became increasingly negative and less arousing toward the end. Instead, poems were perceived as more friendly when words of an increasingly positive character were used toward the final lines of a poem. We conclude that these correlations between word positions and affective values offer a good proxy for how overall affectivity is being continuously created throughout the course of a poem involving either a classical crescendo or a descent of affective intensity toward the end. In addition, this finding complements well with the established idea that readers naturally exert their greatest reading emphasis at the end of a sentence or passage (Gopen & Swan, 1990).

Last but not least, our data corroborate and extend recent findings on how sublexical phonological features influence affective processes during poetry reading. Aryani et al. (2016) have already shown for the same corpus how a sublexical, phonologically defined measure of the *basic affective tone* is significantly associated with both the author-given affective labels of single poems and the readers' evaluations of the *general affective meaning*. That is, for instance, valence ratings of poems get more negative, or spitefulness ratings increase, when poems feature particularly many phonological segments of high arousal potential (i.e., occurring in many words of highly arousing lexical meaning – hence reflecting *phonological iconicity* of language). In the present study, using a huge number of predictors from different text levels, we could show that these effects of *basic affective tone* indeed seem to occur independently of the lexical affective content of the poems, as effects persist even after the very robust effects of lexical affective values have been partialled out in our multiple regression models. Note also that control measures of the *basic affective tone* – not using the phonologically salient but all phonological segments – only rarely account for significant amounts of variance of the ratings in our multiple regression models (and if then only referring to specific subsyllabic units), while the EMOPHON's measures based on phonological salience did so in eight out of nine regression models. This is strong evidence that phonological salience in combination with *phonological iconicity* can be considered an important poetic device. While the choice of words and their arrangement is obviously a major concern for poetic style, our data suggest that affective sublexical phonology may be crucial for choosing the words that best fit a given poetic purpose (see also Jakobson's "subliminal verbal patterning in poetry", 1980b). Importantly, our data also show that readers are obviously sensitive to phonological salience per se: Subjective ratings of poeticity and onomatopoeia were significantly associated to the number of phonological segments qualified as phonologically salient by the EMOPHON tool.

At the level of rating dimensions as dependent variables – and from a general perspective – our study offers an interesting comparison between rather global evaluations of the *general affective meaning* of poems using the terms of dimensional emotion models (valence and arousal), specific affective dimensions presumably best suitable for the given corpus (sadness, spitefulness and friendliness), and the more aesthetic evaluations of liking and poeticity, as well as the further evaluation of onomatopoeia. Goodness of fit for regression models trying to predict the latter three dimensions was clearly less as compared to the other two groups. This is no surprise, as in the case of valence and arousal ratings, criteria and predictor variables are based on identical operationalizations of affect (as all predictors were quantified using valence and arousal values). The author-given labels of spitefulness, sadness, and friendliness deliver even more impressive fits, presumably because they might simply capture the entire variance of affective content of these poems in optimal ways. Still, our approach offers interesting insights on how more abstract evaluations of poetry (such as participants' liking of a poem or the ascription of poeticity and onomatopoeia to a text) relate to the basic affective dimensions of valence and arousal at lexical and sublexical textual levels: A remarkable finding is the decrease in general liking ratings of poems with increasing arousal – concerning both the words (or concepts dealt with) in a poem, and its phonological content (also see Aryani et al., 2016, for the prominent role of sublexical arousal). Note that this might meet a general principle of emotion processing, as already Fechner related aesthetic preference for arousal states according to the "principle of the aesthetic middle", meaning that people prefer "a certain medium degree of arousal, which makes them feel neither overstimulated nor dissatisfied by a lack of sufficient occupation" (Fechner, 1876, vol. 2, pp. 217-218; also see Berlyne, 1971, and Wundt, 1874). As the general arousal level of the poems in the Enzensberger volume is on average very high, a lowered lexical arousal level, as indicated by the regression results for liking, would still be of medium value. This principle also seems to generalize to the evaluation of poeticity by our participants: Both very high and very low levels of lexical arousal go along with lesser ascriptions of poeticity to the poems. Also at the sublexical level, a rather low arousal level coincides with higher poeticity ratings. Hence, any extremes at the phonological and at the lexical content level rather seem to "disturb" the perception of poeticity. A similar pattern is present for the explicit evaluation of phonological content during onomatopoeia ratings: these increased with the number of phonologically salient segments, but decreased with deviations concerning the arousal level of these segments toward either the very exciting or the very calm end of the bipolar arousal scale. Most interestingly, they also decreased with increasing spreads of lexical valence and arousal. Again, the focus – at least the conscious one – of our participants on formal features of poetry appeared to be rather disturbed by a too distracting affective variety at the level of semantic content.

Taken together, while previous studies had reported a range of effects of specific text levels influencing the affective appeal of literature (e.g., Bestgen, 1994, or Whissell, 2009, for the lexical level; Lüdtke & Jacobs, 2015, for the inter-lexical level; Aryani et al., 2016, for the sublexical level), in this study we can show in one conjoint explorative approach how sublexical, lexical, and inter-lexical affective features combine in constituting considerable parts of the perceived *general affective meaning* as well as further aesthetic and onomatopoetic evaluations of poetry.

4.5.1. Limitations and future prospects

What we consider a characteristic strength of the current approach certainly represents a shortcoming when it comes to deliver a comprehensive model of poetry perception: our very basic, or even minimalistic, contrastive approach to the standard investigation of the affective perception of poetry, which normally involves supra-lexical context or readers' personality features as well. While this alternative approach interestingly matches current computer-based approaches to poetic writing (Misztal & Indurkhya; 2014; Kirke & Miranda, 2013), it does not take into account well established phenomena of, e.g., familiarity (Bohrn et al., 2013) or comprehensibility (Leder et al., 2012) for poetry perception, nor does it allow for generalizing over different populations of readers. The latter might especially matter, considering that poetry may be differently "consumed" by expert readers with specific expert poetry reading strategies in comparison to unexperienced readers (see, e.g., Hanauer, 1995, on differences in literariness ratings between expert and novice readers, and Hanauer, 1996, regarding poeticity ratings), whereas our sample represents a randomly selected group of participants. For example, it is important to consider that people naïve to art may generally prefer art work that provides them with warm, i.e., positive and low-arousing, feelings (Winston & Cupchik, 1992). Furthermore, people less experienced with poetry might be less aware of more sophisticated stylistic devices or further meanings on a meta-level. Hence, basic textual features may play a bigger role in forming the *general affective meaning* of poetry for lay people than for experienced poetry readers. It would be interesting to investigate through follow-up studies with expert poetry consumers whether the influence of basic textual levels on affective perception would decrease with expertise. Moreover, future studies trying to complete the "emotion potential function" (Jacobs, 2015a, 2015b) for literary texts might have to include many further contextual and personality features of the readers to come up with a more complex account of affective perception of poetry.

Also, the wide variety of poetic œuvres certainly calls for cross-validations of findings with different text material and in different languages – including prose as well as everyday written and spoken language. Further, also the choice of textual measures could still be extended – for example, integrating morphemic and syntactic text levels – and refined – for example in terms of the inter-lexical measures. The merit of this study might thus just lie in having made first explorative steps toward investigating – or having opened initial insights on – text-based affective potential functions for several aspects of the *general affective meaning*. These innovative insights may also compensate methodological disadvantages of our statistical approach using a large number of predictors in stepwise multiple regression. While we opted for this specific method as it seems optimal when screening for the most influential ones among a wide range of possible candidate measures, future studies may apply more fine-grained methods to disentangle the details concerning the interplay of a restricted number of variables according to more specific research questions.

Future studies should, in particular, extend our investigations to (i) the works of other writers – as some of our findings may in theory result from an idiosyncratic writing style of H. M. Enzensberger, (ii) (non-)literary texts or even everyday speech – in different languages, and (iii) affective ratings from different types of reader groups including expert readers.

4.5.2. Conclusion

In this study we focused on how and to which extent affective connotations of very basic textual measures at the lexical, inter-lexical, and even sublexical level of a poem – that can all be derived from existing normative databases – determine the perception of the *general affective meaning* of poetry in a way that proves quantifiable beyond the specific context of a given poem, author, or recipient. By applying an exhaustive exploratory regression analysis to a comprehensive corpus of poems and their ratings from hundreds of readers, we found that a significant amount of variance in discrete and dimensional affective ratings of poetry can be accounted for solely by text-based affective measures from different levels of processing. In all of the presented statistical models – focusing on different aspects of the *general affective meaning* – variance of each rating dimension is significantly accounted for by affective properties of several text levels: while the lexical one generally explains the biggest amount of variance, further significant effects in explaining residual variance are found for the alternative sublexical and inter-lexical text levels. Thus, our research brings together previous accounts on specific effects of single text levels, showing how they may co-exist each in their own right or interact to constitute the complex holistic framework of poetry perception. Taken together, the affective properties of text elements from all three text levels could account for 43 to 70 % of the variance in the perceived *general affective meaning* of the here utilized poetry and still for 23 to 48 % of the variance in further aesthetic and onomatopoetic evaluations of the poems – a substantial amount purely accounted for by textual elements which should not be neglected in future affective analyses of poetry. This mixed-level approach represents a first step toward quantifying and computationally modeling what Jakobson hypothesized about the “Framework of Language” (1980): “Each [text] level above brings new particularities of meaning...”. Our explorative regression models may guide the way for various future ideas on interrelations between specific textual features and the perception of *general affective meaning* in further poem corpora and other literary work.

4.6. Ethics statement

This study was approved by the ethics committee of the Freie Universität Berlin and was conducted in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). We conducted a non-experimental, voluntary online survey, in which people had to read and judge poems. In the instructions we told the participants that they can skip the survey any time they want to. If they had any questions regarding the survey they could contact us any time (e-mail addresses provided). There was only one participant of the online survey who was only 17 years old. We did not have any additional instructions for minors or their parents. But we assume that rating poetry does not pose a significant difference between teenagers and adults.

4.7. Acknowledgements

This research was funded through the grant 410 "Sound physiognomy in language organization, processing and production" to MC from the German Research Foundation (DFG) via the Cluster of Excellence “Languages of Emotion” at the Freie Universität Berlin. We thank Kai-Michael Würzner for advice on the PoS tagging of the poems. We acknowledge support by the Open Access Publication Funds of the Freie Universität Berlin.

4.8. References

Albers, S. (2008). *Lautsymbolik in ägyptischen Texten* (Vol. 54). Darmstadt: Philipp Von Zabern Verlag GmbH.

Anderson, C. W. and McMaster, G. E. (1982). Computer assisted modeling of affective tone in written documents. *Computers and the Humanities*, *16*, 1-9.

Aryani, A., Jacobs, A. M., and Conrad M., (2013). Extracting salient sublexical units from written texts: Emophon, a corpus-based approach to phonological iconicity. *Frontiers in Psychology*, *4*, 654. doi: 10.3389/fpsyg.2013.00654

Aryani, A. and Kraxenberger, M., Ullrich, S., Jacobs, A.M., and Conrad, M. (2016). Measuring basic affective tone in poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts, 10*(2), 191-204. doi: 10.1037/aca0000033

Astley, N. (Ed.). (2006). *Bloodaxe Poetry Introductions: Enzensberger, Holub, Sorescu, Tranströmer.* Highgreen: Bloodaxe Books.

Auracher, J., Albers, S., Zhai, Y., Gareeva, G., and Stavniychuk, T. (2011). P is for Happiness, N is for Sadness: universals in sound iconicity to detect emotions in poetry. *Dis. Process.*, *48*, 1-25. doi: 10.1080/01638531003674894

Berlyne, D.E., (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.

Bestgen, Y. (1994). Can emotional valence in stories be determined from words? *Cognition & Emotion*, *8*, 21-36. doi: 10.1080/02699939408408926

Bleich, D. (1978). *Subjective Criticism*. Baltimore, MD: Johns Hopkins University Press.

Bohrn, I.C., Altmann, U., Lubrich, O., Menninghaus, W., and Jacobs, A.M. (2013). When we like what we know - A parametric fMRI analysis of beauty and familiarity. *Brain Lang.*, *124*(1), 1-8. doi:10.1016/j.bandl.2012.10.003.

Bradley, M.M., Greenwald, M.K., Petry, M.C., and Lang, P.J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 379-390.

Bradley, M.M. and Lang, P.J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings.* Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A.M., Bölte, J., and Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Exp. Psychol., 58*, 412-424. doi:10.1027/1618-3169/a000123

Bühler, K. (1934). *Sprachtheorie*. Oxford: Fischer.

Conrad, M., Recio, G., and Jacobs, A.M. (2011). The time course of emotion effects in first and second language processing: A cross cultural ERP study with German–Spanish bilinguals. *Frontiers in Psychology*, *2*(351), 1-16. doi:10.3389/fpsyg.2011.00351

Conrad, M., Aryani, A., Schmidtke, D.S., and Jacobs, A.M., (in preparation). The sound of dictators: Encoding of affective meaning via phonological iconicity.

Cupchik, G.C. and A.S. Winston (1992). Reflection and reaction: A dual-process analysis of emotional responses to art. In: L. Ya. Dorfman, D.A. Leontiev, V.M. Petrov and V.A. Sozinov (eds.), *Art and emotions*, *2*, 65-72. Perm, CIS: The Perm State Institute of Culture.

Cupchik, G. C. (1994). Emotion in aesthetics: Reactive and reflective models. *Poetics*, *23*, 177-188. doi:10.1016/0304-422X(94)00014-W

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature neuroscience, 19*(1), 158-164. doi:10.1038/nn.4186

Enzensberger, H.M. (1957). *die verteidigung der wölfe.* Frankfurt am Main: Suhrkamp.

Fechner, G.T. (1876). *Vorschule der Asthetik* (vol. 2). Leipzig: Breitkopf & Härtel.

Fónagy, I. (1961). Communication in poetry. *Word*, *17*, 194-218.

Frijda, N.H. (1989). Aesthetic emotions and reality. *American Psychologist*, *44*, 1546-1547.

Gopen, G.D., and Swan, J. A. (1990). The science of scientific writing. *American Scientist*, *78*(6), 550-558.

Hanauer, D. (1995). The effects of educational background on literary and poetic text categorization judgements. In: G. Rusch (ed.), *The empirical study of literature*, 338-347. Siegen: LUMIS Siegen University Press.

Hanauer, D. (1996). Integration of phonetic and graphic features in poetic text categorization judgements. *Poetics*, *23*(5), 363-380.

Hofmann, M.J., Kuchinke, L., Tamm, S., Võ, M. L.-H, and Jacobs, A.M. (2009). Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(4), 389-397.

Hoffstaedter, P. (1987). Poetic text processing and its empirical investigation. *Poetics, 16*, 75-91.

Hsu, C.-T., Conrad, M., and Jacobs, A.M. (2014). Fiction feelings in Harry Potter: Haemodynamic response in the mid-cingulate cortex correlates with immersive reading experience. *NeuroReport*, *25*(17), 1356-1361. doi:10.1097/WNR.0000000000000272

Hsu, C.-T., Jacobs, A.M., and Conrad, M. (2015). Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. *Cortex*, *63*, 282-295. doi:10.1016/j.cortex.2014.09.002

Hsu, C.-T., Jacobs, A.M., Citron, F., and Conrad, M. (2015). The Emotion Potential of Words and Passages in Reading Harry Potter: An fMRI Study, *Brain & Language*, *142*, 96-114.

Hsu, C.-T., Jacobs, A.M., Altmann, U., and Conrad, M. (2015). The Magical Activation of Left Amygdala when Reading Harry Potter: An fMRI Study on How Descriptions of Supra-Natural Events Entertain and Enchant. *PLoS ONE*, *10*(2): e0118179. doi:10.1371/ journal.pone.0118179

Jacobs, A.M. (2011). Neurokognitive Poetik: Elemente eines Modells des literarischen Lesens [Neurocognitive poetics: Elements of a model of literary reading]. In *Gehirn und Gedicht: Wie wir unsere Wirklichkeiten konstruieren [Brain and poetry: How we construct our realities],* eds. R. Schrott & A. M. Jacobs (München: Hanser), 492-520.

Jacobs, A.M., Lüdtke, J., and Meyer-Sickendiek, B. (2013). Bausteine einer Neurokognitiven Poetik: Foregrounding/Backgrounding, lyrische Stimmung und ästhetisches Gefallen (Elements of a neurocognitive poetics: Foregrounding/backgrounding, typical mood, and aesthetic pleasure). In *Stimmung und Methode (Mood and Method)*, eds. B. Meyer-Sickendiek and F. Reents (Tübingen: Mohr Siebeck), 63-94.

Jacobs, A.M. (2014). Affektive und ästhetische Prozesse beim Lesen: Anfänge einer neurokognitiven Poetik (Affective and aesthetic processes in reading: towards a neurocognitive poetics). In *Sprachen der Emotion (Languages of Emotion)*, eds. G. Gebauer and M. Edler (Frankfurt: Campus), 134-154.

Jacobs, A.M., Võ, M.L.-H., Briesemeister, B.B., Conrad, M., Hofmann, M.J., Kuchinke, L., Lüdtke, J., and Braun, M. (2015). 10 years of BAWLing into affective and aesthetic processes in reading: what are the echoes? *Frontiers in Psychology*, *6*, 714.

Jacobs, A.M. (2015a). Towards a neurocognitive poetics model of literary reading. In *Towards a cognitive neuroscience of natural language use,* ed. R. Willems (New York, NY: Cambridge University Press), 135-159.

Jacobs, A.M. (2015b). Neurocognitive Poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, 186.

Jakobson, R. (1960). Closing statement: Linguistics and poetics. In *Style in language,* ed. T.A. Sebeok (Cambridge, MA: MIT Press), 350-377.

Jakobson, R. (1980a). *The framework of language*. Ann Arbor, MI: University of Michigan.

Jakobson, R. (1980b). Subliminal Verbal Patterning in Poetry. *Poetics Today, 2*(1a), 127-136.

Kirke, A. and Miranda, E. (2013). Emotional and multi-agent systems in computer-aided writing and poetry. in *Procs. of the Artificial Intelligence and Poetry Symposium (AISB'13)*, eds. R. Aylett and G. Michaelson (London: The Society for the Study of Artificial Intelligence and the Simulation of Behaviour), 17-22.

