Eike U. Decker*

# Language change as a scientific construct of a probabilistically organized information system

https://doi.org/10.1515/glot-2020-2004

**Abstract:** The present paper discusses language change from an information and systems theoretical point of view, taking on a diachronic perspective. It is argued that human language has to be regarded as a probabilistically organized information system in which synchronizations of linguistic systems of individuals create unstable (dynamic, ever-changing) collective levels ("language systems"). Therefore, probabilistic organization of language processing on an individual level leads – via bottom-up structure – to probabilistic organization of language systems as a whole. If we thus regard linguistic objects like e.g. a Saussurean sign as generally unstable and defined by probability distributions even from a synchronic point of view, we must understand language change (diachronic developments) as probabilistic as well. Therefore, language change in its "classical sense" (a change in linguistic objects) has to be reinterpreted as a change in probability distribution. Nevertheless, the term *language change* and its meaning then still lack exactness regarding some details; so we have to use this term carefully and be aware of its weaknesses. With a close look at language as an information system with both a synchronic as well as a diachronic dimension, we finally have to admit that language change is a scientific construct serving as a – sometimes quite useful – simplification within the linguistic field.

**Keywords:** information theory, language change, language evolution, probabilistic linguistics, systems theory

## 1 Introduction

In consequence of the fact that human languages are systems of signs, it is clear that what is usually called *language change* either affects the relation between a signifier and a signified (as it is the case e.g. regarding sound changes, changes in spelling

***Corresponding author: Eike U. Decker**, Universität Heidelberg, Germanistisches Seminar, Lehrstuhl für Germanistische Sprachwissenschaft (Sprachgeschichte), Hauptstraße 207–209, Heidelberg, D-69117, Germany, E-mail: eike-decker.linguistics@gmx.net

or semantic changes) or the relation of one linguistic sign to another.[1] This makes language change a matter of both semiotics and systems theory. Furthermore, the language system,[2] of course, is at least two-dimensional: It consists of a synchronic level, representing the status quo of a particular language at a certain time, and a diachronic level in terms of a continuum of a language and its development over a period of time. Understanding the diachronic level as a conglomeration of many "layers" of synchronic levels representing distinct points in time, language change can be described as the process leading to a difference between two of those synchronic levels within such a conglomeration. But, from a systems theoretical point of view, how are synchronic and diachronic levels actually connected? What is the role of a single human individual when it comes to language change? What causes language change and can be regarded as a key factor?

The present paper examines these questions proposing an approach which on the one hand concentrates on the linguistic individual (i.e. normally a human being) as language changing force and on the other hand combines a probabilistic information model of language with linguistic systems theory, arguing that language as a probabilistically organized information system can only be understood by focussing on the linguistic system of linguistic individuals (idiolect[3]). If this system consists of probabilistic structures, language change can be described as a change in probability distribution rather than an actual change of linguistic objects. Thus, language change in a classical sense has to be seen as a scientific construct; nevertheless, the term remains useful under certain conditions: If we adjust it in a way that refers to probability distributions of linguistic objects and their occurrence by including our knowledge of the probabilistic organisation of language as it seems to exist in every linguistic individual.

## 2 Language change and language system

Central aspects of the two-dimensional language systems have already been mentioned. Let us now first take a closer look at the dimension of synchronic

**1** e.g. the English lexeme *red*, which is mainly referring to the colour red, formed a semantic relation to the ideology of communism as red became the colour of many communist parties all over the world in the 20th century. This relation did not exist in former times and, in some contexts, gave *red* even a new negative connotation (e.g. in the society of the USA during the Cold War), which can also be described as a semantic change (for the empiric evidence see Aitchison 2012: 101).

**2** In this paper the term *language system* is only used to refer to a linguistic system on a collective level, especially on the level of languages like English or German.

**3** If we use *idiolect* in a way that refers to the linguistic system of an individual as a whole.

levels. So what is the basic function of language at such a level? What is its nature?

## 2.1 How synchronizations make up a language system

Influenced by Karl Bühler's organon model and extending it, Roman Jakobson distinguishes six functions of language: a referential, a poetic, an emotive, a conative, a phatic, and a metalingual function (cf. Jakobson 1960: 353–357). What these functions have in common is the fact that they essentially all describe a matter of information exchange: A sender sends (linguistic) information, which may reach a receiver[4] and even if it does not, the information sent will have left the sender (e.g. acoustically encoded), which already is an interaction in a physical sense. So the basic function of language may be interpreted as exchange of information.

With a closer look at the sender (and also the possible receiver) of a linguistic information, this assumption leads to systems theory. A basic definition of the term *system* is the one by Robert Fagen and Arthur Hall (1956): "A system is a set of objects together with relationships between the objects and between their attributes. […] Attributes are properties of objects" (Fagen and Hall 1956: 18). Furthermore, a system – may be except for the "system of the universe" – is always a distinct part of reality in the abstract (cf. Schweizer 1979: 37) that has to be distinguished from its environment. The environment of a system can only be described relatively to the system (cf. Luhmann 1987: 36) and is likely to interact with the system (cf. Fagen and Hall 1956: 20). If we turn back to language and its function of information exchange, we can now interpret the sender of linguistic information as a system interacting with its environment by transmitting this information into this environment. Thus, a possible receiver can be seen as a system in the environment of the sender, who constitutes a system him-, her- or itself.