Köhler, W. (1929). *Gestalt Psychology*. New York, NY: Liveright.

Kuchinke, L., Jacobs, A.M., Grubich, C., Võ, M. L.-H., Conrad, M., and Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: An fMRI study. *NeuroImage*, *28*(4), 1022-1032.

Leder, H., Belke, B., Oeberst, A., and Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489-508.

Leder, H., Gerger, G., Dressler, S.G., and Schabmann, A. (2012). How art is appreciated. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 2.

Lehne, M. (2014). *Emotional Experiences of Tension and Suspense: Psychological Mechanisms and Neural Correlates* (Doctoral dissertation, Freie Universität Berlin).

Liu, H., Hu, Z., and Peng, D. (2013). Evaluating Word in Phrase: The Modulation Effect of Emotional Context on Word Comprehension. *Journal of psycholinguistic research*, *42*, 379-391. doi:10.1007/s10936-012-9221-7

Lüdtke, J., Meyer-Sickendieck, B., and Jacobs, A.M. (2014). Immersing in the stillness of an early morning: Testing the mood empathy hypothesis of poetry reception. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(3), 363.

Lüdtke, J. and Jacobs, A.M. (2015). The emotion potential of simple sentences: additive or interactive effects of nouns and adjectives? *Frontiers in Psychology*, *6*, 1137.

Mar, R.A., Oatley, K., Djikic, M., and Mullin, J. (2011). Emotion and narrative fiction: Interactive influences before, during, and after reading. *Cognition & Emotion*, *25*, 818-833. doi:10.1080/02699931.2010.515151

Marković, S. (2012). Components of aesthetic experience: aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *i-Perception*, *3*(1), 1-17.

Menninghaus, W., Bohrn, I.C., Altmann, U., Lubrich, O., and Jacobs, A.M. (2014). Sounds funny? Humor effects of phonological and prosodic figures of speech. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 71-76.

Menninghaus, W., Bohrn, I.C., Knoop, C.A., Kotz, S.A., Schlotz, W., and Jacobs, A.M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48-60.

Miall, D.S. and Kuiken, D. (1994). Foregrounding, defamiliarization, and affect response to literary stories. *Poetics*, *22*, 389-407. doi:10.1016/0304-422X(94)00011-5

Miller, G.A. (1954). Psycholinguistics. In G. Lindzey (Ed.): *Handbook of social psychology.* Reading, Mass.: Addison-Wesley, 693-708.

Misztal, J. and Indurkhya, B. (2014). Poetry generation system with an emotional personality. In *Proceedings of 5th International Conference on Computational Creativity (ICCC 2014)*, eds. S. Colton, D. Ventura, N. Lavrač, and M. Cook (Association for Computational Creativity), 72-81.

Mukařovský, J. (1964). Standard Language and Poetic Language. in *A Prague School Reader on Esthetics, Literary Structure, and Style*, ed. P. L. Garvin (Washington, DC: Georgetown University Press), 17-30.

Oatley, K. (1994). A taxonomy of the emotions of literary response and a theory of identification in fictional narrative. *Poetics*, *23*, 53-74. doi:10.1016/0304-422X(94)P4296-S

Obermeier, C., Menninghaus, W., Von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., and Kotz, S. A. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology*, *4*, 1-10. doi:10.3389/fpsyg.2013.00010

Osgood, C.E. and Suci, G.J. (1955). Factor analysis of meaning. *Journal of experimental psychology*, *50*(5), 325-338.

Osgood, C.E., Suci, G., and Tannenbaum, P. (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press.

Osgood, C.E., May W.H., and Miron, M.S. (1975*). Cross-cultural universals of affective meaning*. Urbana, IL: University of Illinois Press.

Palazova, M., Mantwill, K., Sommer, W., and Schacht, A. (2011). Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. *Neuropsychologia*, *49*, 2766-2775. doi:10.1016/j.neuropsychologia.2011.06.005.

Paltoglou, G. (2014). Sentiment analysis in social media. In *Online Collective Action,* eds. N. Agarwal, M. Lim, and R. T. Wigand (Vienna: Springer), 3-17. doi:10.1007/978-3-7091-1340-0_1

Recio, G., Conrad, M., Hansen, L.B., and Jacobs, A.M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134*, 34-43.

Russell, J.A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of Personality and Social Psychology*, *36*(10), 1152-1168.

Russell, J.A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145-172. doi:10.1037/0033-295X.110.1.145

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225-239. doi:10.1037/h0070931

Schauenburg, G., Ambrasat, J., Schröder, T., von Scheve, C., and Conrad, M. (2014). Emotional connotations of words related to authority and community. *Behavior research methods*, *47*(3), 720-735.

Schmidtke, D.S., Schröder, T., Jacobs, A.M., and Conrad, M. (2014a). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior research methods*, *46*(4), 1108-1118.

Schmidtke, D.S., Conrad, M., & Jacobs, A.M. (2014b). Phonological iconicity. *Frontiers in psychology*, *5*, 80. doi:10.3389/fpsyg.2014.00080

Schrott, R. and Jacobs, A. M. (2011). *Gehirn und Gedicht: Wie wir unsere Wirklichkeiten konstruieren [Brain and poetry: How we construct our realities]*. München: Hanser.

Scott, G.G., O'Donnell, P.J., Leuthold, H., and Sereno, S.C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology*, *80*, 95-104.

Sweeney, K. and Whissell, C.M. (1984). A dictionary of affect in language: I. Establishment and preliminary validation. *Perceptual and Motor Skills*, *59*, 695-698.

Taylor, I.K. and Taylor, M.M. (1965). Another look at phonetic symbolism. *Psychol. Bull.*, *64*, 413-427. doi:10.1037/h0022737

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology, 61*(12), 2544-2558.

Turner, F. & Poeppel, E. (1983). The neural lyre: Poetic meter, the brain, and time. *Poetry Magazine, 142*(5), 277-309.

Tsur, R. (1992) *What makes sound patterns expressive?* Durham, NC: Duke University Press.

van Peer, W. (1986). *Stylistics and Psychology: Investigations of Foregrounding.* London, UK: Croom Helm.

van Peer, W., Hakemulder, J., and Zyngier, S. (2007). Lines on feeling: Foregrounding, aesthetics and meaning. *Language and Literature*, *16*, 197-213. doi:10.1177/0963947007075985

Võ, M. L.-H., Jacobs, A. M., and Conrad, M. (2006). Cross-validating the Berlin Affective Word List. *Behavior research methods*, *38*(4), 606-9.

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., and Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior research methods*, *41*(2), 534-8. doi:10.3758/BRM.41.2.534

Watson, D. and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219-235.

Whissell, C.M., Fournier, M., Pelland, R., Weir, D., & Makarec, K. (1986). A dictionary of affect in language: IV. Reliability, validity, and applications. *Perceptual and Motor Skills*, *62*(3), 875-888.

Whissell, C.M. (1994). A computer program for the objective analysis of style and emotional connotations of prose: Hemingway, Galsworthy, and Faulkner compared. *Perceptual and Motor Skills*, *79*(2), 815-824

Whissell, C.M. (1999). Phonosymbolism and the emotional nature of sounds: evidence of the preferential use of particular phonemes in texts of differing emotional tone. *Perceptual and Motor Skills*, *89*, 19-48. doi:10.2466/pms.1999.89.1.19

Whissell, C.M. (2000). Phonoemotional profiling: a description of the emotional flavour of English texts on the basis of the phonemes employed in them. *Perceptual and Motor Skills*, *91*, 617-648. doi:10.2466/pms.2000.91.2.617

Whissell, C. (2009). Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language. *Psychological reports*, *105*(2), 509-521.

Winston, A.S. and Cupchik, G.C. (1992). The evaluation of high art and popular art by naive and experienced viewers. *Visual Arts Research*, *18*(1), 1-14.

Wiseman, M., and van Peer, W. (2003). Roman Jakobsons Konzept der Selbstreferenz aus der Perspektive der heutigen Kognitionswissenschaft. In *Roman Jakobsons Gedichtanalysen*, eds. H. Birus, S. Donat, and B. Meyer-Sickendiek (Göttingen: Wallstein), 277-306.

Wundt, W.M. (1874). *Grundzüge der physiologischen Psychologie.* Leipzig: Engelmann.

Wundt, W.M. (1896). *Grundriss der Psychologie.* Leipzig: Engelmann.

4.9. Supplementary material

Table S1:

Full listing of all predictor variables of the multiple regression approach

(1) sigma factor of lexical valence

(2) absolute value of the sigma factor of lexical valence

(3) sigma factor of lexical arousal

(4) absolute value of the sigma factor of lexical arousal

(5) minimum of lexical valence

(6) maximum of lexical valence

(7) valence span (max – min)

(8) minimum of lexical arousal

(9) maximum of lexical arousal

(10) arousal span (max – min)

(11) standard deviation of lexical valence

(12) standard deviation of lexical arousal

(13) correlation coefficient of word position and arousal

(14) absolute value of the correlation coefficient of word position and arousal

(15) correlation coefficient of word position and valence

(16) absolute value of the correlation coefficient of word position and valence

(17) correlation coefficient of word position and the absolute value of valence

(18) absolute value of the correlation coefficient of word position and the absolute value of valence

(19) total number of words per poem

(20) total n of salient onsets only

(21) total n of salient nuclei only

(22) total n of salient codas only

(23) total n of all salient subsyllabic segments together

(24) sigma factor of sublexical valence for salient onsets

(25) sigma factor of sublexical valence for salient nuclei

(26) sigma factor of sublexical valence for salient codas

(27) sigma factor of sublexical valence for all salient segments

(28) sigma factor of sublexical arousal for salient onsets

(29) sigma factor of sublexical arousal for salient nuclei

(30) sigma factor of sublexical arousal for salient codas

(31) sigma factor of sublexical arousal for all salient segments

(32) absolute value of the sigma factor of sublexical valence for salient onsets

(33) absolute value of the sigma factor of sublexical valence for salient nuclei

(34) absolute value of the sigma factor of sublexical valence for salient codas

(35) absolute value of the sigma factor of sublexical valence for all salient segments

(36) absolute value of the sigma factor of sublexical arousal for salient onsets

(37) absolute value of the sigma factor of sublexical arousal for salient nuclei

(38) absolute value of the sigma factor of sublexical arousal for salient codas

(39) absolute value of the sigma factor of sublexical arousal for all salient segments

(40) sigma factor of sublexical valence of all onsets

(41) sigma factor of sublexical valence of all nuclei

(42) sigma factor of sublexical valence of all codas

(43) sigma factor of sublexical valence of all segments

(44) sigma factor of sublexical arousal of all onsets

(45) sigma factor of sublexical arousal of all nuclei

(46) sigma factor of sublexical arousal of all codas

(47) sigma factor of sublexical arousal of all segments

(48) absolute value of the sigma factor of sublexical valence of all onsets

(49) absolute value of the sigma factor of sublexical valence of all nuclei

(50) absolute value of the sigma factor of sublexical valence of all codas

(51) absolute value of the sigma factor of sublexical valence of all segments

(52) absolute value of the sigma factor of sublexical arousal of all onsets

(53) absolute value of the sigma factor of sublexical arousal of all nuclei

(54) absolute value of the sigma factor of sublexical arousal of all codas

(55) absolute value of the sigma factor of sublexical arousal of all segments

5. General discussion

In this thesis I investigated the psychophysiological reality of a holistic operationalization of affectivity at the phonological level of the German language, the *sublexical affective potential*, derived from statistical co-occurrences of particular subsyllabic segments in words with a particular affective meaning in the BAWL-E corpus, and its interaction with affectivity at higher language levels – thereby reflecting affective iconicity and its potential functions in language processing. The results of the EEG study could confirm a psychophysiological reality of the *sublexical affective potential*, showing similarities to affective semantic processing. Other sublexical properties such as a low frequency or high complexity of the subsyllabic segments add to the salience of negative high-arousing word sounds. The EEG results further revealed a significant interaction of the *sublexical affective potential* and the *lexical affective content* on a single-word basis around 400 ms after stimulus onset, representing a facilitated lexical access when the affective qualities from the sublexical and lexical level are consistent. An applied poetry perception study further showed that affective information from the sublexical level interactively climbs the ladder of language levels, adding to the impacts of affective features from the lexical as well as the inter-lexical level. Hence, the *general affective meaning* of a text (here, poems) is interactively created during the reading process by influences from all language levels. In the following, I will further discuss these findings and integrate them with accounts of neural semantic representations, affectivity in language, interactive reading processes, and iconicity and arbitrariness in general.

5.1. The *sublexical affective potential* and its neurophysiological reality

As the German affective language corpus analyses in chapter 1.6 showed, the choice of phonological units to constitute words of particular affective meaning is not arbitrary. This led to my working hypothesis that sound and meaning are linked in language – not only in some confined examples of obviously iconic words, but unobtrusively in a holistic manner throughout the lexicon. The basic affective dimensions valence and arousal seem to be suitable to subsume a wide variety of semantic dimensions that have been shown to be involved in sound-meaning correspondences (see chapter 1.5). To empirically corroborate the *sublexical affective potential* of words, which was statistically derived in a top-down manner from the *lexical affective content* of words in the BAWL-E database (namely, word valence and arousal ratings) as a function of the sound-meaning correspondences, an EEG study was conducted to investigate the psychophysiological reality of this *sublexical affective potential*, for example, whether it is actually affective in nature.

In the EEG study, words with a *sublexical affective potential* of negative valence and high arousal were contrasted with words of a neutral, low-arousing *sublexical affective potential*. To be able to compare the results to known EEG components of lexical affective processing, the same affective contrast was manipulated at the lexical level. These combinations of negative valence with high arousal and neutral valence with low arousal constitute a large affective contrast with reference to the U-shaped distribution of lexical valence and arousal found in the BAWL (see Figure 1.2) and affective word databases of other languages (Schmidtke, Schröder, Jacobs, & Conrad, 2014). A similar relationship can be assumed for the sublexical level (see chapter 1.6 and Figure 1.4). The subjects had to perform a lexical decision task with no knowledge about the affective manipulations of the word material.

Using a stimulus set that was highly controlled for more than 40 potentially confounding factors at the sublexical as well as the lexical level, the EEG revealed a small but significant effect of the s*ublexical affective potential*, represented by an early posterior negativity (EPN) at around 250 ms after word onset, with a slightly higher negativity for sublexically negative high-arousing words compared to neutral low-arousing words (see chapter 2). This component seems to be an analogue of the lexical EPN, which appears at around 200 – 300 ms after stimulus onset and is known to stand for affective semantic processing (Kissler, Herbert, Peyk, & Junghofer, 2007). Hence, the fact that there are also detectable differences for the *sublexical affective potential* of words in the same time range and with a similar topography indexes that this early processing step is sensitive to affective information at different levels of language input. In terms of the DMF model (see chapter 1.8) the lexical EPN and the comparable sublexical EPN reflect the emotive-evaluative modalities of the word's concept (the *lexical affective content*) and the word form (the *sublexical affective potential*) respectively. These results confirm my hypothesis that while reading a word the emotive-evaluative features of both the concept and the word form become co-activated due to the neural network character of word processing.

The second main finding of the study was that the *sublexical affective potential* seems to be highly entangled with sublexical frequencies and complexity. This became already obvious during stimulus selection when trying to manipulate the *sublexical affective potential* and *lexical affective content* while controlling for dozens of other lexical as well as sublexical variables. For example, when words with complex syllable onsets or phonological units of low frequency were excluded, also the manipulation strength got diminished, meaning that the affective differences between contrasting groups became smaller. To avoid losing the manipulation strength to the strict control and to find out whether the negative high-arousing *sublexical affective potential* might actually be created by low-frequent and complex phonological units – i.e., their perceptual salience – we decided to prepare the two stimuli sets, one maximally controlled and one maximally manipulated. As shown in chapter 2, for the maximally manipulated set a robust higher negativity for words with negative high-arousing compared to neutral low-arousing *sublexical affective potential* emerged at around the time of the EPN and lasted for 400 ms. This suggests that formal sublexical systematicities such as the frequency and complexity of phonological units may be part of the effect constituted by a negative high-arousing *sublexical affective potential* by interacting with and thus potentially intensifying it. One potential explanation of how this prolonged negativity – reflecting enhanced neural processing efforts – might come about, could be that complex phonological units activate further neurons in the semantic network, for example those that are responsible for complex articulatory movements. According to the Action Perception Theory of language (Pulvermüller et al., 2006, 2014), motor aspects of a word get co-activated even when we are just perceiving a word without any action intention. Due to the dual aspect of the neural semantic network I assume this not only to be the case for the word concept, for example an action such as kicking or lifting, but as well for the word form, hence its articulation (or even writing movements). And the articulation of a complex onset such as /kr/ might surely need more neural activation in the part of the motor cortex which is responsible for articulation representation than a simple onset such as /k/.

Regarding phonological units of low frequency, it might just take longer to activate the respective neural node, for example, due to prediction errors when other, more probable phonemes were predicted to follow an antecedent phoneme, which then did not turn out to appear. These and presumably further different co-activation mechanisms that are triggered by the uncontrolled formal sublexical features, might interact ongoingly and – if they all work in the same direction – build up the effect that we see as the prolonged negativity in the EEG. Hence, several sublexical properties might partly be involved in creating the *sublexical affective potential*.

The *lexical affective content* contrast also elicits a late positive complex (LPC) around 400 – 700 ms after stimulus onset, which is not the case in the *sublexical affective potential* contrast. The LPC is assumed to indicate more elaborated and task-dependent semantic processing of emotional stimuli, which includes, for example, continued stimulus evaluation such as categorization or memory updating. Thus, it seems reasonable that the *sublexical affective potential* as implicit word feature and not relevant for the lexical decision task, does not elicit an LPC. Future EEG studies using tasks that directly address the *sublexical affective potential* (for example, in iconicity or affective word sound ratings or decision tasks) might expect an LPC.