Broadening the scope, one can connect such smaller systems with the bigger system of a particular language as a whole by bringing in Joachim Herrgen and Jürgen Schmidt's model of synchronizations, which aims to explain dynamics in language (cf. Herrgen and Schmidt 2011; Schmidt 2010). To reach an exchange of linguistically encoded information and to reach from an individual to a collective level – such as e.g. a speech community – a synchronization is needed. By this term

4 I do not use the terms *addresser* and *addressee* for they suggest that the sender (the addresser) always aims to exchange information with a certain receiver (the addressee), which, of course, is not always true (the sender might aim to exchange information with a receiver who does not receive, or the sender exchanges information with a receiver who was not target of the exchange as the sender wants it to be).

Herrgen and Schmidt "mean 'the calibration of competence differences in the performance act', which results in a 'stabilization and/or modification of the active and passive competences involved' ([Herrgen and Schmidt 2011: 28])" (Schmidt 2010: 212). This happens because speakers "actively and interactively 'synchronize' their complex and distinct systems of linguistic knowledge" (Schmidt 2010: 212) to be understood or rather to raise the probability of being understood. This "elementary synchronization act just described (the calibration of the individual competencies within a single interaction) is referred to as *microsynchronization*. It underlies all of the other types" (Schmidt 2010: 212). We can also describe it as a process that synchronizes two linguistic systems, which both exist on individual levels, to a linguistic system on a collective level. The term *mesosynchronization* is used to refer to "a series of parallel acts of synchronization, performed by individuals in personal contact situations, which lead to the establishment of common context-dependent linguistic knowledge" (Schmidt 2010: 213) and those synchronization acts "are responsible for the establishment of group and context-dependent linguistic conversations and thus, in the final instance, for the formation of varieties" (Schmidt 2010: 213), whereas *macrosynchronizations* "refer to synchronization acts via which the members of a linguistic community orient themselves to a common norm" (Schmidt 2010: 214). So mesosynchronizations, meaning a synchronization of two systems on collective levels to one another, work in a bottom-up way, macrosynchronizations in a top-down way.

It is important to point out here that every system on a collective level is unstable for it is changeable every time a synchronization takes place. With regard to systems theory a collective level brought to existence through a synchronization affects the organization of the linguistic system of the individuals, who are involved in the process of synchronization. The reason for this is the fact that such individuals may get a feedback from each other about whether they think they comprehend each other or not. This "feedback effects a modification or stabilization of the applied language production strategy" (Schmidt 2010: 212) and, of course, also of the applied language comprehension strategy. As a result, the future linguistic behaviour of the individuals is influenced by every preceding linguistic act. So the linguistic systems on individual levels[5] are carrying an "image" of this collective level containing information about what seems to be helpful in order to have a "successful" conversation (or rather a successful exchange of information – i.e. an exchange in a way the partners in communication intended it to be). The process that leads to such an image can be called "imaging", but is also similar to a macrosynchronization in terms of Herrgen and Schmidt, because an

5 Meaning the linguistic systems on individual levels of those individuals, who are part of a certain collective level, they synchronized.

image of a collective level created in and carried by a linguistic system on an individual level is comparable to a system on a collective level affecting a system on an individual level.

## 2.2 The role of probability

Every linguistic information – and even every information about how, when, where (etc.) to use it – can be located in the linguistic systems on individual levels, which may synchronize with each other and thus generate a(n unstable) system on a collective level. So the key to understanding the linguistic memory on collective levels – as we assume it by regarding the diachronic dimension of language as a continuum – is to understand how language works on individual levels.

In this paper I do not want to discuss in detail what types of information should be assumed to be part of the linguistic information a linguistic system consists of, but rather how they are organized; I will do this by using a probabilistic approach. If we define the so-called "mental lexicon" not only as some sort of neural dictionary, but as a memory with all information about linguistic objects the carrier of the mental lexicon (normally a human being) bears, we will still have to add information on this information about linguistic objects to complete the information package we may identify with the elements of a linguistic system.

This additional information can be described numerically: Besides linguistic objects,[6] relations of such an object to others are significant regarding the organization of language in a linguistic system – just like a system mainly consists of its elements and the relations between them (cf. again Fagen and Hall 1956: 18). A basic linguistic model paying attention to this insight is the model of semantic networks developed by Ross Quillian (cf. especially Quillian 1967; Quillian 1968, and Quillian 1969), which got extended several times (cf. e.g. Simmons 1972; Simmons 1973). The idea is that a semantic (and linguistic) object gets most or even all of its meaning – i.e. the information it bears – from the objects it is related to. But relations cannot only be distinguished by whether they exist or not, but also by how strong they are. In artificial neural networks (ANNs) a weighting is used to simulate the intensity of a relation between two objects (neurons) by connecting every relation with a number or rather a numeric value: the bigger this value is, the stronger is the relation (cf. e.g. Rey

**6** With the term *linguistic object* I refer – depending on context – to all different possibilities of linguistic information to appear such as e.g. semes, phonemes, graphemes, morphemes, lexemes, syntactic constructions, and other compositions; thus, linguistic objects can be regarded as union of linguistic information (in terms of elements) or in some cases even as a single linguistic information itself.

and Wender 2011: 15). Such a weighting can explain why in word-association tests most people react e.g. to the English lexeme *husband* with the lexeme *wife* (Aitchison 2012: 100) and not with a lexeme like *man* or *adult*: The latter lexemes are related to *husband* as well, but apparently not as strongly as *wife*. Also other phenomena like semantic priming could be explained by such weightings, which, of course, represent a probability distribution (cf. McNamara 2005: 26); finally, we see that in semantics probability plays an important role and that even the semantic of a whole sentence is judged by its probability (cf. Cohen 2009: 279). Generally, the weighting-number (i.e. the numeric value of the weighting) of a relation will get bigger, if the related objects co-occur; the reason for this can be found in Hebb's rule of learning, often referred to as "what fires together wires together": Just as "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" (Hebb 1949: 70), the relation of co-occurring linguistic objects gets stronger every time they appear together – this connection between biological and linguistic systems is not surprising as linguistic systems are integrated in biological systems like a human brain.[7]

If the relations of linguistic objects are organized in a way that can be described by numeric values – which would allow to sort these relations by quality (represented by a probability distribution) – there can be no doubt that language as a whole is organized in a probabilistic way. As seen, probabilistic linguistics is based on biological plausibility and able to shed light on some of the main issues of linguistics (cf. e.g. Bod et al. 2003): How does both language production and perception work? Why is there language variation? And furthermore: Why and how does language change?

Regarding language production and perception, it is now to assume that a checking of the probability distribution takes place, which mainly consists of checking how likely the connection of a certain signifier with certain signified, is in a certain context. So the linguistic sign as proposed by Ferdinand de Saussure (cf. Saussure 1916: 99–105) is not stable; instead, it is represented by a probability distribution connecting either all possibilities of signifiers a certain signified can have with this signified or all possibilities of signifieds a certain signifier can have with this signifier.[8]

---

**7** But, of course, a change is just about to happen regarding the development of AIs ("artificial intelligence"), in which a linguistic system is possibly integrated as well; nevertheless those systems are based on artificial biological systems like ANNs (cf. Rey and Wender 2011: 14).

**8** This is, of course, depending on whether it is focussed on language production or language perception and on the point of view chosen.

## 2.3 Language change in a probabilistically organized system

In order to examine the nature of language change, it seems to be important to explain how the probabilistic structure of linguistic systems on individual levels, which participate in collective levels, affects the structure of systems on collective levels. Once again, the key to this is synchronization: Linguistic systems on individual levels create a new collective level by (micro)synchronization and possibly using existing images of collective levels created in the past. So if systems on collective levels exist due to certain synchronizations of systems on individual levels, which work probabilistically, and are memorized by systems on individual levels as an image, there cannot be a way that collective levels are stable and unambiguously defined; instead, they are created each time a conversation[9] takes place by probabilistic language production or perception and collapse as the conversation is over leaving just an image in the linguistic systems involved (i.e. systems on individual levels).

Probabilistic approaches to language change are not new. Examining frequent verbs in several Germanic languages Damaris Nübling (2000) noticed that there seems to be a positive connection between frequency of linguistic objects and their possible shortening (cf. Nübling 2000: 290). Similarly, William Labov observes the relevance of probabilities regarding chain shifts meaning that one (sound) change taking place probably causes certain changes to follow (cf. especially Labov 1994: 113–291). So chain shifts seem to be series of language changes, which are highly probable as every change increases the probability for the next change in the chain to happen by changing the probability distributions of the linguistic object in question. Chain shifts are possible because every language change primarily changes probability distributions regarding the probability of connections between certain signifiers and certain signifieds. And once again, such a change in probability distribution cannot be assumed to take place on a "real" collective level, but rather on an individual level and thus affecting the collective one. Therefore, in order to understand language change, we have to focus on linguistic systems on individual levels and this is what some approaches on language change already do. The so-called *Piotrowski's law*, for example, predicts a s-shaped function of language change arguing that language change starts slow after the innovation took place the first time and increases its speed of spreading while reaching new individuals (i.e. linguistic systems on an individual level), which then are starting to use the innovation and taking part in spreading it; this happens until so many individuals are participating in the language change that the speed of spreading slows down (cf. Leopold 2005: 627–628). We will see that it can be

9 Or rather an exchange of linguistic information.

helpful to emphasize the importance of systems on individual levels even more by discussing the theory of linguistic recessivity.

Finally, we should keep in mind that by talking about probability distributions the existence of a range of possibilities is implied, so the easiest case of a probability distribution is a binary one: An event either takes place or it does not.[10] Regarding language change this is consistent with the claim of Roger Lass (1980) "that linguistic change is entirely a domain of options, including the zero option" (Lass 1980: 131), in which the "zero option" corresponds to continuity (or rather "non-happening" language change).