In this study, the focus was on investigating the psychophysiological reality of the *sublexical affective potential*, which had best chances using a strong affective contrast of naturalistic relevance – here, contrasting potentially alerting concepts and sounds with innocuous ones. Also with reference to the asymmetrical U-shaped distribution of valence and arousal word ratings (Figure 1.2) the contrast of negative high-arousing words and neutral low-arousing words is the largest possible one, as very positive concepts are not rated as arousing as negative ones. Yet, this contrast cannot be used to distinguish between the separate contributions of both affective dimensions to the effect(s). Such a disentanglement between valence and arousal effects at the sublexical level could be the subject of future studies. A helpful study to develop hypotheses for the respective contributions of valence and arousal at the sublexical level – now that we know that the affective processing at the sublexical level is similar to that at the lexical level – could be the fine-grained EEG study by Recio, Conrad, Hansen, & Jacobs (2014), which orthogonally contrasted lexical valence and arousal in three gradations each. Interestingly, they showed that there are no valence effects anymore, when the arousal level is high. This points into the direction that the effect of arousal might be the dominating one, meaning that valence effects are rather to be found when arousal is medium to low. Furthermore, Recio et al. (2014) could only find arousal effects within the neutral and negative valence conditions but not among the positive words. Hence, in order to contrast different arousal levels, positive words should rather not be involved. A high salience in terms of enhanced visual processing could be shown particularly for negative words of high arousal (Recio et al., 2014) – the same combination of valence and arousal that was used in our study. This makes sense biologically, as it is one of the brain's tasks to detect and avoid potential danger. Therefore, it must process relevant stimuli in an enhanced way. All in all, this strengthens the environmental validity of the affective contrast we used in our study.

Last but not least, that no behavioral effects were found in our EEG study, could also be explained by the study of Recio et al. (2014), at least with regard to the *lexical affective content*: comparing negative high-arousing words with neutral low-arousing words also did not yield behavioral differences in their study. Competing mechanisms as outlined in the Automatic Evaluation Hypothesis (formerly known as Affective Primacy Hypothesis: Zajonc, 1980, 1984) and the Automatic Vigilance Hypothesis (Pratto & John, 1991) might be a potential reason for this (see discussion in chapter 2.4).

5.2. The interactive basis of affective iconicity

Now that we know that the corpus-statistically derived *sublexical affective potential* possesses a psychophysiological reality and thus actually represents affective word processing going on at the sublexical level, the second research question was concerned with potential interactions of such affectivity at the sublexical level with affective features of higher language levels, foremost the lexical level. Roman Jakobson already had postulated that features of language at each level of language signs participate in the meaning of the whole (Jakobson, 1980, p. 20). And the Interactive Activation Model (IAM) by McClelland & Rumelhart (1981) also predicts bidirectional mappings between the sublexical level (phonemes or letters) and lexical level of a word (see chapter 1.9), hence also between the respective affective connotations.

The 2x2 design of the EEG experiment (Figure 5.1) allows to contrast words that are consistent in their *lexical affective content* and their *sublexical affective potential* with those that are inconsistent. In terms of EEG components, inconsistent semantic relationships between or within words often yield differences in the N400 component which most likely reflects the ease of lexical access (Lau, Phillips, & Poeppel, 2008). It is known from priming studies or the Stroop task that inconsistencies lead to higher negativities in the N400 time window (300 – 500 ms after stimulus onset) than consistent relations between properties (Rebai, Bernard, & Lannou, 1997; Brown & Hagoort, 1993; Liotti, Woldorff, Perez III, & Mayberg, 2000). Such a higher negativity might reflect how lexical-conceptual prediction errors impede the activation of representational features of the word in the mental lexicon (Lau, Holcomb, & Kuperberg, 2013). Hence, we had predicted a higher negativity (i.e., a hindered lexical access) for the words that are inconsistent in their *lexical affective content* and their *sublexical affective potential* compared to those that are consistent. Indeed, as chapter 3 shows, a highly significant interaction of the affective information at the lexical and sublexical level appeared between 280 and 430 ms after stimulus onset with the topography of an N400 component – a fronto-central negativity on the scalp (Conrad, Ullrich, Schmidtke, & Kotz, 2022). According to the hypothesis, we see a reduced negativity – hence, a facilitated lexical access – for words which are consistent in their affective characteristics of the *lexical affective content* and the *sublexical affective potential* compared to inconsistent words.

Sublexical affective potential ➡ / Lexical affective content ⬇	negative valence + high arousal	neutral valence + low arousal
negative valence + high arousal	**1** **consistent** for example *„Krieg“ („war“)*	**2** **inconsistent** for example *„Waffe“ („weapon“)*
neutral valence + low arousal	**3** **inconsistent** for example *„Strumpf“ („stocking“)*	**4** **consistent** for example *„Boden“ („soil“)*

Figure 5.1: 2x2 design matrix of the EEG experiment presented in chapter 2 and 3.

A colleague re-analyzed the same EEG data conducting a frequency analysis (Glim, Ullrich, Rummer, & Conrad, in prep.). She found an increased oscillatory phase synchronization in the theta frequency band between 180 and 640 ms for the words inconsistent in their *lexical affective content* and *sublexical affective potential* compared to the consistent ones. This time frame includes the time window in which we found the N400 (280 – 430 ms). Such a synchronization of oscillatory phases represents an intensified neural communication. As it could be detected particularly at left frontal sites, where language processing takes place, this corroborates the assumption that the reinforced neural communication is necessary due to the hindered lexical processing. Here again, the inconsistent words need an intensified neural processing to resolve the mismatch between affective features at different levels of the linguistic input. All in all, the results from the EEG experiment show that formal sublexcial systematicities as represented by the *sublexical affective potential* influence affective word processing.

Yet how could the potential neural mechanisms behind such a semantic integration process look like? The Interactive Activation Model by McClelland & Rumelhart (1981) helps to visualize the initial processing steps (also see Figure 5.2): When the word stimulus pops up on the screen, the visual input activates neurons at the feature level. Particular feature constellations activate neurons representing letters. Between the letter and the word level may actually lie further levels, for example for letter combinations such as syllable onsets and codas that are typical for a particular language. In parallel, the neurons representing the spoken sound of such letters (i.e., the phonemes and phoneme clusters) become co-activated as well and function as inner speech (McClelland, Rumelhart, & Hinton, 1986).

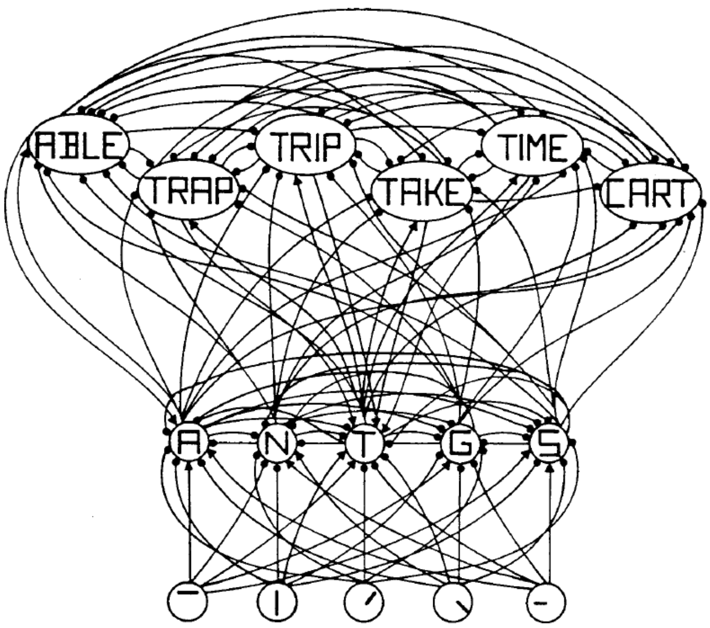

Figure 5.2: A few of the neighbors of the node for the letter *T* in the first position in a word, and their interconnections. (From McClelland, James L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account on Basic Findings. *Psychological Review*, *88*(5), 375–407. Copyright © 1981 by *American Psychological Association*. Reproduced with permission.)

When a word's onset is assessed, its neuronal activation pattern leads to pre-activations of potential word forms that begin with this onset, which may lead to the following (speculative) mechanism in terms of the Predictive Processing framework (Bar, 2007; Huang & Rao, 2011; Lupyan & Clark, 2015): These pre-activated word forms are neurally connected to their meanings which thus become slightly co-activated as well. The affective connotations of such co-activated word meanings on average function as prediction for the affective content of the word being processed. Based on this prediction, the search for the presented word form in the mental lexicon is presumably conducted among words of the respectively narrowed affective spectrum – that is, not among all words in the mental lexicon starting with the same onset, but only those whose *lexical affective content* matches the overall affective potential of that onset. If the affective meaning of the actually presented word is consistent with this affective anticipation, the word form will be quickly found in this affectively confined spectrum – hence, the prediction facilitated the lexical access. However, if the *lexical affective content* of the word being processed is inconsistent with the affective prediction, the word form cannot be found and a prediction error occurs (potentially involving some sense of saliency). In that case, a new prediction needs to be created, for example, the affective search spectrum needs to be shifted or broadened – thus increasing the overall neural processing effort. In summary, our perceptual system explores all possibilities until all constraints, including affective ones, are taken into account, before it commits itself to the most probable possibility. Finally, the integration of the information from all phoneme (cluster) positions leads to the network-like activation of the targeted word form and its conceptual meaning (with all connotative meanings being co-activated as well). While the other pre-activated word forms that did not match the targeted one get suppressed, neuronal connections from the activated word level downwards to the letter and phoneme level are still getting reinforced.

Such active neural firing within a semantic network retains the information in the working memory (Fuster & Alexander, 1971; Kubota & Niki, 1971).

If affectively inconsistent words require an effortful recruitment of additional brain resources, one might wonder why such non- or anti-iconic words exist in human vocabularies at all. It could be that the intensified neural processing of a word reflects a degree of salience. This salience may draw attention to the word's specific meaning, thereby allowing efficient conceptual distinctions between words with similar meanings in the lexicon (see Lockwood & Dingemanse, 2015). It has been shown that other factors than iconicity (usually subsumed under the label arbitrariness, also see chapter 5.4.3.) – one of them could be affective inconsistency – are coming into play in expanding lexicons to avoid the overlap of word meanings, which would easily happen if most words were based on iconicity only (Gasser, 2004; Monaghan, Christiansen, & Fitneva, 2011). For example, if a word that denotes a negative concept does not sound negative, the respective concept might not be as dangerous as a negative word concept also accompanied by a negative word sound. Considering the examples in Figure 5.1, a weapon ("Waffe" in German) is dangerous but not necessarily as deadly as a war ("Krieg") in total can be. Or if a neutral object has a negative word sound, the salience of this inconsistency might trigger attention towards a conceptual peculiarity. For example, a stocking ("Strumpf" in German) is conceptually related to a sock ("Socke"), but can be distinguished from it in particular by its length. And the word onset /str/ is actually a phonestheme indicating length (see chapter 1.4). It elicits perceptual salience mainly by the complex consonant clustering (see chapter 2) which adds to the affective inconsistency, thus hinting at the conceptual difference. In English, however, this distinction is rather made by an iconic prolonging of the word (*sock* → *stocking*; see chapter 5.4.3. for the discussion of differential developments of iconic patterns in different languages).

All in all, the significant interaction effect between the *lexical affective content* and the *sublexical affective potential* is not only highly interesting due to its potential functionality in language processing, it also corroborates the existence of a neurophysiological correlate of the *sublexical affective potential* per se – because something that does not exist cannot not interact with something else. This just being mentioned as the rather weak EPN finding of the affective comparison of the sublexically manipulated groups (see chapter 2) might not be strong enough evidence. Although the information from the sublexical features of a word might be very weak, throughout the many stages of neuronal reinforcements and interpretations it might still be able to influence perception in a subliminal way.

5.3. The role of sublexical affectivity in affective poetry perception

To examine the influence of *sublexical affective values* on real-life text processing instead of a laboratory setting where single words have to be read while wearing an EEG cap with electrodes on one's head, a poetry reading study was conducted. Reading poems triggers highly affective processes in the brain. By the artful organization of words and sounds in rhythm, rhymes, or alliterations experiences of beauty and awe can arise in the reader. In its own way, poetry unifies "thought, language, music, and images with play, pleasure, and emotion" (Jacobs, 2015a). Hence, it is a good starting point to look for interactive effects of lexical and sublexical affectivity.

For affective comparisons, texts of different affective content are needed. Therefore, a poem collection by H. M. Enzensberger (1957), which the author had subdivided into chapters of happy, sad, and spiteful poems (57 poems in total), seemed to be a suited corpus to work with. To validate the author's categorization on the one hand and to create dependent variables for the investigation of the influence of sublexical and further text levels onto the general affective perception of the poems, an online rating study of the poems was conducted.

More than 200 people rated the poems for their *general affective meaning* on eight scales capturing different aspects of meaning: valence and arousal as the basic affective dimensions; friendliness, sadness, and spitefulness as the emotional categories of the author; poeticity and liking as higher aesthetic evaluations; and onomatopoeia as a measure of how much the readers perceive the match of sound and meaning in the poems.

As it is known from text sentiment analyses, a major part of the *general affective meaning* of a text can already be predicted from the affective meanings of the words in it (Anderson & McMaster, 1982; Bestgen, 1994; Paltoglou, 2014; Thelwall, Buckley, Paltoglou, & Cai, 2010). Now that we know that there is even affective meaning below this lexical level, the question arises, how much of the *general affective meaning* could already be predicted by this affectivity at the sublexical level. Thus, in a first approach, such a potentially direct influence of sublexical affectivity onto the *general affective meaning* was investigated using a novel way of quantifying sublexical affectivity as the *basic affective tone* of a poem (Aryani et al., 2016). As fully explicated in their publication, the *basic affective tone* uses the *sublexical affective values* (SAV) in the context of foregrounding (Hakemulder, 2004; Miall & Kuiken, 1994; van Peer, 1986), i.e., only taking into account deviant phonological segments which appear in a poem with a higher frequency of occurrence than would be expected from standard language use. Such a higher than expected frequency is assumed to create perceptual salience, i.e., shifting the readers attention to these phonological segments and their affective properties. However, this is referring to a rather subliminal attention level regarding sublexical feature processing, perhaps comparable to the N400-related salience created by affective inconsistencies between the lexical and the sublexical level as shown in chapter 3. Thus, the reader would not explicitly become aware of salient phonological segments. Yet, overall throughout the reading experience, the *sublexical affective values* of all salient phonological segments in the poem together may create a subtle impression of the affective sound qualities in the reader's perception.

To determine the salient phonological segments, each poem was analyzed by the computational tool EMOPHON (Aryani et al., 2013), which compares the actual frequency of occurrence of each subsyllabic segment to its expected frequency of occurrence derived from a reference linguistic corpus. Segments which occurred significantly more often than expected were treated as salient and only their *sublexical affective values* were taken into consideration. However, not simply the mean of the *sublexical affective values* of the salient phonological segments in a poem is referred to as the *basic affective tone*, but a sigma factor indicating how much this salient SAV mean deviates from a null model in which the same number of salient segments in a poem could be loaded with randomly drawn combinations of *sublexical affective values* – representing expected values of sublexical affectivity (see the statistical explanations for this approach in Aryani et al., 2016). Hence, the *basic affective tone* is constituted by *Salient-SAV-Sigma* measures for arousal as well as valence.

Getting back to the question, how much of the *general affective meaning* of a text can already be predicted by the *basic affective tone*, Aryani et al. (2016) found that in particular sublexical arousal seems to be related to the overall emotional perception of a poem. Among the sad and spiteful poems, *Salient-SAV-Sigma* measures for arousal had higher means (i.e., stronger deviations from to be expected arousal values) than among the friendly poems. Also, the sublexical arousal measures showed higher means among the poems rated as high-arousing compared to low-arousing poems. *Salient-SAV-Sigmas* means of valence only were significantly higher in poems with a positive compared to a negative *general affective meaning*. In stepwise regressions, 9.5 to 22 % of the variance in the ratings of the *general affective meaning* of the poems were accounted for by the *Salient-SAV-Sigmas* of arousal, with the highest percentage in the model predicting spitefulness. All in all, sublexical arousal, in particular when deviating from an expected level, proved to be significantly influencing the affective perception of poetry.

So much for the astonishingly big role of the *basic affective tone* in constituting the *general affective meaning*. Yet, what about the interaction of affective properties of different text levels? It might be the case that the effects of the *basic affective tone* are caught up in the large predictive potential of lexical affective meaning. Yet, following Jakobson's (1980) view that characteristics from different hierarchical text levels vividly interact, I had hypothesized that the *general affective meaning* of the poems should be influenced by affective features from different text levels. Hence, in Ullrich, Aryani, Kraxenberger, Jacobs, & Conrad (2017, here chapter 4) the poem ratings were modelled as functions of sublexical, lexical, and inter-lexical properties of the texts in a stepwise multiple regression approach. At the sublexical level, not only the *basic affective tone* was used (Aryani et al., 2016) but also the number of salient phonological segments only as well as the sigma factors for valence and arousal of all – not only the salient – segments. Further, we differentiated between the specific types of subsyllabic segments – onsets, nuclei, and codas. Lexical affectivity was not directly operationalized by averaging the valence and arousal ratings of the words taken from the extended BAWL database (Schmidtke & Conrad, 2018), but by respective sigma values reflecting the deviation of their means from an expectable lexical affective level for texts of an equivalent length (for details see the methods part of Ullrich et al., 2017, here chapter 4.3.). And for the inter-lexical level, i.e., potential fluctuations of the lexical affectivity throughout the poem, lexical valence and arousal minima and maxima as well as ranges and the rise or decline of affectivity from the start to the end of a poem were used as measures.

Looking at the results of the several multiple regression analyses, the affective properties from all three text levels together account for 43 to 70 % of the variance in the affective ratings of the poems (Ullrich et al., 2017). Up to 50 % of the variance can be accounted for by lexical affective properties, but sublexical and inter-lexical measures significantly add up on this, which reflects the predicted interplay of the different text levels. All by itself, the *basic affective tone* can account for up to 10 % of the variance in affective poem ratings, which underlines the psychological reality of sublexical affective properties and that they play a relevant role in the perception of poetry. Hence, the sublexical affectivity is not caught up in the affectivity of higher language levels, but they all interactively exert their own influence on the *general affective meaning* of the poems.