# 3 Terminology matters: The theory of linguistic recessivity

The chapters above revealed that due to the complexity of language systems, their diverse dimensions and levels, so-called *language change* is not easy to catch. Furthermore, it becomes obvious that the centre of a language system is the linguistic systems on individual levels. Indeed, language change cannot be examined without a carefully used terminology and touching not only interactions between linguistic individuals (i.e. linguistic systems on individual levels), but also language processing. In addition to that, terms like *linguistic object* and the nature of what is meant by them need to be defined clearly. The fact that many of the linguistic terms used above were developed and established at different times and in different disciplines of linguistics leads to some confusion and uncertainty in their use, especially when they are used together. This is why a new terminological approach is needed and that is the aim of the following explanations introducing the theory of linguistic recessivity.[11]

## 3.1 What brings linguistic objects to life: A perceptual approach

What is required for a linguistic object to be considered "existent"? A possible answer could be: It has to be part of a linguistic system and recognizable within

**10** It is clear that if the probability for one event to happen is one, this cannot be called a probability distribution for there is only one possible event, which is guaranteed to happen.

**11** Those explanations are in some aspects a short-cut of my dissertation ("'Rezessive' Information in Sprache" ["'Recessive' information in language"]) at University of Heidelberg. For full justification of the presented theory I refer readers to this dissertation (i.e. Decker, n.d.).

this system through (successful) use in communication. But such an answer leads to another question: How should "recognizable" be defined in this context?

Because human language is an information system developed and used within or by mankind it is a fact that an exchange of information via language can only be possible, if linguistically encoded information is perceived by (the linguistic system of) a sender or (the linguistic system of) a receiver, or ideally both: This normally happens in an auditory or a visual way (speech perception; reading).[12] But this definition of perception fails to integrate the process of language production, suggesting that only linguistically encoded information that reaches a receiver is part of the linguistic system. To solve this problem, I propose to extend the definition of perception to include linguistically represented thoughts, arguing that such thoughts are both linguistic and perceived due to their nature of being mentally represented.[13] The idea behind this is in some aspects based on central assumptions of the linguistic approaches by Richard Montague (the so-called "Montague grammar"; cf. e.g. Montague 1970, Montague 1973). Montague assumed that "there is […] no important theoretical difference between natural languages and the artificial languages of logicians; indeed, [he considered] it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory" (Montague 1970: 373). This assumption can be applied not only to language, but also to perception: We know – especially regarding humans – that a (physical) stimulus effects a transduction, which means that a sensory receptor converts that stimulus into a neural action potential, and is normally followed by a transformation of this action potential (cf. Jäncke 2013: 209–210). So we can regard language perception as a process of transducing and transforming language encoded in form of a physical stimulus – like speech or a written text – into language encoded in form of a neural action potential[14] and thus there is no evidence for a qualitative difference between mentally (and linguistically) represented thoughts, which can

**12** Of course, not only a partner, who is involved in a process of language production, can receive linguistically encoded information; a sender can instantly be receiver of the sounds he/she/it uttered too.

**13** This guarantees for example that a linguistic utterance (e.g. of the English sentence *Help me!*) by a deaf person, which reaches no partner in communication (i.e. no receiver), also counts as a linguistic action.

**14** This is why I propose the use of an "extended definition of language" that refers to every interaction in a physical sense as a sort of communication (for an exchange of information takes place); perhaps linguistics should then be extended as well by developing a new sub-discipline of "physical linguistics" examining not only the biological foundation of linguistic systems – as biolinguistics does – but concentrating on underlying physical processes.

also be identified with action potentials, and recognized perception of a physical stimulus.[15]

Based on these assumptions, perception of a linguistic object is a mental representation of a linguistic object and thus can be effected by a physical stimulus as well as by thought – or rather: A perception of a linguistic object can be caused by processes within the prevailing linguistic system or by interactions between this system and its environment. For our understanding of linguistic objects, we could say that they are "brought to life" by being perceived. But, because we defined perception in a way that may refer to some steps in both language production and language perception, we obviously need to readjust those terms, for what is generally meant by *language perception* should not be confused with the understanding of *perception* we have just established. This is why I propose to replace the term *language perception* by *genesis of a signified* (triggered by the physical stimulus of a signifier) and the term *language production* by *genesis of a signifier* (triggered by a signified, which is meant to be uttered in order to exchange information regarding this signified).

Turning back to so-called language change one could now ask how linguistic knowledge on collective levels can be identified. If a system on a collective level has to be synchronized by systems on individual levels or rather is the product of such a synchronization, information that was never part of an interaction between linguistic systems on individual levels can never be part of a system on a collective level. This means a linguistic object in a system on a collective level has to be perceived by at least two linguistic systems on an individual level in at least one communicative interaction. Following terms of biology (especially genetics), this paper proposes to call linguistic information in a linguistic system that are actually perceived part of a *linguistic phenotype* (or *phenotype of language*)[16] of this system and opposes this phenotype to a *linguistic genotype* (or *genotype of language*) meaning all information in this linguistic system as a whole, even those that are

**15** Whereas *perception* often is defined as processing of information that comes from the environment and enters the (cognitive) system in question, the use of this term in the present paper is also referring to processing of information within a system. A distinction between *perception* and *attention* in linguistic context is not made for there seems to be no convincing study of the question whether linguistic objects are always part of only one of those cognitive processes or not; nevertheless, such analyses are always depending on questions of definition.

**16** Due to this definition the understanding of the term *linguistic phenotype* differs from those other researches in linguistics have. Stephen Anderson and David Lightfoot (2000), for example, refer to an actual genetic basis of human language by using the term *linguistic genotype* and to grammatical competences of a human individual by using the term *linguistic phenotype* (cf. Anderson and Lightfoot 2000: 702, 709).

not perceived. Such information that enters a linguistic system without being consciously perceived could be referred to as *recessive*.

The reason such a distinction is needed is the fact that linguistic information remains in the system even if it is not perceived.[17] A linguistic system on an individual level (as it can be found in a human being) is able to use linguistic information that somehow came into the system in the past without having to learn about this information again as a precondition for its use in the present or in future. Because of our memory one can use (and thus perceive) a word like *dog* today even if the last time it was perceived (e.g. said or heard) was yesterday and it was not perceived in the meantime (which also means that it could not get into the linguistic system again, if it were lost after the last perception of it). That is why I am proposing to call linguistic information that is actually not perceived "recessive" – if it is, or is considered to be, existent within the linguistic system in question. This definition makes the linguistic phenotype an unstable category; its composition is variable and depends on the perception(s) actually made at a certain time. Regarding the linguistic phenotype ($P$), the linguistic genotype ($G$), and all of the recessive information ($R$) in a linguistic system as set in mathematical sense, they – or rather their relations to one another – can be defined as follows:[18]

1. $|P| \leq |G| \geq |R|$
2. $G = P \dot{\cup} R$
3. $P \cap R = \varnothing$

So recessive linguistic information can be transferred to the linguistic phenotype by a process that leads to a perception of this information in terms of a mental representation of this information. I propose to call this process (and also its result) *phenotypification* and the linguistic system phenotypificating (or rather the carrier of this system) *phenotypificator*. Such phenotypifications can only be found on individual levels because a (human) collective cannot perceive as a whole, but individuals can (independently from each other for we e.g. do not necessarily feel, what somebody else feels).

---

**17** If this were true, this would mean that there is no memory for linguistic information at all.

**18** Regarding the first definition, it should be remarked that although it is possible in theory that $|P| = |G|$ it has to be assumed that this is never happening in reality (at least regarding human language use) for a phenotypificator would have to phenotypificate every information of his, her, or its linguistic system at the same time; a much more complex system of language processing than a human brain – and its possibilities to create an output (such as using the vocal tract) – would be necessary to do so. Therefore, a definition as follows seems to be more likely: $|P| < |G|$ or even $|G| \geq |R| > |P|$ (then assuming that phenotypifications normally phenotypificate just a small part of the elements of $G$).

## 3.2 Language in the making: Combining the perceptual approach with a probabilistic one

We now have observed the important role of perception in a language system, for it is perception on individual levels that allows linguistic objects to be recognized in an epistemological sense and, in accordance with the terminology presented above, a phenotypification is precondition for the perceptibility of a linguistic object. Due to synchronizations of systems on individual levels, a linguistic object can become part of a system on a collective level in terms of shared knowledge (e.g. via micro- and mesosynchronization) affecting the linguistic systems on individual levels involved in that synchronization (e.g. via macrosynchronization): Those systems adjust their package of linguistic information and the relations between such information, which may form the linguistic object in question, depending on the context and on whether they regard their exchange of information as successful or not (this is what this paper refers to as "imaging").

But let us now take a closer look at how linguistic objects are formed according to the theory of linguistic recessivity; once again, a combination of Saussurean semiotics, probabilistic approaches to linguistics, information theory, and systems theory is the basis of the following assumptions. The central thesis – as mentioned in Section 2.2 – is that linguistic signs are an unstable connection of a signifier and a signified, which can only be stabilized in and for a certain phenotypification, which always takes place in a certain usage, a certain context, at a certain time etc. Both the signifier and the signified can be understood as a package of smaller information (which one might identify e.g. with phonemes or semes). With regard to a linguistic object independent of a state of phenotypification, we have to assume all possible connections between signifiers on the one hand and signifieds on the other hand to be relevant at the same time. This state can be called a "state of superposition" (as adapting a physical term that describes the sum of states that are valid at the same time or at least might be valid at the same time, for a decision cannot be made [see e.g. Gerjuoy 1993: 1384]). Such a broad understanding of linguistic superposition can be interpreted as a claim that language has to be regarded as just one single sign taking new shape each time it is used. To make such an assumption helpful, one could now add the probabilistic approach already presented: A linguistic sign – free of a state of phenotypification – consists of both a set of signifiers and a set of signifieds plus a probability distribution by which is meant that every combination of a signifier with a signified has a certain probability (and the probability distribution represents all of those probabilities); such a probability