An interesting aspect of this study is the role that sublexical saliency plays in affective poetry perception: It were mostly the *basic affective tone* values (focusing on foregrounded phonological segments only) that appeared as significant predictors in the stepwise regression models, but rarely the sigma factors of the *sublexical affective values* of all phonological segments in a poem. Furthermore, the number of salient segments alone (without their affective values attached) correlated positively with the poeticity and onomatopoeia ratings. Hence, the readers of poetry detect and evaluate such statistical irregularities from standard phoneme usage, even if at a rather implicit level. I assume these foregrounding effects to be based on prediction errors in the brain (see chapter 1.9), i.e., unexpected language patterns require enhanced neural processing effort to assign meaningful relevance to them (also compare section 5.2). In the context of poetry consumption, hence, foregrounding aspects at all language levels, including the sublexical one, seem to contribute to the poetic evaluation of a poem as was predicted by the Neurocognitive Poetics model (NCPM; Jacobs, 2015a, 2015b).

Looking at the EEG studies on affective phonological iconicity (chapters 2 and 3), one might wonder why the *sublexical affective potential* of the words was not used in the poetry study (i.e., the mean of the *sublexical affective values* of all subsyllabic segments in a word) to operationalize sublexical affectivity in the poems. As we have learned from study 2, the *sublexical affective potential* can be crucially influenced by specific rare or structurally complex phonological segments. Averaging the *sublexical affective values* of all phonological segments in a word might thus reduce the distinct affective impact, i.e., the foregrounding effect, of these segments. As specifically poets are known to pay attention to the "right" choice of particular sounds for their poems, we focused on single phonological segments in the poems' analysis. However, when thinking about sentiment analyses of prose texts, where writers may not particularly choose words with regard to their sound, one might consider to compare both approaches. Or further, one could even calculate some kind of consistency score for each word – reflecting the match of lexical and sublexical affectivity – and investigate how the use of inconsistent words influences the *general affective meaning* of a text passage. All in all, such *sublexical affective values* – as can be derived from affective word corpora – offer plenty of research opportunities, for example, but not only, for sentiment analyses of poetic as well as prose texts.

The literary genre poetry may allow easy access to affective text perception. However, in an applied reading task, further context aspects might influence how deeply a text or poem is processed, and thus, how emotionally sensitive the reading experience gets. Specifically, subjective aspects of the readers may come into play here, e.g., how attentive one is, how immersed one gets, what literary preferences and experiences one has, etc. Reading a poem at home, relaxed on the sofa with a glass of wine surely allows a deeper involvement and appreciation of the text than at the computer in the context of a research study. Furthermore, if someone is generally not interested in (the presented type of) poetry, salient foregrounding features can also lead to an experience of weirdness regarding the text. Thus, personal context plays an important role for the appraisal of poetry. But meta aspects like this were not taken into account in my purely text-based approach, because I assumed that people taking part in a poetry reading study are at least somewhat interested in that genre. Future studies should consider short accompanying surveys to shed light on the influence of personal reading experiences.

5.4. Implications for phonological iconicity

A very broad definition of iconicity is the resemblance of form and meaning (Dingemanse et al., 2015). Phonological iconicity in particular is concerned with language sounds making up the form of a word, as opposed to hand signs or written word forms (although these automatically activate the sound of a word in literate people). For the following discussion, it is important to make another distinction regarding two aspects of iconicity:

First, at the phonological level, iconicity is reflected by particular language sounds representing broad aspects of meaning. Empirical evidence for such sound-meaning correspondence used to come from studies using pseudowords or selected words of foreign languages (see chapter 1.3). Only recently, more studies look at the distribution of language sounds across whole language corpora to find patterns of sound-meaning mappings there as well (see chapter 1.4). However, such patterns (from all types of empirical approaches) refer to many different dimensions of meaning. In the introduction of this thesis I suggest to use affective dimensions such as valence and arousal in order to categorize and quantify this large variety of meaning dimensions in a psychologically plausible way. The *sublexical affective potential* that I derived holistically from affective sound-meaning correspondences in the German language, was further examined for its psychophysiological reality in the first study of this thesis (chapter 2). As the EEG results showed that this operationalization of iconicity is actually of an affective nature, this approach could then be used for further studies on how meaningful sounds affect language processing.

Second, at the lexical level, a word is considered iconic when its overall sound and its conceptual meaning resemble each other in some more or less obvious way. Comparing iconic and non-iconic (or even anti-iconic) words regarding their processing in our minds provides important insights about the role that iconicity plays in language processing. This is what the second part of the EEG study (chapter 3) was about. It showed that the lexical access of iconic words is facilitated, which means that the meaning encoded at the sublexical level is neurally processed and interacts with the conceptual meaning at the lexical level. That such interactions of meaning at different language levels can climb the ladder up to the text level was shown in the applied poetry study (chapter 4), which further underlined the role of salience, i.e., that not all language sounds influence higher language levels equally but that foregrounded sounds, e.g., rare ones, receive a higher attention and thus can exert a bigger influence.

5.4.1. The role of affective dimensions at the phonological level

This section is concerned with the first aspect of iconicity mentioned above – affective iconicity at the phonological level. As research on the semantic differential had shown that at the lexical level most dimensions of meaning can be accounted for by valence and arousal (Osgood, 1952; Osgood & Suci, 1955), I have assumed a similar relationship for the sublexical level. Empirical evidence that hints into this direction, comes from a study looking at cross-modal associations between different tastes or flavors and pseudowords (Crisinel, Jones, & Spence, 2012). There, subjects had to savor 12 gustatory stimuli and rated each on 22 scales. Among the scales were four with taste descriptors (*sweet – not sweet, sour – not sour,* etc.), twelve adjective pairs (*bad – good, soft – hard, active – passive, high – low,* etc.), and also four scales with contrastive pseudoword anchors derived from iconicity studies (*kiki – bouba, maluma – takete, lula – ruki,* and *decter – bobolo*). They found consistent cross-modal associations between the tastes/flavors and the pseudowords, and also conducted a principal component analysis which showed that those associations rested upon the basic dimensions valence and arousal but not potency (Crisinel et al., 2012). Hence, valence and arousal have the potential to holistically encompass most dimensions of meaning – even at the phonological level – due to their connotative resonance; they do not just represent two of many dimensions that can be represented iconically in language.

Furthermore, valence and arousal are core dimensions of several emotion theories: Wilhelm Wundt, one of the founders of modern psychology, had stated that every stimulus we perceive becomes evaluated by our nervous system, and the resulting affect determines our further behavior (Wundt, 1904). And also the Circumplex Model of Affect characterizes emotions using the dimensions valence and arousal (Russel, 1980). Research on affective word processing shows that the "appraisal of valence is an automatic process that is neither elicited nor enhanced by instructions to report the outcome of these judgments" (Jacobs, Hofmann, & Kinder, 2016). Arousal, on the other hand, has been found to increase the amount of attention that affective stimuli receive, and influences memory encoding and consolidation (Hamann, 2001; McGaugh, 2006).

At the phonological level, I could show that the *sublexical affective potential* as operationalized in this thesis has a psychophysiological reality that is comparable to that of the *lexical affective content* of words (Ullrich, Kotz, Schmidtke, Aryani, & Conrad, 2016, here chapter 2). As their mutual ERP component, an early negative potential (EPN), indexes automatic attention allocation to affective word characteristics, one can assume that an automatic affective classification takes place at the sublexical as well as at the lexical level during word processing. Thus, as valence and arousal become automatically assessed by the brain with each word stimulus, they represent psychologically plausible tools to unify the vast number of semantic dimensions at both the lexical and the sublexical level.

Yet, one answered question always poses several new ones. As a next step, it would be very interesting to disentangle the effects of sublexical valence and arousal, for example with regard to how, where, and when they get processed in the brain – in relation to each other as well as to the sensory processing of the respective stimulus being evaluated. As is known from neuro-imaging studies on affective word processing, arousal seems to be assessed first and valence slightly later (Citron, 2012; Recio et al., 2014). Sometimes, arousal differences lead to very early effects after less than 100 ms (Hofmann, Kuchinke, Tamm, Võ, & Jacobs, 2009), which is most likely rooted in a sublexical effect, as lexical meaning is not accessed before 200 ms (Kissler & Herbert, 2013). Together with information about neural processing loci assessed in fMRI studies, one can roughly say that arousal is more stimulus-driven and perceived without much effort, whereas valence rather represents higher-order cognitive evaluations (Citron, 2012). Yet, they interact in complex ways and at several points of time during word recognition.

At the sublexical level, only very few empirical findings do exist for valence and arousal processing of linguistic stimuli so far: Regarding arousal, Aryani, Hsu, & Jacobs (2018) showed with an fMRI study how the differently arousing sound of words is processed in the brain. The involved brain areas were similar to those processing nonverbal emotional expressions and affective prosody (Aryani, Hsu, et al., 2018), which favors the assumption that actual sound characteristics underly the arousal effect and not just learned statistical relations of phonemes to particular affective meanings. Furthermore, this underlines the theory that language evolved from affective vocalizations and networks (Panksepp, 2008). Hence, arousal – in linguistic as well as non-linguistic utterances – is naturally relevant for signaling importance of stimuli, negative as well as positive ones (remember the U-shaped distribution mentioned in chapter 1.5).

For sublexical valence, as far as I know, no EEG or fMRI study results have been reported apart from Ullrich et al. (2016). Yet, we did not disentangle valence from arousal effects. Hints from research on prosody point into a similar direction as our findings, i.e., that valence differences in language utterances can be detected around 200 ms after stimulus onset (Paulmann, Bleichner, & Kotz, 2013; Paulmann & Kotz, 2008; Wambacq, 2004).

However, there is a strong account for arousal being the stronger and more consistent dimension of affective iconicity at the sublexical level, whereas valence is harder to map consistently, presumably due to its higher cognitive processing level (Aryani, Conrad, Schmidtke, & Jacobs, 2018).

Yet, what about potency? In the semantic differential, it appeared as the second factor after evaluation (valence), similarly important as activity (arousal) (Osgood, 1952; Osgood & Suci, 1955). However, potency does not play a role in the Circumplex Model of Affect (Russell, 1980) as it is not describing an emotion per se but rather beliefs about the source or consequences of an emotion (Russell, 1978). The main difference between the semantic differential and dimensional theories of affect is that the first one is concerned about the stimulus characteristics as perceived by observers, whereas the latter ones are concerned with the feelings environmental stimuli elicit within the observer (Bakker, van der Voord, Vink, & Boon, 2014). In affective word ratings both approaches somehow get mixed, as subjects judge the characteristics of a given concept with regard to how it would affect themselves ("How positive/negative/arousing is X to you?"). Hence, the potency of a concept (or the dominance of the subject towards it, see Schmidtke, Schröder, et al., 2014) would only become relevant if a subject directly had to deal behaviorally with it and would thus need to evaluate how big, strong, or dangerous the matter is. However, as, for example, size is a relevant dimension in the research on sound-meaning correspondences (Sapir, 1929; Thompson & Estes, 2011) and loads onto the potency factor of the semantic differential (Osgood & Suci, 1955), it might be worth investigating potency at the sublexical level as well and how it relates to lexical meanings. For example, as size ratings do not make sense for all words in a language corpus (Winter & Perlman, 2021), one could let subjects rate potency instead and later examine whether and how the phonemes that were found to be characterizing potency relate to phonemes typically related to size – I would assume a big overlap here.

This example leads me to the general question how denotative and connotative dimensions of meaning relate to each other, in particular regarding iconicity. As discussed earlier in this section, whenever concrete sensory experiences drive an iconic relationship in language, affective evaluations in terms of valence or arousal (and presumably also potency) are automatically happening at a neural level, too. Let's take the bouba-kiki paradigm as an example: Shape is the primary perceptual quality here, i.e., represents the denotative meaning. Bar & Neta (2006) could show that sharp transitions in contour convey a sense of threat, and therefore trigger a negative bias. And threat is generally accompanied by high arousal, which has also been found to mediate the bouba-kiki effect (Aryani et al., 2020). Hence, spiky shapes should be rather negative and high-arousing, whereas round shapes are rather positive and low-arousing, which matches the sound characteristics of the respective names *kiki* and *bouba*, in particular of the consonants (see the arousal ranking of onsets derived from the BAWL-E in chapter 1.6, and also see Fort et al., 2015, for the stronger role of consonants in the bouba-kiki effect).

Yet, that most semantic properties can be associated with connotative dimensions such as valence, arousal, or potency does not mean that all types of meaning can be reduced to these affective connotations. Osgood & Suci (1955) wrote that “this is not taken to imply that these three, largely connotative factors represent an exhaustive description of the meaning space. There is evidence in our data for a large number of "specific" factors, quite possibly denotative in nature and representative of the ways in which our sensory nervous systems are capable of differentiating input signals. When used connotatively, such descriptive scales tend to rotate into one of the first three factors, but when used denotatively in judging sensorily relevant concepts such scales represent independent factors”.

Hence, connotative dimensions such as valence or arousal are not driving or mediating denotative, in particular perceptual dimensions of meaning, but exist independently of them. When iconicity research is directed at specifying the language sounds that are related to a particular dimension of meaning (for example, roughness, Winter et al., 2017), information about valence and arousal do not help much. However, whereas iconicity for perceptual denotative dimensions such as shape, size, brightness, etc., can only be detected in those parts of the vocabulary that deal with the respective dimension (e.g., see Winter & Perlman, 2021, for size iconicity, or Sidhu et al., 2021, for shape iconicity), connotative-affective iconicity can be quantified for a whole language corpus, because affective dimensions resonate with each concept. The resulting affective iconicity patterns may be more subtle and less specific in this approach, but these general operationalization of iconicity across a whole language corpus in terms of valence and arousal render holistic investigations about how iconicity affects language processing or learning possible (see section 5.4.2.).

Last in this section, I just want to stress that apart from being related to most language sounds as a connotative dimension of meaning, valence can also be a primary iconic feature of particular phonemes. For example, a strong connection between valence and some vowel sounds seems to arise from articulatory overlaps with emotional facial expressions. In particular /i:/ is related to positive valence due to a similar muscular activity as during smiling, whereas /o:/, /y/, or /ʌ/ are related to negative valence (Körner & Rummer, 2021; Rummer & Schweppe, 2019; Yu, McBeath, & Glenberg, 2021). Also for olfactory iconicity it was shown that some specific language sounds represent disgust, i.e., a rather denotative affective meaning, due to articulatory similarities of the speech sounds and facial disgust displays (Speed, Atkinson, Wnuk, & Majid, 2021). Hence, it would be interesting for future studies to investigate overlaps between the denotative and connotative role of valence, i.e., what patterns from the denotative version are conserved in the connotative version.

5.4.2. The role of iconicity in language processing

This section is concerned with the second aspect of iconicity mentioned above – affective iconicity of words, i.e., the fit of their meaning at the conceptual and the sound level. Hence, during the cognitive processing of words the semantic qualities of the lexical and the sublexical level, i.e., the word sound, need to be compared. That this comparison takes place in an interactive way (as predicted by the IAM, McClelland & Rumelhart, 1981) and that the results influence the word processing – in form of a facilitated processing of iconic words – has been shown in our EEG study focusing on affective iconicity (chapter 3).

Also, one fMRI study of a colleague had looked at the interaction of affective meaning, specifically arousal, between the sublexical and the lexical level and found parts of the amygdala that are responsible for the multimodal representation of emotions being activated when arousal values matched at both language levels (Aryani, Hsu, & Jacobs, 2019). Furthermore, a handful of behavioral studies could show that phonological iconicity influences word processing (Aryani, Conrad, et al., 2018; Aryani & Jacobs, 2018; Schmidtke & Conrad, 2018; Sidhu, Vigliocco, & Pexman, 2020). This is still a quite recently discovered function of iconicity which can only be investigated further if real words – instead of artificial pseudowords – can be used for psycholinguistic studies.

By then, iconicity had mainly been shown to facilitate language learning, not only in children (Imai et al., 2008; Perry, Perlman, Winter, Massaro, & Lupyan, 2017) but also in adults (Lockwood et al., 2016; Nygaard et al., 2009). Even great apes learn iconic gestures faster than arbitrary ones (Bohn, Call, & Tomasello, 2016).

Furthermore, iconicity is assumed to make conversations livelier and more figurative and direct (Dingemanse & Thompson, 2020; Lu & Goldin-Meadow, 2018; Perniss & Vigliocco, 2014). Specifically affective iconicity has been shown to play a role in a more direct and efficient communication, particularly in the case of threat: Adelman et al. (2018) found that word initial phonemes predict the word's valence best, so that potential danger can already be perceived before the whole word content is fully understood. The higher behavioral relevance of negative signs that they found was also relevant in our EEG study, where we had chosen to contrast_negative high-arousing *sublexical affective potential* (instead of positive high-arousal) with a neutral low-arousing *sublexical affective potential* for the same reason.

And iconicity is believed to have played an important role during language evolution (Cuskley & Kirby, 2013; Dellert, Erben Johansson, Frid, & Carling, 2021; Perlman, 2017), which goes along with the biological principle that the phylogenetic development, here individual language acquisition, parallels the ontogenetic development, i.e., language evolution (Imai & Kita, 2014; Lamendella, 1976). Several studies showed that a communication system can be built on iconicity (Fay, Arbib, & Garrod, 2013; Perlman, 2017). However, due to further evolutionary forces such as conventionalization, sound changes, etc., the iconic origin of most words became obscured (Flaksman, 2019, 2020; Gasser, 2004). Yet, the facilitatory role of iconicity for language processing might have been one factor that preserved iconic patterns to some extent (Sidhu et al., 2020). Hence, iconic alignments between the sound and meaning of words are not just potential relicts of the evolutionary origin of language, but play an active role in present-day language processing.

Concerning methodological aspects, iconicity needs to be operationalized at the word level to investigate word processing. In our EEG study this was done by contrasting the *lexical affective content*, i.e., valence and arousal ratings, with the *sublexical affective potential*, derived from iconicity patterns of phoneme segments across the BAWL-E database. Due to our orthogonal 2x2 design we created categories of iconic and non-iconic word that were either consistent or inconsistent, respectively, in their lexical and sublexical affective meanings. For other analyses that afford continuous variables, one would need to calculate a kind of consistency score.

Another – quite new – corpus-based approach, which sounds less effortful to pursue, is to collect direct ratings of iconicity for each word in a corpus. This was recently done with over 3,000 English words (Perry, Perlman, & Lupyan, 2015; Winter, Perlman, et al., 2017). Such iconicity ratings were, for example, used in studies looking at language acquisition (Massaro & Perlman, 2017) or language processing during lexical decisions (Sidhu et al., 2020). However, rating the fit of sound and meaning of natural words is a quite artificial task for people that are naïve to the subject of iconicity. And it is not apparent, on what dimensions people are making the ratings. It might be easy to rate an object that is emitting sounds itself, such as a rattle, but might be harder for shaped objects such as a triangle or a ball (where the expert eye spots the bouba-kiki pattern right away but the lay person may not), not to mention abstract concepts.

In comparison to iconicity ratings, our approach has the advantage that word ratings for valence and arousal come more naturally to people than ratings for iconicity. However, then one still needs to calculate the *sublexical affective potential* and possibly the above-mentioned consistency score between the lexical and sublexical variables. Another option to gather affective sound values would be a word sound rating (Aryani, Conrad, et al., 2018, study 2a). Although the conceptual meaning of a word cannot be completely ignored by the rater, it can be partialled out of the rating values subsequently (see methods of Aryani, Conrad, et al., 2018). Still, the degree of agreement between raters is quite low, which is why this approach can rather be seen as a quick & dirty option, however, with quite high correlations with the *phonological affective potential* – an operationalization similar to the *sublexical affective potential* (Aryani, Conrad, et al., 2018).

5.4.3. General thoughts on iconicity

<u>… and universality</u>

One question always being asked in the context of iconicity research regards the universality of the presented findings. The studies conducted for this thesis do not touch that question, as they are all concerned with phonological iconicity within the German language in particular. However, there is a bunch of evidence that points towards some universal sound-meaning patterns. Almost all languages in the world possess sound-meaning correlations in some way (Ciccotosto, 1991). But it is not absolutely clear yet, whether there are universally basic sound-meaning references stemming from our evolutionary roots that might have developed divergently over time and places into different languages but still show high correlations among languages, or if this is due to the common underlying cognitive-associative mechanism of human language but with very distinct relations between sounds and meaning in each language (see Taylor & Taylor, 1962).

Most likely, it is a mixture of both options. Recently, a study comparing the speech sounds appearing in 100 basic vocabulary items of more than 4,000 languages found some strong association patterns between word meanings and speech sounds across languages, for example, smallness being consistently related to /i:/, and nose words to /n/, or tongue words to /l/ (Blasi et al., 2016). However, they also found indications that such universal patterns emerged independently at different spaces and times, hence, were not inherited from the very first name givers. This is most likely due to the above mentioned cognitive-associative mechanisms which are the same in all human beings and lead to similar outputs if similar inputs, such as natural relationships, are given.

Often in the iconicity literature, only such universal patterns are defined as truly iconic, whereas language-specific patterns are classified as systematicities with unknown origin (Dingemanse et al., 2015). However, it could also be that different patterns of iconicity across languages arose at different points in time or were based on different environmental sources of association. For example, body-related sources (e.g., articulatory feedback or sensory input from near-body senses such as touch and taste) are common to all humans and should thus lead to universal patterns. For example, as all human beings possess the same composition of the articulatory tract, the proprioception of speech sounds should be the same across individuals and lead to similar iconic patterns in different languages, e.g., the feeling of abruptness when producing a stop sound may have led to its use in verbs for stopping behavior (Klenovšak, 2014). Hence, also the inherent connection of particular vowel sounds to facial emotion displays (which are universal according to Ekman, Sorenson, & Friesen, 1969), can be regarded as universal – e.g., remember the association between /i:/ and positive valence being mediated by the same facial muscle performing the articulation of /i:/ and smiling (see the Articulatory Feedback Hypothesis by Körner & Rummer, 2021; Rummer et al., 2014; Rummer & Schweppe, 2019).

Yet, considering the influence of environmental factors, different iconicity patterns could have emerged in different surroundings. On the one hand, relations based on natural laws – such as sounds moving closer to oneself sounding higher than sounds moving away (Doppler effect) – are the same across the globe and should also yield similar iconicity patterns in language. On the other hand, specific environments may have led to specific iconicity patterns. For example, hissing sounds may neurally activate the concept *snake* but would have only been loaded with negative and high-arousing relevance if there actually were poisonous snakes living in the area. And water-related sounds, such as pouring rain, could have had a positive or negative imprint on humans' cognition depending on whether they used to live in an area where water could be a danger (for example, in terms of flooding) or a blessing (e.g., in dry regions with little rain).

Furthermore, in the process of naming something, many ideocratic associations can come into play, which surely may have led to iconicity patterns that are specific to a language or language family. I like to use the concept *tree* as an example (also see my little invented story in the introduction), which is often used as a counter-example for a natural relatedness of sound and meaning: If the properties of the concept *tree* had influenced the word form, how come that the word forms are so different across languages? Because the properties from which name associations were derived might have been different ones – a tree has manifold features! Looking up the several possible meanings of the German translation "Baum" in a dictionary, they mostly refer to the trunk of a tree, as for example used in "Schlagbaum", a horizontally positioned trunk functioning as a barrier for people or vehicles, interestingly also labelled as "boom" in English (more a nautical/technical term). Using phonemes such as /b/, /au/ or /u:/, and /m/ to name the most massive part of a tree, being long, thick and round, sounds quite iconic to me. The French word "arbre" sounds similarly trunk-like, yet the /r/'s let it appear somehow more delicate. And indeed, the translations reveal further meanings such as "spindle" or "shaft", something made from a trunk, yet using some delicate treatment – there even is the technical English word "arbor" for objects like this. The English word "tree", on the contrary, rather refers to the branching pattern of the twigs, such as in its use as decision tree or in hierarchical informatic or phylogenetic trees – with /i:/ denoting something smaller and slenderer than the trunk. Looking up translations for the Polish "drzewo" shows that this word's meaning is mainly associated to wood, hence the material a tree is made of, which can be used for building something. My last but not least tree example is the Mandarin word "shù", which also means "bundle" or, as an adjective, "vertical". Here, one can easily picture in one's mind the diffusely sticking out branches of a treetop being bunched together as the trunk by a huge invisible hand. Thus, whatever the original name givers saw in the tree – its shape or steadiness at the bottom or top, its material, etc. – inspired them to create a word form that was iconic to this particular feature. But still, behind all that was the same underlying cognitive-associative principle of how human brains integrate external cues into language patterns.

<u>… and arbitrariness</u>

We quantified statistical correspondences between phoneme clusters and their appearance in affectively connotated words. Further, we showed that these correspondences can exert influence on the affective perception of words in the brain as well as on the affective reception of poetry. Yet, these findings do not tell us where these sound-meaning correspondences come from. From the way our brain, and thus language, works (Christiansen & Chater, 2008), it can be concluded that such correspondences must be the result of many steps of language evolution, for example associative ones.

Yet, the first principle of linguistics is still that language is arbitrary (Hockett, 1960; Saussure, 1916), which dismisses sound-meaning correspondences in language as an exceptional phenomenon. But what does "arbitrary" actually mean and how could that have evolved? By colloquial definition, "arbitrary" means "at random". Yet surely, people some 100,000 years ago did not use dice to arbitrarily choose some sounds to make up a new word when they needed one. According to de Saussure (1916), „The word arbitrary […] implies that the signal is unmotivated […] in relation to its signification, with which it has no natural connexion in reality". Yet still, I cannot imagine how unmotivated words or utterances might have emerged. Everything in language is motivated somehow, yet these reasons might be manifold and we just cannot trace back all influential factors anymore (for an approach towards that see Flaksman, 2019). Hence, for me "arbitrary" just means "we do not know the original motivation". An object or an event in our environment or even a new idea in our mind possesses an abundance of perceivable parameters (see the tree example above). Any one or a combination of a few of these parameters might have been involved in associatively naming the object or event originally. Hence, a name only represents one or a few aspects of an object, depending on the idiosyncratic associations of the name giver, which is then blurred by subsequent associative and habitual changes made by the next thousands of generations of language users, until only traces of an originally iconic sound association are left in today's word for this concept. We know from archaeological evidence that written alphabets have passed through a process from iconic pictographic signs into mostly arbitrary signs that can be used more flexibly and efficiently (Gelb, 1963), and can only assume a similar process for spoken language.

It has been shown that so-called arbitrariness offers a variety of cognitive advantages for the evolution of language such as flexibility, abstractness, generativity, or referential specificity

(Christiansen & Chater, 2008). In particular for growing vocabularies, only using iconic words would severely restrict the number of concepts that could be named, which is why iconicity decreases the larger vocabularies get, and arbitrariness takes over (Gasser, 2004). Yet again, this just means that we do not know which alternative motivational factors are taking over – it is surely not becoming random. Perhaps, just to imagine one example, sound contrasts are maximized to differentiate between similar concepts that are important to be distinguished from an evolutionary perspective. For example, a poisonous snake species needs to be named very differently compared to harmless species for people to be able to remember its name well and to react quickly when hearing the name due to its salient sound. As we saw in the first part of the EEG study (chapter 2), a complex syllable onset or a very rare sound could be used to increase the alerting potential of a word's sound. Thereby, an iconicity shift might have taken place, from one dimension, such as shape for example, to another, such as arousal in the poisonous snake example. Just because we cannot grasp all the potentially underlying reasons for the innumerable language changes that had happened since language originated, it does not mean that they are arbitrary. In my view, all established motivations for language changes so far during evolution – among them flexible iconicity patterns – are just the tip of the iceberg. Empirical findings have already moved from the most obvious sound-meaning relations to more hidden ones that can be made visible only by statistical measures. Yet, with time, many more relations will be revealed, so that perhaps finally no "arbitrariness", i.e., unknown motivations, will be left.

Interestingly, even if we are not aware of subtle patterns of iconicity in a word, our brains can still implicitly detect a match or mismatch between its sound and meaning, as the N400 results of the second part of our EEG study showed. Keeping in mind that there are words in our lexicon which are clearly inconsistent in their sound-meaning mapping, suggests that arbitrariness – in terms of no recognizable sound-meaning correspondence – is not the end of the scale but a kind of mid-point, and the opposite of iconicity would be something like anti-iconicity – each with its own distinct function for language processing that is still to be researched further.

5.5. Limitations and research directions

Language is a specifically human way of communication and unique in its way compared to other animals' communication means (as far as we understand them through research). It allows us to transfer knowledge and ideas across time and global distances, which is why humans were able to develop the complex and connected world we live in today. But although we use language every day to communicate our needs, ideas, or emotions to other people in oral, written, or even signed form, we do not know, how it has developed – neither during our own ontogenesis nor during human evolution. Research about the origin of language will always stay very speculative, for we cannot look back in time, and spoken language does not fossilize. But with sophisticated methods one can at least get closer to what might have played a role (and still plays) during the evolutionary process.

Our approach, however, was not of a retrospective kind but looked at a contemporary corpus of the German language. In there, we found systematic sound-meaning correspondences, and from these, we derived *sublexical affective values* – in terms of valence and arousal – for each phoneme segment based on its appearance in words with specific affective connotations. In opposition to traditional iconicity or sound symbolism research, we did not quantify the iconic potential of language sounds in a bottom-up manner, for example in reference to their pitch or formant frequencies. Hence, our top-down approach does not represent a form of perceptual iconicity but could rather been described as conventional iconicity. For this reason, not all of the values or phoneme segment orders (see chapter 1.6) seem to conform to previous evidence from iconicity research at first sight. However, previous findings also contradict each other sometimes and we do not know all evolutionary factors yet, that may have led to affective connotations of language sounds. Thus, apparently deviant patterns should be researched in more depth, e.g., in different dialects of a language or across related languages.

Furthermore, the affective word database we used (the BAWL-E, Schmidtke & Conrad, 2018) might not be perfectly representative for the German lexicon. The approximately 6,000 words had originally been chosen from the CELEX database for the purpose of research on affective words and not iconicity. Thus, the numbers of positive, negative and neutral words are roughly similar (to provide a balanced choice of stimulus material for experiments) which might not be the case in everyday language. Also, around one third of the words are composita or contain pre- or suffixes, which is not an ideal premise for corpus-based calculations, as some particular phoneme segments might be overrepresented due to their appearance in frequent pre- or suffixes, for example. However, our approach was a first attempt into this direction of quantifying iconicity and surely needs refinement. Yet, that we nonetheless got effects for the *sublexical affective potential* and its interaction with the *lexical affective content* speaks for the validity of the operationalization.

Although I argue in this thesis that affective dimensions of meaning are suited to quantify iconicity holistically across a whole language corpus, the *sublexical affective potential* does not represent iconicity directly (as captured by iconicity ratings) but still only two dimensions of meaning, valence and arousal, at the level of language sounds. Yet, as delineated in chapter 1.5, those two dimensions have the potential to encompass most of the other dimensions of meaning in a connotative way. But still, interrelations between denotative and connotative dimensions of meaning surely deserve further research at different language levels and regarding potential interactions (also see chapter 5.4.1.). Also, the interplay between valence and arousal in terms of their role in iconicity needs to be investigated further.

Another point of criticism refers to the EEG study where we presented the manipulated word stimuli visually. Phonological iconicity is about language sounds, which is why an acoustic presentation seems obvious. But then several other factors need to be considered, such as how to produce these acoustic stimuli – by human or an artificial voice? With human voices one has to decide between male or female voices or create a controlled balance. Also voices per se lead to automatic evaluations in the listener – mostly negative ones in the case of artificial voices. And also factors like intonation and prosody carry affective information. To control for all these potentially confounding variables in addition to the already very long list of variables that needed to be controlled for would have been almost infeasible. Hence, visual presentation was chosen for the EEG stimuli, as it is known from reading research that the visual presentation of a word also activates the sound of a word mentally in form of a more or less conscious inner voice (Abramson & Goldinger, 1997; Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009). This can be explained by contextual learning – when learning the written word form of a concept, it becomes integrated into the neural network that comprises the acoustic representation of that word as well as semantic, grammatical, and further contextual properties related to that word concept – as I demonstrated before by means of the DFM model (chapter 1.8) and the IAM (chapter 1.9).

However, when seeing the written word forms, potential iconic relations between the shape of a letter and properties of its sound could also act as confounding factors. For example, “b” or “o” have rounded features and were empirically shown to signify round shape, such as in the bouba-kiki comparison, where bouba denotes the curvy shape. Presumably, such iconic relations might have developed due to mental associations between sound, meaning, and written form, and thus potentially go into the same direction as the sound-meaning correspondences (see, e.g., Jee, Tamariz, & Shillcock, 2018, showing that the visual forms of letters and their pronunciations are systematically related). Hence, these relations might have enhanced the resulting effects slightly, but this is actually not bad for a pioneering study like ours, where it was quite doubtful whether we could detect such presumably very small effects at all. It should be the aim of future studies to disentangle the influences of visual and acoustic features of words on affective iconicity.

Furthermore, future studies should extend the concept of affective iconicity to other languages, which would then show which affective connotations of phonemes or phoneme clusters are universal and which vary from language to language.

5.6. Conclusion

Most studies on iconicity in language start with challenging the principle of arbitrariness in language. Yet, as they mainly look at iconicity in one dimension of meaning and often use artificial pseudowords, they finally cannot conclude more than to suggest that iconicity might be "a necessary linguistic principle rather than an occasional, exceptional one" (Nuckolls, 1999).

Focusing on the affective dimensions valence and arousal, which connotatively encompass the majority of semantic dimensions according to the semantic differential, allowed us to detect subtle affective sound-meaning correspondences across a whole corpus of German words being rated for valence and arousal. The *sublexical affective potential* represents an operationalization of affective iconicity based on these holistic sound-meaning correspondences, and has been verified by showing that it reflects affective word processing at the phonological level and that it interacts with affectivity at other language levels. These insights demonstrate that iconicity actually is a basic mechanism of language, "a necessary linguistic principle" in Nuckoll's words.

And although iconicity expresses itself in all kinds of meaning dimensions, valence and arousal turned out to be useful dimensions to operationalize iconicity holistically due to their connotative characteristics. They will be helpful tools in future iconicity studies to either narrow down the search for iconic words in a whole language corpus, or for investigations on the relevance of iconicity for language processing. The findings of this thesis just represent the starting signal for further promising research on phonological iconicity.

I want to finish my exciting journey through the realm of affective iconicity with another statement from Otto Jespersen, who had paved the way for this research 100 years ago: "The development of our ordinary speech has been largely an intellectualization, and the emotional quality which played the largest part in primitive utterances has to some extent been repressed; but it is not extinct, and still gives a definite coloring to all passionate and eloquent speaking and to poetic diction. Language, after all, is an art – one of the finest of arts" (Jespersen, 1922, p. 441).

6. References

Abramson, M., & Goldinger, S. D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics*, *59*(7), 1059–1068. https://doi.org/10.3758/BF03205520

Ackrill, J. L. (1963). *Aristotle's Categories and De Interpretatione*. Oxford University Press.