depends on the context of each phenotypification. Nevertheless, those probability distributions are – due to synchronizations as described by Herrgen and Schmidt (2011) – stabilizing on a collective level, which does not mean that full stability is reached. For example, the probability that the (English) signifier *mouse* is phenotypificated with reference to a little rodent as signified is quite high, if the phenotypificator is a speaker of English. But still, depending on the context of phenotypification there might be an even higher probability that the same signifier is phenotypificated referring to a computer mouse as signified. If we talk about this signifier on a scientific, linguistic level without regarding a certain context, we cannot decide whether the phenotypification with reference to a rodent or to a computer mouse is more likely. But we can tell that both possibilities are quite probable as long as we refer to an English context.[19] So what we see is that we need such a contextualization to make the presented approach work in a useful way. Of course, the contextualization used here can be made more precise: We can investigate probability distributions of English in a certain geographic area like Ireland or add a sociolinguistic filter by examining only the language of English speaking farmers in Ireland. But if we add no such filter, we actually cannot make another statement than arguing that language are just one single linguistic sign consisting of all possible connections between all possible signifiers on the one hand and all possible signifieds on the other hand.

Filters, as those discussed above, are a tool for researchers to create a relevant context to investigate and to get useful data regarding the probability distribution of the possibilities of phenotypification of a linguistic object.[20]

Taking a look again at the signifier *mouse* using a filter that restricts its context of appearance to the collective level of an English-speaking community, we can assume, for example, the following semantic possibilities of phenotypification to have the highest probabilities of being phenotypificated together with the signifier *mouse*: (1.) a little rodent like a house mouse (*mus musculus*), (2.) a computer mouse, and (3.) a shy person; nevertheless, other semantics are possible, but can be assumed to have a lower probability of being phenotypificated with the signifier *mouse*.[21] By using such a filter, the probability distribution for a phenotypification of a bilateral sign with the signifier *mouse* allows to

**19** In terms of referring to a linguistic system on a collective level of a community that uses a linguistic code we are used to calling "English".

**20** While being aware that most researchers use a comparable approach this argument aims to shed light on the underlying processes and assumptions.

**21** This scenario is an example; if we wanted to really investigate it, we would have to examine how often the English signifier *mouse* gets phenotypificated together with each of the semantics mentioned above as signified.

call all possible semantics to be in sum the signified with the probability distribution providing insight into the weighting of those possible semantics in relation to each other: these semantics superpose each other (i.e. they are in a state of superposition). Removing or adding a filter always affects this probability distribution.[22] As already mentioned, language can be regarded as a "single sign" structured by probability distributions, and filters are incapable of "destroying" this sign; nevertheless, filters prime the probability distribution of phenotypifications of this single sign in a way that makes it useful with regards to linguistic studies by paying attention to the context of a linguistic phenotypification in terms of mathematical or rather numeric (i.e. quantitative) language descriptions. Linguistic studies using this tool are working on probabilities and uncertainty (for they refer to superposition), while language processing in the human brain works similar, but not exactly the same way: Language processing also works probabilistically, but working with superposition can be a problem as it makes language processing more complex than it needs to be, and seems to be; this is why the human brain normally makes a decision for example on what is most likely meant by a perceived signifier (cf. e.g. the probabilistic hypothesis by Egon Brunswik, arguing that the environment of an organism is uncertain and probabilistic, which is why decisions are based on probabilism as well [cf. Wirtz 2017: 1324]); the phenotypification of the semantics of such a signifier then is done via probability matching – and is not a phenotypificating of all superposing semantics at the same time.[23]

Regarding a linguistic object free of state of phenotypification – but may be by using filters – we have to readjust our (metaphorical) claim, which was made in Section 3.1, "that they are 'brought to life' by being perceived".[24] Following the theory of linguistic recessivity, linguistic objects exist in a recessive way – i.e. as part of the linguistic genotype, but not of its phenotype – as the information they consist of is part of the linguistic system in question; that is the precondition of this information getting phenotypificated.

---

**22** If we, for example, added a filter restricting on conversations between zoologists, the probability distribution of the possible semantics above would likely change in a way that increases the probability of a rodent being the semantic information becoming signified together with the signifier *mouse* as a phenotypification actually takes place.

**23** A process that at first keeps most promising possibilities via analysing their probabilities and thus finally leads to a decision via probability matching is, for example, an algorithm called "beam search"; beam search is actually used in software that is used for natural language processing (NLP) and based on ANNs, and can be found e.g. in NLP-software by Google (cf. Andor et al. 2016: 1).

**24** I would like to append to this claim that linguistic objects can (or even have to) exist before they can be "brought to life".