Adelman, J. S., Estes, Z., & Cossu, M. (2018). Emotional sound symbolism: Languages rapidly signal valence via phonemes. *Cognition*, *175*, 122–130. https://doi.org/10.1016/j.cognition.2018.02.007

Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Systems Studies*, *38*(1/4), 298–348. https://doi.org/10.12697/SSS.2010.38.1-4.11

Anderson, C. W., & McMaster, G. E. (1982). Computer assisted modeling of affective tone in written documents. *Computers and the Humanities*, *16*(1), 1–9. https://doi.org/10.1007/BF02259727

Aryani, A., Conrad, M., Schmidtke, D., & Jacobs, A. M. (2018). Why "piss" is ruder than "pee"? The role of sound in affective meaning making. *Plos One*, *13*(6), e0198430. https://doi.org/10.1371/journal.pone.0198430

Aryani, A., Hsu, C. T., & Jacobs, A. M. (2018). The sound of words evokes affective brain responses. *Brain Sciences*, *8*, 94. https://doi.org/10.3390/brainsci8060094

Aryani, A., Hsu, C. T., & Jacobs, A. M. (2019). Affective iconic words benefit from additional sound–meaning integration in the left amygdala. *Human Brain Mapping*, *40*(18), 5289–5300. https://doi.org/10.1002/hbm.24772

Aryani, A., Isbilen, E. S., & Christiansen, M. H. (2020). Affective Arousal Links Sound to Meaning. *Psychological Science*, *31*(8), 978–986.

https://doi.org/10.1177/0956797620927967

Aryani, A., & Jacobs, A. M. (2018). Affective congruence between sound and meaning of words facilitates semantic decision. *Behavioral Sciences*, *8*, 56. https://doi.org/10.3390/bs8060056

Aryani, A., Jacobs, A. M., & Conrad, M. (2013). Extracting salient sublexical units from written texts: “Emophon,” a corpus-based approach to phonological iconicity. *Frontiers in Psychology*, *4*(OCT), 1–15. https://doi.org/10.3389/fpsyg.2013.00654

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. (2016). Measuring the basic affective tone of poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(2), 191–204. https://doi.org/10.1037/aca0000033

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex*, *63*, 196–205. https://doi.org/10.1016/j.cortex.2014.08.025

Atkin, A. (2013). “Peirce’s Theory of Signs.” In *The Stanford Encyclopedia of Philosophy* (Summer2013 ed.). Edward N. Zalta (ed.). http://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/

Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, *332*, 346–349. https://doi.org/10.1126/science.1199295

Auracher, J. (2017). Sound iconicity of abstract concepts: Place of articulation is implicitly associated with abstract concepts of size and social dominance. *PLoS ONE*, *12*(11), 1–25. https://doi.org/10.1371/journal.pone.0187196

Auracher, J., Albers, S., Zhai, Y., Gareeva, G., & Stavniychuk, T. (2011). P is for happiness, N Is for sadness: Universals in sound iconicity to detect emotions in poetry. *Discourse Processes*, *48*(1), 1–25. https://doi.org/10.1080/01638531003674894

Bakker, I., van der Voord, T. J. M., Vink, P., & Boon, J. De. (2014). Pleasure, Arousal, Dominance: Mehrabian and Russell revisited. *Current Psychology*, *33*(3), 405–421. https://doi.org/10.1007/s12144-014-9219-4

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289. https://doi.org/10.1016/j.tics.2007.05.005

Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science*, *17*(8), 645–648. https://doi.org/10.1111/j.1467-9280.2006.01759.x

Bayer, M., Sommer, W., & Schacht, A. (2010). Reading emotional words within sentences: The impact of arousal and valence on event-related potentials. *International Journal of Psychophysiology*, *78*(3), 299–307. https://doi.org/10.1016/j.ijpsycho.2010.09.004

Bergen, B. K. (2004). The Psychological Reality of Phonaesthemes. *Language*, *80*(2), 290–311. http://www.jstor.org/stable/4489664%0Ahttp://about.jstor.org/terms

Berlin, B. (1994). Evidence for pervasive synesthetic sound symbolism in ethnozoological nomenclature. In Leanne Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound Symbolism* (pp. 76–93). Cambridge University Press.

Berlin, B. (2006). The First Congress of Ethnozoological Nomenclature. *Journal of the Royal Anthropological Institute*, *12*, S23–S44. https://doi.org/10.1111/j.1467-9655.2006.00271.x

Bestgen, Y. (1994). Can emotional valence in stories be determined from words? *Cognition and Emotion*, *8*(1), 21–36. https://doi.org/10.1080/02699939408408926

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages.

Proceedings of the National Academy of Sciences, *113*(39), 10818–10823. https://doi.org/10.1073/pnas.1605782113

Bloomfield, L. (1909). A semasiologic differentiation in Germanic secondary ablaut. *Modern Philology*, *7*(2), 245–288.

Bloomfield, L. (1910). A semasiologic differentiation in Germanic secondary ablaut – concluded. *Modern Philology*, *7*(3), 345–382.

Bohn, M., Call, J., & Tomasello, M. (2016). Comprehension of iconic gestures by chimpanzees and human children. *Journal of Experimental Child Psychology*, *142*, 1–17. https://doi.org/10.1016/j.jecp.2015.09.001

Boissy, A., Manteuffel, G., Jensen, M. B., Moe, R. O., Spruijt, B., Keeling, L. J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I., & Aubert, A. (2007). Assessment of positive emotions in animals to improve their welfare. *Physiology and Behavior*, *92*(3), 375–397. https://doi.org/10.1016/j.physbeh.2007.02.003

Bolinger, D. L. (1950). Rime, Assonance, and Morpheme Analysis. *WORD*, *6*(2), 117–136. https://doi.org/10.1080/00437956.1950.11659374

Bouissac, P. (1995). Syntactic iconicity and connectionist models of language and cognition. In M. E. Landsberg (Ed.), *Syntactic Iconicity and Linguistic Freezes* (pp. 393–418). Mouton de Gruyter. https://doi.org/10.1515/9783110882926.393

Bradley, M. M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Semantic Differential Manikin and the Semantic Differentail. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(I), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (1999). Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. In *Technical Report C-1, The Center for Research in Psychophysiology, University of Florida*. https://doi.org/10.1109/MIC.2008.114

Brand, M., Giroux, I., Puijalon, C., & Rey, A. (2007). Syllable onsets are perceptual reading units. *Memory & Cognition*, *35*(5), 966–973.

Braun, M. (2015). Emotion and language – When and how comes emotion into words? Comment on "The quartet theory of human emotions: An integrative and neurofunctional model" by S. Koelsch et al. *Physics of Life Reviews*, *13*, 36–37. https://doi.org/10.1016/j.plrev.2015.04.006

Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, *30*(7), 1977–1989. https://doi.org/10.1002/hbm.20643

Briefer, E. F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, *288*(1), 1–20. https://doi.org/10.1111/j.1469-7998.2012.00920.x

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion effects on lexical decision response times. *PLoS ONE*, *6*(8), e23743. https://doi.org/10.1371/journal.pone.0023743

Brown, C., & Hagoort, P. (1993). The Processing Nature o f the N400: Evidence from Masked Priming. *Journal of Cognitive Neuroscience*, *5*(1), 34–44. https://doi.org/10.1162/jocn.1993.5.1.34

Brown, R. W., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. *The Journal of Abnormal and Social Psychology*, *50*(3), 388.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424.

https://doi.org/10.1027/1618-3169/a000123

Buchner, A. (1665). *Anleitung Zur Deutschen Poeterey*.

Burnage, G. (1990). *CELEX - A guide for users*.

Carling, G., & Johansson, N. (2014). Motivated language change: Processes involved in the growth and conventionalization of onomatopoeia and sound symbolism. *Acta Linguistica Hafniensia*, *46*(2), 199–217. https://doi.org/10.1080/03740463.2014.990293

Cassidy, K. W., Kelly, M. H., & Sharoni, L. J. (1999). Inferring gender from name phonology. *Journal of Experimental Psychology: General*, *128*(3), 362–381. https://doi.org/10.1037//0096-3445.128.3.362

Chang-Liang, Q. U. (2018). Sources of Otto Jespersen's Sound Symbolism From Various Disciplines: A Case of Linguistic Historiography. *Journal of Literature and Art Studies*, *8*(4), 628–634. https://doi.org/10.17265/2159-5836/2018.04.011

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, 489–558. https://doi.org/10.1017/S0140525X08005268

Ciccotosto, N. (1991). *Sound Symbolism in Natural Languages. Dissertation*. University of Florida.

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*(3), 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Comenii, I. A. (1966). *De Rerum Humanarum Emendatione Consultations Catholicae Pars V. Panglottia*.

Conrad, M., Ullrich, S., Schmidtke, D., & Kotz, S. A. (2022). ERPs reveal an iconic relation between sublexical phonology and affective meaning. *Cognition*, *226*, 105182. https://doi.org/10.1016/j.cognition.2022.105182

Crisinel, A. S., Jones, S., & Spence, C. (2012). “The sweet taste of maluma”: Crossmodal associations between tastes and words. *Chemosensory Perception*, *5*(3–4), 266–273. https://doi.org/10.1007/s12078-012-9133-9

Cuskley, C. (2013). Mappings between linguistic sound and motion. *Public Journal of Semiotics*, *5*(1), 39–62.

Cuskley, C., & Kirby, S. (2013). Synesthesia, Cross-Modality, and Language Evolution. In Julia Simner & E. Hubbard (Eds.), *The Oxford Handbook of Synesthesia* (pp. 869–899). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0043

Cutler, A., McQueen, J., & Robinson, K. (1990). Elizabeth and John: Sound patterns of men’s and women’s names. *Journal of Linguistics*, *26*, 471–482.

Dellert, J., Erben Johansson, N., Frid, J., & Carling, G. (2021). Preferred sound groups of vocal iconicity reflect evolutionary mechanisms of sound stability and first language acquisition: evidence from Eurasia. *Philosophical Transactions of the Royal Society B*, *376*, 20200190. https://doi.org/10.1098/rstb.2020.0190

Ding, J., Wang, L., & Yang, Y. (2014). The dynamic influence of emotional words on sentence processing. *Cognitive, Affective and Behavioral Neuroscience*, *15*(1), 55–68. https://doi.org/10.3758/s13415-014-0315-6

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, *19*(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., & Thompson, B. (2020). Playful iconicity: Structural markedness underlies the relation between funniness and iconicity. *Language and Cognition*, *12*, 203–224. https://doi.org/10.1017/langcog.2019.49

Eich, E., Kihlstrom, J. F., Bower, G. H., Forgas, J. P., & Niedenthal, P. M. (2000). *Cognition and*

Emotion. Oxford University Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-Cultural Elements in Facial Displays of Emotion. *Science*, *164*, 86–88. https://doi.org/10.1126/science.164.3875.86

Elsen, H. (2014). Lautsymbolik – ein vernachlässigter Forschungsgegenstand der Sprachwissenschaft. *Glottotheory*, *5*(2), 185–218. https://doi.org/10.1515/glot-2014-0016

Elsen, H. (2017a). Ist das Phonästhem eine morphologische oder eine lautsymbolische Erscheinung? *Zeitschrift Für Wortbildung*, *1*(2), 9–29.

Elsen, H. (2017b). The two meanings of sound symbolism. *Open Linguistics*, *3*(1), 491–499. https://doi.org/10.1515/opli-2017-0024

Engelkamp, J., & Rummer, R. (1998). The Architecture of the Mental Lexicon. In A. D. Friederici (Ed.), *Language Comprehension: A Biological Perspective* (pp. 133–175). Springer-Verlag. https://doi.org/10.1007/978-3-642-97734-3_5

Enzensberger, H. M. (1957). *Die Verteidigung der Wölfe*. Suhrkamp.

Etzi, R., Spence, C., Zampini, M., & Gallace, A. (2016). When sandpaper is “kiki” and satin is “bouba”: An exploration of the associations between words, emotional states, and the tactile attributes of everyday materials. *Multisensory Research*, *29*(1–3), 133–155. https://doi.org/10.1163/22134808-00002497

Fay, N., Arbib, M., & Garrod, S. (2013). How to Bootstrap a Human Communication System. *Cognitive Science*, *37*, 1356–1367. https://doi.org/10.1111/cogs.12048

Filippi, P., Congdon, J. V., Hoang, J., Bowling, D. L., Reber, S. A., Pašukonis, A., Hoeschele, M., Ocklenburg, S., de Boer, B., Sturdy, C. B., Newen, A., & Güntürkün, O. (2017). Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: Evidence for acoustic universals. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1859), 1–9. https://doi.org/10.1098/rspb.2017.0990

Firth, J. R. (1930). *Speech*. Oxford University Press.

Flaksman, M. (2019). From ie *u̯e to English window: on the age and number of imitative words in English. *Indo-European Linguistics and Classical Philology*, *XXIII*, 1066–1075. https://doi.org/10.30842/ielcp230690152379

Flaksman, M. (2020). Pathways of de-iconization: How borrowing, semantic evolution, and regular sound changes obscure iconicity. In P. Perniss, O. Fischer, & C. Ljungberg (Eds.), *Iconicity in Language and Literature, Vol. 17: Operationalizing Iconicity* (pp. 76–103). John Benjamins Publishing Company. https://doi.org/10.1075/ill.17.05fla

Fónagy, I. (1961). Communication in Poetry. *WORD*, *17*(2), 194–218. https://doi.org/10.1080/00437956.1961.11659754

Fónagy, I. (1963). *Die Metaphern in der Phonetik*. Mouton.

Fort, M., Martin, A., & Peperkamp, S. (2015). Consonants are More Important than Vowels in the Bouba-kiki Effect. *Language and Speech*, *58*(2), 247–266. https://doi.org/10.1177/0023830914534951

Frishberg, N. (1975). Arbitrariness and Iconicity: Historical Change in American Sign Language. *Language*, *51*(3), 696–719. https://www.jstor.org/stable/412894

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Fritsch, N., & Kuchinke, L. (2013). Acquired affective associations induce emotion effects in word recognition: An ERP study. *Brain and Language*, *124*(1), 75–83. https://doi.org/10.1016/j.bandl.2012.12.001

Fuster, J. M., & Alexander, G. E. (1971). Neuron Activity Related to Short-Term Memory. *Science*, *173*(3997), 652–654. https://doi.org/10.1126/science.173.3997.652

Gabelentz, G. von der. (2016). *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse* (M. Ringmacher & J. McElvenny (eds.)). Language Science Press. https://doi.org/10.17169/langsci.b97.143

Gallace, A., Boschin, E., & Spence, C. (2011). On the taste of "Bouba" and "Kiki": An exploration of word-food associations in neurologically normal participants. *Cognitive Neuroscience*, *2*(1), 34–46. https://doi.org/10.1080/17588928.2010.516820

Gasser, M. (2004). The origins of arbitrariness in language. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 434–439). Lawrence Erlbaum Associates, Inc. https://doi.org/DOI: 10.1098/rstb.2013.0299

Gebels, G. (1969). An investigation of phonetic symbolism in different cultures. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 310–312. https://doi.org/10.1016/S0022-5371(69)80083-6

Gelb, I. J. (1963). *A Study of Writing*. University of Chicago Press.

Glim, S., Ullrich, S., Rummer, R., & Conrad, M. (n.d.). Enhanced neural connectivity during the processing of words with affective sound-meaning inconsistencies. *In Preparation*.

Gnatchuk, H. (2015). Phonosemantic features of English and German consonants. *Glottometrics*, *30*, 1–18.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, *3*(11), 419–429. https://doi.org/10.1016/S1364-6613(99)01397-2

Greenberg, Joseph H., & Jenkins, J. J. (1966). Studies in the Psychological Correlates of the Sound System of American English. WORD, *22*(1–3), 207–242. https://doi.org/10.1080/00437956.1966.11435451

Grimm, J. (1864). Über den Ursprung der Sprache. In *Reden und Abhhandlungen, Kleinere*

Schriften, Erster Band (pp. 255–298). Ferd. Dümmlers Verlagsbuchhandlung. https://babel.hathitrust.org/cgi/pt?id=uc1.b4066066&view=1up&seq=9

Hakemulder, J. F. (2004). Foregrounding and its effect on readers' perception. *Discourse Processes*, *38*(2), 193–218. https://doi.org/10.1207/s15326950dp3802_3

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hebb, D. O. (1949). *The organization of behavior*. John Wiley & Sons.

Heise, D. R. (1966). Sound-Meaning Correlations Among 1,000 English Words. *Language and Speech*, *9*(1), 14–27. https://doi.org/10.1177/002383096600900102

Herrmann, T., & Grabowski, J. (1994). *Sprechen - Psychologie der Sprachproduktion*. Spektrum-Verlag.

Hinton, Leanne, Nichols, J., & Ohala, J. (1994a). Introduction: sound-symbolis processes. In Leanne Hinton, J. Nichols, & J. Ohala (Eds.), *Sound Symbolism* (pp. 1–12). Cambridge University Press.

Hinton, Leanne, Nichols, J., & Ohala, J. J. (1994b). *Sound Symbolism* (Leanne Hinton, J. Nichols, & J. J. Ohala (eds.)). Cambridge University Press.

Hirata, S., Ukita, J., & Kita, S. (2011). Implicit phonetic symbolism in voicing of consonants and visual lightness using Garner's speeded classification task. *Perceptual and Motor Skills*, *113*, 929–940.

Hockett, C. F. (1960). The Origin of Speech. *Scientific American*, *203*, 88–111.

Hofmann, M. J., Kuchinke, L., Tamm, S., Võ, M. L.-H., & Jacobs, A. M. (2009). Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective and Behavioral Neuroscience*, *9*(4), 389–397. https://doi.org/10.3758/9.4.389

Hsu, C.-T., Jacobs, A. M., Citron, F. M. M., & Conrad, M. (2015). The emotion potential of words and passages in reading Harry Potter - An fMRI study. *Brain and Language*, *142*, 96–114. https://doi.org/10.1016/j.bandl.2015.01.011

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580–593. https://doi.org/10.1002/wcs.142

Huey, E. B. (1908). *The Psychology and Pedagogy of Reading*. The Macmillan Company.

Humboldt, W. von. (1836). *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwickelung des Menschengeschlechts*. Königliche Akademie der Wissenschaften.

Ikegami, T., & Zlatev, J. (2007). From non-representational cognition to language. In T. Ziemke, J. Zlatev, & R. M. Frank (Eds.), *Body, Language and Mind, Vol 1: Embodiment* (pp. 241–283). Mouton de Gruyter.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B*, *369*, 20130298. https://doi.org/10.1098/rstb.2013.0298

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. https://doi.org/10.1016/j.cognition.2008.07.015

Jackendoff, R. (2006). How Did Language Begin? *Linguistic Society of America*. https://www.linguisticsociety.org/content/how-did-language-begin

Jacobs, A. M. (2015a). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, Article 186. https://doi.org/10.3389/fnhum.2015.00186

Jacobs, A. M. (2015b). Towards a neurocognitive poetics model of literary reading. In R. M. Willems (Ed.), *Cognitive Neuroscience of Natural Language Use* (pp. 135–159).

Cambridge University Press. https://doi.org/10.1017/CBO9781107323667.007

Jacobs, A. M., Hofmann, M. J., & Kinder, A. (2016). On elementary affective decisions: To like or not to like, that is the question. *Frontiers in Psychology*, *7*, Article 1836. https://doi.org/10.3389/fpsyg.2016.01836

Jacobs, A. M., Võ, M. L. H., Briesemeister, B. B., Conrad, M., Hofmann, M. J., Kuchinke, L., Lüdtke, J., & Braun, M. (2015). 10 years of BAWLing into affective and aesthetic processes in reading: What are the echoes? *Frontiers in Psychology*, *6*, Article 714. https://doi.org/10.3389/fpsyg.2015.00714

Jakobson, R. (1980). *The framework of language*. University of Michigan.

Jakobson, R., & Waugh, L. R. (1979). *The Sound Shape of Language*. Indiana University Press.

Jee, H., Tamariz, M., & Shillcock, R. (2018). The substructure of phonics: The visual form of letters and their paradigmatic English pronunciation are systematically related. *PsyArXiv*. https://doi.org/10.31234/osf.io/n85mb

Jespersen, O. (1917). Anmeldelser af F. de Saussure, Cours de linguistique générale. *Nordisk Tidsskrift for Filologi*, *6*, 37–41.

Jespersen, O. (1922). *Language: Its Nature, Development, and Origin*. George Allen & Unwin Ltd.

Jespersen, O. (1933). Symbolic Value of the vowel i. In *Linguistica. Selected Papers of O. Jespersen in English, French, and German* (pp. 283–303). Levin and Munksgaard.

Jiang, Z. qing, Li, W. hui, Liu, Y., Luo, Y. jia, Luu, P., & Tucker, D. M. (2014). When affective word valence meets linguistic polarity: Behavioral and ERP evidence. *Journal of Neurolinguistics*, *28*(1), 19–30. https://doi.org/10.1016/j.jneuroling.2013.11.001

Johnson, R. C., Suzuki, N. S., & Olds, W. K. (1964). Phonetic symbolism in an artificial language. *Journal of Abnormal and Social Psychology*, *69*(2), 233–236.

https://doi.org/10.1037/h0043851

Jowett, B. (1892). *The Dialogues of Plato, Translated into English with Analyses and Introductions* (H. Milford (ed.); 3rd ed.). Oxford University Press.

Kawahara, S., Noto, A., & Kumagai, G. (2018). Sound symbolic patterns in Pokémon names. *Phonetica*, *75*(3), 219–244. https://doi.org/10.1159/000484938

Kelly, M. H. (1992). Using Sound to Solve Syntactic Problems: The Role of Phonology in Grammatical Category Assignments. *Psychological Review*, *99*(2), 349–364.

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. In S. Anders, G. Ende, M. Junghöfer, J. Kissler, & D. Wildgruber (Eds.), *Progress in Brain Research* (Vol. 156, pp. 147–183). Elsevier. https://doi.org/10.1016/S0079-6123(06)56008-X

Kissler, J., & Herbert, C. (2013). Emotion, Etmnooi, or Emitoon? – Faster lexical access to emotional than to neutral words during reading. *Biological Psychology*, *92*(3), 464–479. https://doi.org/10.1016/j.biopsycho.2012.09.004

Kissler, J., Herbert, C., Peyk, P., & Junghofer, M. (2007). Buzzwords: Early cortical responses to emotional words during reading: Research report. *Psychological Science*, *18*(6), 475–480. https://doi.org/10.1111/j.1467-9280.2007.01924.x

Kita, S. (2008). World-view of protolanguage speakers as inferred from semantics of sound symbolic words: A case of Japanese mimetics. In N. Masataka (Ed.), *The Origins of Language* (pp. 25–38). Springer. https://doi.org/10.1007/978-4-431-79102-7_3

Kjetsaa, G. (1974). Lomonosov's sound characteristics. *Scando-Slavica*, *20*(1), 77–93. https://doi.org/doi.org/10.1080/00806767408600652

Klenovšak, D. (2014). Weak Cratylus: Aspektuelle Ikonizität. In C. Trautmann, B. Sonnenhauser, & P. Noel Aziz Hanna (Eds.), *Diskussionsforum Linguistik in Bayern:*

Interaktionen (Vol. 3, p. 23). Ludwig-Maximilians-Universität München.

Köhler, W. (1947). *Gestalt Psychology*. Liveright.

Körner, A., & Rummer, R. (2021). Articulation Contributes to Valence Sound Symbolism. *Journal of Experimental Psychology: General*, Advance online publication. https://doi.org/doi:10.1037/xge0001124

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, *114*(1), 19–28. https://doi.org/10.1016/j.cognition.2009.08.016

Kubota, K., & Niki, H. (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *Journal of Neurophysiology*, *34*(3), 337–347. https://doi.org/10.1152/jn.1971.34.3.337

Kuchinke, L., Jacobs, A. M., Grubich, C., Võ, M. L.-H., Conrad, M., & Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: an fMRI study. *NeuroImage*, *28*(4), 1022–1032. https://doi.org/10.1016/j.neuroimage.2005.06.050

Lamendella, J. T. (1976). Relations between the ontogeny and phylogeny of language: a neo-recapitulationist view. In S. R. Harnad, H. D. Steklis, & J. Lancaster (Eds.), *Annals of the New York Academy of Sciences, Vol. 280: Origins and Evolution of Language and Speech* (pp. 396–412). The New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.1976.tb25504.x

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated Attention: Affect, Activation, and Action. In P. J. Lang, R. F. Simons, & M. Balaban (Eds.), *Attention and Orienting: Sensory and Motivational Processes* (pp. 97–136). Lawrence Erlbaum Associates, Inc.

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 Effects of Prediction from Association in Single Word Contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*, 920–933. https://doi.org/10.1038/nrn2532

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural Correlates of Processing Valence and Arousal in Affective Words. *Cerebral Cortex*, *17*, 742–748. https://doi.org/10.1093/cercor/bhk024

Liotti, M., Woldorff, M. G., Perez III, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, *38*(5), 701–711. https://doi.org/10.1016/S0028-3932(99)00106-2

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology*, *6*, 1246. https://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(8), 1274–1281. https://doi.org/10.1037/xlm0000235

Lomonosov, M. V. (1748). Kratkoe rukovodstvo krasnorečiju. In *Lomonosov: Polnoe sobranie sočinenij. 1952* (pp. 91–388). AN SSSR.

Lowrey, T. M., & Shrum, L. J. (2007). Phonetic Symbolism and Brand Name Preference. *Journal of Consumer Research*, *34*(3), 406–414. https://doi.org/10.1086/518530

Lu, J. C., & Goldin-Meadow, S. (2018). Creating Images With the Stroke of a Hand: Depiction of Size and Shape in Sign Language. *Frontiers in Psychology*, *9*, Article 1276.

https://doi.org/10.3389/fpsyg.2018.01276

Lüdtke, J., & Jacobs, A. M. (2015). The emotion potential of simple sentences: additive or interactive effects of nouns and adjectives? *Frontiers in Psychology*, *6*, Article 1137. https://doi.org/10.3389/fpsyg.2015.01137

Lupyan, G., & Clark, A. (2015). Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Psychological Science*, *24*(4), 279–284. https://doi.org/10.1177/0963721415570732

Maglio, S. J., Rabaglia, C. D., Feder, M. A., Krehm, M., & Trope, Y. (2014). Vowel Sounds in words affect mental construal and shift preferences for targets. *Journal of Experimental Psychology: General*, *143*(3), 1082–1096. https://doi.org/10.1037/a0035543

Magnus, M. (1999). *Gods of the Word: Archetypes in the Consonants*. Thomas Jefferson University Press. https://books.google.de/books?id=rIFsAAAAIAAJ&dq=editions:EPwQiXSCn0UC&hl=de

Magnus, M. (2001). *What's in a Word? Studies in Phonosemantics* [Norwegian University of Science and Technology]. https://doi.org/10.1017/CBO9781107415324.004

Magnus, M. (2013). A History of Sound Symbolism. In K. Allan (Ed.), *The Oxford Handbook of the History of Linguistics* (pp. 191–208). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199585847.013.0010

Maltzman, I., Morrisett Jr., L., & Brooks, L. O. (1956). An investigation of phonetic symbolism. *The Journal of Abnormal and Social Psychology*, *53*(2), 249–251. https://doi.org/10.1037/h0048406

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch and loudness. *American Journal of Psychology*, *87*, 173–188.

Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 384–394.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45. https://doi.org/10.1146/annurev.psych.57.102904.190143

Massaro, D. W., & Perlman, M. (2017). Quantifying Iconicity's Contribution during Language Acquisition: Implications for Vocabulary Learning. *Frontiers in Communication*, *2*, Article 4. https://doi.org/10.3389/fcomm.2017.00004

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound– shape correspondences in toddlers and adults. *Developmental Science*, *9*(3), 316–322.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986). The Appeal of Parallel Distributed Processing. In D. E. Rumelhart, J. L. Mcclelland, & T. P. R. Group (Eds.), *Parallel Distributed Processing* (pp. 3–44). MIT Press.

McClelland, James L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account on Basic Findings. *Psychological Review*, *88*(5), 375–407.

McGaugh, J. L. (2006). Make mild moments memorable: add a little arousal. *Trends in Cognitive Sciences*, *10*(8), 345–347. https://doi.org/10.1016/j.tics.2006.06.001

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. MIT Press.

Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics*, *22*, 389–407. https://doi.org/https://doi.org/10.1016/0304-422X(94)00011-5

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, *96*, 143–182. https://doi.org/10.1016/j.cognition.2004.09.001

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, *55*, 259–305. https://doi.org/10.1016/j.cogpsych.2006.12.001

Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*(3), 325–347. https://doi.org/10.1037/a0022924

Newman, S. S. (1933). Further Experiments in Phonetic Symbolism. *The American Journal of Psychology*, *45*(1), 53–75.

Nielsen, A., & Rendall, D. (2011). The Sound of Round: Evaluating the Sound-Symbolic Role of Consonants in the Classic Takete-Maluma Phenomenon. *Canadian Journal of Experimental Psychology*, *65*(2), 115–124. https://doi.org/10.1037/a0022268

Nielsen, A., & Rendall, D. (2013). Parsing the Role of Consonants Versus Vowels in the Classic Takete-Maluma Phenomenon. *Canadian Journal of Experimental Psychology*, *67*(2), 153–163. https://doi.org/10.1037/a0030553

Nuckolls, J. B. (1999). The Case for Sound Symbolism. *Annual Review of Anthropology*, *28*, 225–252. https://doi.org/10.1146/annurev.anthro.28.1.225

Nuerk, H.-C., Rey, A., Graf, R., & Jacobs, A. M. (2000). Phonographic sublexical units in visual word recognition. *Current Psychology Letters*, *2*, 25–36.

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, *112*, 181–186. https://doi.org/10.1016/j.cognition.2009.04.001

Ohala, J. J. (1984). An Ethological Perspective on Common Cross-Language Utilization of Fo of Voice. *Phonetica*, *41*, 1–16.

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L.

Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound Symbolism* (pp. 325–347). Cambridge University Press.

Opitz, M. (1624). *Buch von der Deutschen Poeterey*. David Müllers Buchhandlung.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, *49*(3), 197–237.

Osgood, C. E., & Suci, G. J. (1955). Factor analysis of meaning. *Journal of Experimental Psychology*, *50*(5), 325–338.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, *114*(2), 173–186. https://doi.org/10.1016/j.jecp.2012.05.004

Paltoglou, G. (2014). Sentiment Analysis in Social Media. In N. Agarwal, M. Lim, & R. T. Wigand (Eds.), *Online Collective Action*. Springer. https://link.springer.com/chapter/10.1007/978-3-7091-1340-0_1

Panksepp, J. (2008). The power of the word may reside in the power of affect. *Integrative Psychological and Behavioral Science*, *42*, 47–55. https://doi.org/10.1007/s12124-007-9036-5

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987. https://doi.org/10.1038/nrn2277

Paulmann, S., Bleichner, M., & Kotz, S. A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, *4*, Article 345. https://doi.org/10.3389/fpsyg.2013.00345

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *NeuroReport*, *19*(2), 209–213.

https://doi.org/10.1097/WNR.0b013e3282f454db

Peirce, C. S. (1931). Principles of Philosophy. In C. Hartshorne, P. Weiss, & A. W. Burks (Eds.), *The Collected Papers of Charles Sanders Peirce*. Harvard University Press.

Perlman, M. (2017). Debunking two myths against vocal origins of language. *Interaction Studies*, *18*(3), 376–401. https://doi.org/10.1075/is.18.3.05per

Perniss, P., Thompson, R. L., Vigliocco, G., Christiansen, M. H., & Wilcox, S. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, *1*(December), 1–15. https://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B*, *369*, 20130300. https://doi.org/10.1098/rstb.2013.0300

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLoS ONE*, *10*(9), e0137147. https://doi.org/10.1371/journal.pone.0137147

Perry, L. K., Perlman, M., Winter, B., Massaro, D. W., & Lupyan, G. (2017). Iconicity in the speech of children and adults. *Developmental Science*, *00*, e12572. https://doi.org/10.1111/desc.12572

Pratto, F., & John, O. P. (1991). Automatic Vigilance: The Attention-Grabbing Power of Negative Social Information. *Journal of Personality and Social Psychology*, *61*(3), 380–391.

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *PNAS*, *103*(20), 7865–7870. https://doi.org/10.1073/pnas.0509989103

Pulvermüller, F., Moseley, R. L., Egorova, N., Shebani, Z., & Boulenger, V. (2014). Motor cognition-motor semantics: Action perception theory of cognition and communication. *Neuropsychologia*, *55*(1), 71–84. https://doi.org/10.1016/j.neuropsychologia.2013.12.002

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia - A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Rebai, M., Bernard, C., & Lannou, J. (1997). The Stroop's test evokes a negative brain potential, the N400. *International Journal of Neuroscience*, *91*(1–2), 85–94. https://doi.org/10.3109/00207459708986367

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language*, *134*, 34–43. https://doi.org/10.1016/j.bandl.2014.03.009

Rhodes, R. (1994). Aural images. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound Symbolism* (pp. 276–292). Cambridge University Press.

Rich, A. N., & Mattingley, J. B. (2010). Out of sight, out of mind: The attentional blink can eliminate synaesthetic colours. *Cognition*, *114*, 320–328. https://doi.org/10.1016/j.cognition.2009.10.003

Richards, I. A. (1932). *The Philosophy of Rhetoric*. Oxford University Press.

Rummer, R., & Schweppe, J. (2019). Talking emotions: vowel selection in fictional names depends on the emotional valence of the to-be-named faces and objects. *Cognition and Emotion*, *33*(3), 404–416. https://doi.org/10.1080/02699931.2018.1456406

Rummer, R., Schweppe, J., Schlegelmilch, R., & Grice, M. (2014). Mood Is Linked to Vowel Type: The Role of Articulatory Movements. *Emotion*, *14*(2), 246–250. https://doi.org/10.1037/a0035752

Russell, J. A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of*

Personality and Social Psychology, *36*(10), 1152–1168. https://doi.org/10.1037/0022-3514.36.10.1152

Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225–239. https://doi.org/10.1037/h0070931

Saussure, F. de. (1916). *Cours de linguistique générale* (C. Bally, A. Schehaye, & A. Riedlinge (eds.)). Payot.

Schmidtke, D., & Conrad, M. (2018). Effects of affective phonological iconicity in online language processing: Evidence from a letter search task. *Journal of Experimental Psychology: General*, *147*(10), 1544–1552. https://doi.org/10.1037/xge0000499

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in Psychology*, *5*, Article 80. https://doi.org/10.3389/fpsyg.2014.00080

Schmidtke, D. S., Schröder, T., Jacobs, A. M., & Conrad, M. (2014). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior Research Methods*, *46*(4), 1108–1118. https://doi.org/10.3758/s13428-013-0426-y

Schourup, L. (1993). Nihongo no kakikotoba-hanashikotoba ni okeru onomatope no bunpu nitsuite [On the distribution of onomatopoeias in spoken and written Japanese]. In H. Kakehi & I. Tamori (Eds.), *Onomatopia: Giongo-gitaigo no rakuen [Onomatopoeia: Paradise of mimetics]* (pp. 48–55). Keiso Shobo.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, C. H., Hamm, A. O., & Lang, P. J.

(2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion*, *18*(5), 593–611. https://doi.org/10.1080/02699930341000239

Shanahan, D. (2008). A New View of Language, Emotion and the Brain. *Integrative Psychological and Behavioral Science*, *42*, 6–19. https://doi.org/10.1007/s12124-008-9052-0

Sidhu, D. M., Deschamps, K., Bourdage, J. S., & Pexman, P. M. (2019). Does the Name Say It All? Investigating Phoneme-Personality Sound Symbolism in First Names. *Journal of Experimental Psychology: General*, *148*(9), 1595–1614. https://doi.org/10.1037/xge0000662

Sidhu, D. M., & Pexman, P. M. (2015). What's in a Name? Sound Symbolism and Gender in First Names. *PLoS ONE*, *10*(5), e0126809. https://doi.org/10.1371/journal.pone.0126809

Sidhu, D. M., Vigliocco, G., & Pexman, P. M. (2020). Effects of iconicity in lexical decision. *Language and Cognition*, *12*(1), 164–181. https://doi.org/10.1017/langcog.2019.36

Sidhu, D. M., Westbury, C., Hollis, G., & Pexman, P. M. (2021). Sound symbolism shapes the English language: The maluma/takete effect in English nouns. *Psychonomic Bulletin & Review*, *28*(4), 1390–1398. https://doi.org/10.3758/s13423-021-01883-3

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, *39*(4), 553–569. https://doi.org/10.1068/p6591

Sonesson, G. (1997). The ecological foundations of iconicity. In I. Rauch & G. F. Carr (Eds.), *Semiotics Around the World: Synthesis in Diversity. Proceedings of the Fifth International Congress of the IASS, Berkeley, June 12-18, 1994* (pp. 739–742). Mouton de Gruyter. https://doi.org/10.1201/9781420041514.sec1

Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. *I-Perception*, *4*, 239–241.