# 4 Discussing "language change" with regard to the theory of linguistic recessivity

After examining probabilistic elements and structures in language processing and synchronizations of collective levels by taking a look at the basis of language in the human brain and its function of exchanging information, we can now turn back to the case of language change. The most obvious problem regarding language change is may be the fact that this term normally refers to a collective level assuming some kind of system with clear rules and constraints, which are stable for at least a longer timespan (approximately several years, decades or even longer); language change then is the process of destroying this stability. With the insights collected or made in the sections above, one has to admit that language change cannot be found in language systems over time in the way just described: We observed language to be unstable and probabilistically organized, and collective levels of linguistic systems as results of systems on individual levels being synchronized. So, do these insights allow to stick to a concept like language change?

First of all, let us take another look at the two levels of linguistic systems. By claiming that linguistic systems work in a probabilistic way we argue for them being unstable; nevertheless, macrosynchronizations – in terms of Herrgen and Schmidt (2011) – are obviously a stabilizing factor within the linguistic world: Micro- and mesosynchronizations of systems on individual levels create systems on collective levels in order to establish a linguistic code, which is likely to be understood by all systems on individual levels (i.e. individuals that carry a linguistic system) involved in a synchronization in approximately the same way (at least this is the goal), whereas macrosynchronizations use such micro- and mesosynchronizations to affect systems on individual levels in orientation to those "smaller" processes of synchronization (cf. again Schmidt 2010: 212–214; Herrgen and Schmidt 2011: 28–33). In order to use those "smaller" processes, macrosynchronizations can only rely upon images of collective levels that were created by synchronizations in the past and that can be found in linguistic systems on individual levels affected by macrosynchronization. Using images created in the past, macrosynchronizations stabilize linguistic systems over time, even though they do not prevent them from still being somehow unstable. This is why there are still good reasons for systems on collective levels to be understood as stable – at least to some (maybe even quite relevant) extent – nevertheless, the organization of such a collective level remains probabilistic, but the probability distributions of the possibilities of phenotypification of a linguistic object may

vary only little over time.[25] Therefore, we can still argue for linguistic systems on collective levels to exist in a way that allows us to stick to the concept of so-called language change, if we keep in mind that the stability of linguistic systems on collective levels is never a complete one.

But for the concept of language change the claim of linguistic systems being organized in a probabilistic way seems to be more important than the question of stability in language over time: If we assume that e.g. what is said to be the Modern English diphthong /aɪ/ is not stable – which the theory of linguistic recessivity and other probabilistic approaches to linguistics do – this means that /aɪ/ is actually just one possible and quite probable phenotypification of a linguistic object that could also be phenotypificated as /iː/ (as it usually was the case in Middle English),[26] even though the probability for this to happen is obviously not as high as it is for /aɪ/. From this perspective, the Great Vowel Shift does not describe language changes on phonological level as it is said to do according to the classical understanding of language change; but it does describe a change in the probability distribution of the possibilities of phenotypification of certain phonological objects and their probability of being phenotypificated. In the schemes normally used to illustrate the Great Vowel Shift these changes are shown only by the prevailing possibilities of phenotypification with the highest probability of the linguistic system in question[27] being named (i.e. for example Middle English /iː/ and Modern English /aɪ/ with both phonemes standing for the same linguistic object).

The theory of linguistic recessivity also allows us to analyse some special issues of historical linguistics offering new approaches. For example, the German lexeme *auch* is usually used as a particle in the meaning of 'also, as well, too' in all German dialects except for one: Transylvanian Saxon. Transylvanian Saxon is spoken by the descendants of German settlers who came to Transylvania (Romania) from the twelfth century onwards and they use *auch* not only as a particle, but also as coordinating conjunction in the meaning of 'and' (cf. Shinohara 2016: 58). Sara Shinohara argues that this additional meaning of *auch* in Transylvanian Saxon can be seen as consequence of language contact with Romanian, which uses its comparable form *şi* in the same way – i.e. in the meaning of 'and; also, as well, too' (cf. Shinohara 2016: 71–72). This conclusion seems

**25** This is the reason why it is possible for us, for example, to recognize a text written in 2018 (like this paper) as an English text and also a text written in 1980 (like the already cited work of Lass [1980]).

**26** Regarding the phonological differences between Modern English and its earlier forms (like e.g. Middle English), which are mainly summed up in the so-called "Great Vowel Shift", see e.g. Krug (2017) and Schlüter (2017: 34–44); they also show the development of Middle English /iː/ to Modern English /aɪ/ in more detail.