Speed, L. J., Atkinson, H., Wnuk, E., & Majid, A. (2021). The Sound of Smell: Associating Odor Valence With Disgust Sounds. *Cognitive Science*, *45*(5), e12980. https://doi.org/10.1111/cogs.12980

Svantesson, J. O. (2017). Sound symbolism: the role of word sound in meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*, e1441. https://doi.org/10.1002/wcs.1441

Tarte, R. D., & Barritt, L. S. (1971). Phonetic Symbolism in Adult Native Speakers of English: Three Studies. *Language and Speech*, *14*(2), 158–168.

Taylor, I. K., & Taylor, M. M. (1962). Phonetic symbolism in four unrelated languages. *Canadian Journal of Psychology*, *16*(4), 344–356.

Taylor, I. K., & Taylor, M. M. (1965). Another look at phonetic symbolim. *Psychological Bulletin*, *64*(6), 413–427.

Taylor, S. E., & Fiske, S. T. (1978). Salience, Attention, and Attribution: Top of the Head Phenomena. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 11, pp. 249–288). Academic Press. https://doi.org/10.1016/S0065-2601(08)60009-X

Thelwall, M., Buckley, K., Paltoglou, G., & Cai, D. (2010). Sentiment Strength Detection in Short Informal Text. *Journal of the American Society for Information Science and Technology*, *61*(12), 2544–2558.

Thompson, A. L., & Do, Y. (2019). Defining iconicity: An articulation-based methodology for explaining the phonological structure of ideophones. *Glossa: A Journal of General Linguistics*, *4*(1), 72, 1–40. https://doi.org/10.5334/gjgl.872

Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *The Quarterly Journal of Experimental Psychology*, *64*(12), 2392–2404. https://doi.org/10.1080/17470218.2011.605898

Titze, I. R. (1994). *Principles of voice production*. Prentice Hall.

Tzeng, C. Y., Nygaard, L. C., & Namy, L. L. (2017). Developmental change in children's sensitivity to sound symbolism. *Journal of Experimental Child Psychology*, *160*, 107–118. https://doi.org/10.1016/j.jecp.2017.03.004

Ullrich, S., Aryani, A., Kraxenberger, M., Jacobs, A. M., & Conrad, M. (2017). On the relation between the general affective meaning and the basic sublexical, lexical, and inter-lexical features of poetic texts-a case study using 57 Poems of H. M. Enzensberger. *Frontiers in Psychology*, *7*, Article 20173. https://doi.org/10.3389/fpsyg.2016.02073

Ullrich, S., Kotz, S. A., Schmidtke, D. S., Aryani, A., & Conrad, M. (2016). Phonological iconicity electrifies: An ERP study on affective sound-to-meaning correspondences in German. *Frontiers in Psychology*, *7*, Article 1200. https://doi.org/10.3389/fpsyg.2016.01200

Ultan, R. (1978). Size-Sound Symbolism. In J. H. Greenberg (Ed.), *Universals of Human Language, Vol. 2: Phonology* (pp. 525–568). Stanford University Press.

van Peer, W. (1986). *Stylistics and Psychology: Investigations of Foregrounding*. Routledge. https://doi.org/10.4324/9781003130437

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods*, *41*(2), 534–538. https://doi.org/10.3758/BRM.41.2.534

Võ, M. L.-H., Jacobs, A. M., & Conrad, M. (2006). Cross-validating the Berlin Affective Word List. *Behavior Research Methods*, *38*(4), 606–609.

Võ, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, *45*, 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wambacq, I. J. A. (2004). Non-voluntary and voluntary processing of emotional prosody: an event-related potentials study. *NeuroReport*, *15*(3), 555–559.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Westbury, C. (2005). *Implicit sound symbolism in lexical access : Evidence from an interference task*. *93*, 10–19. https://doi.org/10.1016/j.bandl.2004.07.006

Whissell, C. (1999). Phonosymbolism and the emotional nature of sounds: Evidence of the preferential use of particular phonemes in texts of differing emotional tone. *Perceptual and Motor Skills*, *89*(1), 19–48. https://doi.org/10.2466/pms.1999.89.1.19

Winter, B., & Perlman, M. (2021). Size sound symbolism in the English lexicon. *Glossa: A Journal of General Linguistics*, *6*(1), p.79. https://doi.org/10.5334/GJGL.1646

Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic? Iconicity in English sensory words. *Interaction Studies*, *18*(3), 443–464. https://doi.org/https://doi.org/10.1075/is.18.3.07win

Winter, B., Sóskuthy, M., & Perlman, M. (2017). *R is for rough: Iconicity in English and Hungarian surface descriptors.* Peer-reviewed poster at the Iconicity Focus Group Workshop “Types of iconicity in language use, development, and processing” at the MPI Nijmegen, Netherlands.

Wundt, W. (1904). *Völkerpsychologie* (2nd ed.). Verlag von Wilhelm Engelmann.

Yu, C. S.-P., McBeath, M. K., & Glenberg, A. M. (2021). The Gleam-Glum Effect: /i:/ Versus /Λ/ Phonemes Generically Carry Emotional Valence. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *Advance online publication.* https://doi.org/http://dx.doi.org/10.1037/xlm0001017

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Zajonc, R. B. (1984). On the Primacy of Affect. *American Psychologist*, *39*(2), 117–123. https://doi.org/10.1037/0003-066X.39.2.117

Żuchowski, R. (1998). Stops and other sound-symbolic devices expressing the relative lenght of referent sounds in onomatopeia. *Studia Anglica Posnaniensia*, *33*, 475–485. https://repozytorium.amu.edu.pl/bitstream/10593/18209/1/34zuchowski.pdf

7. List of figures

Figure 2.1: Electrode positions of the applied 10–20 system with marked topographic clusters (ROIs) as used in the analyses: red = exploratory topographic clusters, blue = EPN ROI.

Figure 2.2: ERP effects of the *lexical affective content* (top: EPN, bottom: LPC) in the *maximally manipulated* stimulus set at selected electrode sites. For the topographic maps neutral low-arousing words were subtracted from negative high-arousing words.

Figure 2.3: ERP effect of the *sublexical affective potential* in the *maximally controlled* stimulus set at selected electrode sites. For the topographic map sublexically neutral low-arousing words were subtracted from sublexically negative high-arousing words.

Figure 2.4: ERP effects of the *lexical affective content* (top: EPN, bottom: LPC) in the *maximally controlled* stimulus set at selected electrode sites. For the topographic maps neutral low-arousing words were subtracted from negative high-arousing words.

Figure 2.5: ERP effect of the *sublexical affective potential* in the *maximally controlled* stimulus set at selected electrode sites. For the topographic map sublexically neutral low-arousing words were subtracted from sublexically negative high-arousing words.

Figure 3.1: ERP interaction effect between the *sublexical affective potential* and the *lexical affective content* at selected electrode sites representative for the N400 component. For the topographic map consistent words were subtracted from inconsistent words.

Figure 3.2: Interaction matrices of the N400 interaction effect of *sublexical affective potential* and *lexical affective content* across the whole head (A) as well as across the ROI electrodes (B).

Figure 3.3: ERP effect of the *sublexical affective potential* among lexically neutral low-arousing words at selected electrode sites representative for the N400 component. For the topographic

map sublexically neutral low-arousing words were subtracted from sublexically negative high-arousing words.

Figure 3.4: ERP effect of the *sublexical affective potential* among lexically negative high-arousing words at selected electrode sites representative for the N400 component. For the topographic map sublexically neutral low-arousing words were subtracted from sublexically negative high-arousing words.

Figure 4.1: Bivariate correlations between valence ratings and four predictors: the sigma factors for lexical valence (A), the sigma factors for lexical arousal (B), the sigma factors for sublexical arousal of all salient segments (C), and the total number of salient segments per poem weighted by its length – note that the correlation gets significant after partialling out the influence of the other predictors (D).

Figure 4.2: Bivariate correlations between the absolute values of valence ratings and two predictors: the sigma factors for lexical arousal (A) and the sigma factors for sublexical arousal of all salient segments (B).

Figure 4.3: Bivariate correlations between arousal ratings and three predictors: the maximum values of lexical arousal per poem (A), the sigma factors for lexical arousal (B), and the correlation coefficients between lexical valence values and word positions in a poem – note that the correlation gets significant after partialling out the influence of the other predictors (C).

Figure 4.4: Bivariate correlations between friendliness ratings and four predictors: the sigma factors for lexical valence (A), the sigma factors for lexical arousal (B), the sigma factors for sublexical arousal of all salient nuclei (C), and the correlation coefficients between lexical

valence values and word positions in a poem – note that the correlation gets significant after partialling out the influence of the other predictors (D).

Figure 4.5: Bivariate correlations between spitefulness ratings and four predictors: the sigma factors for lexical valence (A), the absolute values of the sigma factors for sublexical arousal of all salient segments (B), the total numbers of words in a poem (C), and the sigma factors for lexical arousal (D).

Figure 4.6: Bivariate correlations between sadness ratings and four predictors: the minimum values of lexical valence per poem (A), the sigma factors for sublexical arousal of all salient segments (B), the correlation coefficients between lexical arousal values and word positions in a poem – note that the correlation gets significant after partialling out the influence of the other predictors (C), and the correlation coefficients between lexical valence values and word positions in a poem – again, note that the correlation gets significant after partialling out the influence of the other predictors (D).

Figure 4.7: Bivariate correlations between liking ratings and two predictors: the sigma factors for lexical arousal (A) and the sigma factors for sublexical arousal of all salient segments (B).

Figure 4.8: Bivariate correlations between poeticity and four predictors: the sigma factors for lexical arousal (A), the absolute values of the sigma factors for lexical arousal (B), the number of all salient segments per poem weighted by its lengths (C), and the sigma factors for sublexical arousal of all salient nuclei – note that the correlation gets significant after partialling out the influence of the other predictors (D).

Figure 4.9: Bivariate correlations between onomatopoeia ratings and four predictors: the standard deviations of lexical valence (A) and lexical arousal in each poem (B), the number of

all salient nuclei per poem weighted by its lengths (C), and the absolute values of the sigma factors for sublexical arousal of all nuclei – note that the correlation gets significant after partialling out the influence of the other predictors (D).

Figure 5.1: 2x2 design matrix of the EEG experiment presented in chapter 2 and 3.

Figure 5.2: A few of the neighbors of the node for the letter *T* in the first position in a word, and their interconnections. (From McClelland, James L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account on Basic Findings. *Psychological Review*, *88*(5), 375–407. Copyright © 1981 by *American Psychological Association*. Reproduced with permission.)

8. List of tables

9. Documentation of research data

The extended version of the BAWL database with 5,695 words rated for valence and arousal, which constitutes the basis from where the *lexical affective content* and the *sublexical affective potential* are derived, is published as supplemental material of Schmidtke & Conrad (2018) here: http://dx.doi.org/10.1037/xge0000499.supp

All calculation steps are reproducibly described in the respective "Materials and methods" sections of the three studies in this thesis (chapters 2 to 4). Of the experimental data collected in the EEG study or the poetry study, only the data used in the second EEG study (chapter 3) are publicly accessible under https://osf.io/gam9y/. I do not own the proprietary rights to publish all the raw data. I collected the data while I was employed in a research project of the *Languages of Emotion* research cluster at the Freie Universität Berlin. The research project 401 (Sound physiognomy in language organization, processing and production) was led by Dr. Markus Conrad, who needs to be contacted for access to the data, for example for further analyses of interest: maconrad@ull.es

10. Statement of originality

I herewith give assurance that I completed this dissertation independently without prohibited assistance of third parties or aids other than those identified in this dissertation. All passages that are drawn from published or unpublished writings, either word-for-word or in paraphrase, have been clearly identified as such. Third parties were not involved in the drafting of the content of this dissertation; most specifically I did not employ the assistance of a dissertation advisor. No part of this thesis has been used in another doctoral or tenure process.

Berlin, 25.08.2021

place, date Susann Ullrich

11. Own contribution to publications

Erklärung zur kumulativen Dissertation im Promotionsfach Psychologie über den Eigenanteil an den veröffentlichten oder zur Veröffentlichung vorgesehenen wissenschaftlichen Schriften innerhalb meiner Dissertationsschrift, Ergänzung zu §5 Abs. 4 Satz 1 der Allgemeinen Bestimmungen für Promotionen an der Universität Kassel vom 13. Juni 2011

Allgemeines

Name: Susann Ullrich

Institut: Institut für Psychologie, Universität Kassel

Thema: “Sound Matters – The Affective Nature of Phonological Iconicity”

Aufstellung der eingereichten Schriften

1. **Ullrich, S.**, Kotz, S. A., Schmidtke, D. S., Aryani, A., Conrad, M. (2016). Phonological Iconicity Electrifies: An ERP Study on Affective Sound-to-Meaning Correspondences in German. *Frontiers in Psychology, 7,* Article 1200. doi:10.3389/fpsyg.2016.01200

2. Conrad, M.*, **Ullrich, S.***, Schmidtke, D. S., Kotz, S. A., (2022). ERPs reveal an iconic relation between sublexical phonology and affective meaning. *Cognition, 226,* Article 105182, doi: 10.1016/j.cognition.2022.105182

   * Both authors contributed equally.

3. **Ullrich, S.**, Aryani, A., Kraxenberger, M., Jacobs, A. M., Conrad, M. (2017). Textual features as basis of the general affective meaning: Exploring the interplay of lexical affective features, dynamic inter-lexical shifts, and the basic affective tone in poetry. *Frontiers in Psychology, 7,* Article 2073. doi:10.3389/fpsyg.2016.02073

Darlegung des Eigenanteils an diesen Schriften

Zu Nr. 1:

Konzeption: in Teilen

Literaturrecherche: vollständig

Methodenentwicklung: in Teilen

Vorbereitung des Versuchsdesigns: mehrheitlich

Datenerhebung: überwiegend

Datenauswertung: vollständig

Ergebnisdiskussion: mehrheitlich

Erstellen des Manuskripts: überwiegend

Bewältigung des Review-Prozesses: vollständig

Zu Nr. 2:

Konzeption: in Teilen

Literaturrecherche: in Teilen

Methodenentwicklung: in Teilen

Vorbereitung des Versuchsdesigns: mehrheitlich

Datenerhebung: überwiegend

Datenauswertung: vollständig

Ergebnisdiskussion: in Teilen

Erstellen des Manuskripts: in Teilen

Bewältigung des Review-Prozesses: in Teilen

Zu Nr. 3:

Konzeption: in Teilen

Literaturrecherche: überwiegend

Methodenentwicklung: in Teilen

Vorbereitung des Versuchsdesigns: vollständig

Datenerhebung: vollständig

Datenauswertung: vollständig

Ergebnisdiskussion: mehrheitlich

Erstellen des Manuskripts: überwiegend

Bewältigung des Review-Prozesses: vollständig

Further publications

As a co-author I was involved in two further articles that are related to this dissertation's topic, which is why they deserve mentioning. The first one is already published, the second one is about to be submitted:

Aryani, A., Kraxenberger, M., **Ullrich, S.**, Jacobs, A. M., & Conrad, M. (2016). Measuring the basic affective tone of poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(2), 191–204. https://doi.org/10.1037/aca0000033

Glim, S., **Ullrich, S.**, Rummer, R., & Conrad, M. (in preparation). Enhanced neural connectivity during the processing of words with affective sound-meaning inconsistencies.

12. Acknowledgments

First, I want to very much thank my supervisors Prof. Dr. Ralf Rummer and Prof. Dr. Judith Schweppe, who stepped into the breach when things with my original supervision turned complicated and supported me in getting this long-time project finished.

Also, I want to thank Dr. Markus Conrad for giving me the opportunity to pursue a doctoral thesis on this exciting topic and for introducing me to the methodological aspects.

I am further much obliged to the dissertation committee, first and foremost Prof. Dr. Sonja Kotz, who has accompanied me as a mentor and co-author throughout the whole journey.

Also, I am very thankful to my colleagues from the Languages of Emotion research cluster at the Freie Universität Berlin, Gesche Schauenburg, Lea Gutz, Jens Ambrasat, Chun-Ting Hsu, Arash Aryani, David Schmidtke, Hauke Blume, Francesca Citron, Kerstin Becker, Marcel Humar, Julia Schmidt, and Niloofar Keshtiari, for their support during the ups and downs of data collection and evaluation – on the one hand by listening and giving valuable feedback, and on the other hand by distracting me during breaks or leisure meetings. Best thanks also go to the student assistants who helped with the data collections: Luisa Egle, Maren Luitjens, Mariam Murusidze, Kathrin Schreiter, Susanne Löhne, and Johannes Ecker.

I further want to thank my colleagues from later academic jobs, Martin Katzorreck, Almut Ketzer-Nöltge, Franziska Machleb, Friederike Schütte, Silke Schunack, Astrid Haase, Anne Werkle, and Renáta Péter-Szabó, who became friends and mentally supported me in keeping up my aim.

I am also very grateful to my current employer Prof. Dr. Annegret Thieken for being so understanding and giving me the necessary flexibility regarding my working hours that allowed me to finish the thesis writing in parallel to working in a very exciting research project on climate adaptation.

A big help for me with structuring and finishing the discussion section of this thesis constituted the opportunity to present my topic in the virtual seminar "IcoSem", organized by Ian Joo from the Facebook group "Linguistic Iconicity". In particular, I want to thank Bodo Winter for the continued e-mail discussions afterwards, deepening and disentangling my thoughts on the topic.

Getting to the friends' section, I send 1,000 thanks to my dear English friend Caroline, who read through the whole introduction and discussion sections correcting my language mistakes and always cheering me up in her WhatsApp messages.

Around one year before finishing my dissertation, I took part in a three-months academic writing coaching which included weekly online meetings with other participants in a so-called motivation group. My group has kept up the weekly meetings since then to continuously reflect our progress and targets regarding our dissertations, and to motivate one another. Thank you so much, Johanna, Judith, and Matthias – without these meetings I might not have made it this far!

Many thanks also go to my dissertation buddy and close friend Maria B. for our long reflection walks in Berlin parks and several online co-working sessions during these crazy pandemic times!

Last, but definitively not least, I want to warmly thank my family and friends who believed in me and the realization of this doctoral thesis over so many years – first and foremost my dearest mother, my partner Julian, as well as Astrid R., Nathalie, Katharina, Sylvie, Sara, Janina, Benno and Maria G.!