**27** i.e. in this case either the system on a collective level of Middle or Modern English.

plausible; nevertheless, it might be adjusted a little, for an interpretation of such a change of meaning as a result of language contact could entice into assuming that the Romanian language "added" a new meaning to Transylvanian Saxon *auch*, which this lexeme did not mean before. With a look at the history of the Transylvanian Saxon language (or rather dialect), it seems to be more plausible that this meaning ('and') has not been adopted from Romanian, but rather "reactivated". Comparative and historical linguistics show that it is very likely that Transylvanian Saxon *auch* has to be seen as a descendant of the Proto-Germanic lexeme **auke* meaning 'and' (amongst others) – as its Gothic (*auk*) and Old Norse (*ok*) descendants meant the same (see, regarding Gothic, Köbler [1989: 70–71], regarding Old Norse, Zoëga [2004: 320–321], and, regarding Proto-Germanic, Kroonen [2013: 42]). The meaning of 'and' for this lexeme obviously got lost in most of the West-Germanic languages like German in the Middle Ages – as for Middle High German *ouch*, which later became *auch* (cf. e.g. Benecke et al. 1990: 449–451).[28] So, the theory of linguistic recessivity allows to regard the signifier (**auke*, *ouch*, *auch* etc.) as connected or rather connectable with a range of possible signifieds (e.g. meanings like 'and', 'also', 'as well', and 'too') assuming no change of these possible signifieds, but rather in the probability distribution representing the connection between signifier and signified if it comes to phenotypification. This means that the probability of the signifier as linguistic object (realized/phenotypificated as **auke*, *ouch*, *auch* etc.) being phenotypificated together with the meaning 'and' was quite high in Proto-Germanic, fell in medieval German, when the probability of phenotypification of meanings like 'also', 'as well', and 'too' rose. Whereas most of the German dialects keep this change in probability distribution on their system on a collective level, Transylvanian Saxon underwent another significant change in this probability distribution, increasing the probability of phenotypification of the meaning 'and' again.

This new approach does not mean that Shinohara's assumption that language contact between Romanian and Transylvanian Saxon caused the semantic "change" in *auch* is wrong, but it changes the role Romanian plays: According to the approach based on the theory of linguistic recessivity, Romanian did not add a "new" meaning to the lexeme *auch*; nevertheless, it can be assumed that the

**28** If the conventional connection of the signifier Proto-Germanic **auke* / Middle High German *ouch* / Transylvanian Saxon *auch* and its (most probable) signified, which normally got phenotypificated together with the signifier, changed in the way described, we could compare this to the pattern of recessive inheritance as it can be found in genetics: Something that forms the phenotype stops doing that and seems to have got lost at some point, but later forms the phenotype again (because it was not actually lost). With regard to the insights already made in this paper, it is obvious that claiming such a "recessive pattern" is a simplification of a more complex development.

language contact increased the probability of phenotypification of the signifier *auch* with the meaning of 'and' in Transylvanian Saxon. Furthermore, this meaning seems to have been recessive to that signifier during the whole Middle Ages.[29]

The reason why this approach to language change caused by language contact can be regarded as more plausible, is the fact that it is based on mathematical and thus logical insights, whereas the assumption that the semantic "change" discussed was not caused by a change in probability distribution and already existing recessive information makes it a coincidence that Transylvanian Saxon *auch* got the meaning 'auch' from Romanian *şi*. It could be asked why Romanian *şi* (meaning 'and; also, as well, too') led to an addition of 'and' to Transylvanian Saxon *auch* and not to an addition of 'also, as well, too' to Transylvanian *und*; the theory of linguistic recessivity allows us to assume that the probability for the latter case was quite low, whereas older approaches fail to explain this. Moreover, it seems to be implausible that the signifier in question (∗*auke*, *ouch*, *auch* etc.) got connected with the meaning of 'and' two times in two thousand years without a connection between those events. This is why it seems that the theory of linguistic recessivity best explains semantic "change" like this.[30] Thus, we can record the fact that so-called language change is obviously not a change in linguistic objects, but rather in probability distributions determining the appearance of linguistic objects when getting phenotypificated; finally, a change in probability distribution within linguistic systems can be caused by changes in determinants that affect the system from the outside, i.e. the environment of the system (e.g. cultural evolution or historical events).[31]

**29** Of course, we cannot be absolutely sure here: It is possible that a single phenotypification of *auch* (or *ouch* etc.) 'and' took place in medieval German (or even more than one phenotypification) without us knowing this today (see for more details Decker, n.d.: Section 3.1).

**30** Similar approaches based on this theory can be made regarding some phenomena of so-called "syntactic loan" like e.g. the *dativus absolutus* in Old English or Old High German, which is said to be a syntactic loan due to influence of the Latin *ablativus absolutus* (cf., regarding those constructions, e.g. Mitchell and Robinson 2007: 106; Schrodt 2004: 95). Nevertheless, it is very likely that the *dativus absolutus* was already common in Proto-Germanic (cf. Euler and Badenheuer 2009: 180–181); if it really got lost before appearing in Old English or Old High German – as often assumed – recessivity would be a possible explanation for its "return".

**31** The fact that there were no computers in the Middle Ages is a good example for such changes in determinants of linguistic systems: It was not completely impossible to phenotypificate *computer* with the meaning 'computer' in the eleventh century (for there are no laws of nature that would have made humans unable to phenotypificate this), but the probability for this was nearly 0 – but not exactly 0 – for there were no such things as computers at that time; the invention of the computer in the twentieth century increased this probability significantly.

# 5 Conclusion: Language change with regard to the theory of linguistic recessivity – a construct?

This paper examines the basis of linguistic systems on individual and collective levels, and also its consequences for the concept of language change. Using evidence from language processing and diachronic analysis, we recognized probabilistic approaches to language – as they can be found in Bod et al. 2003 or even Lass 1980 – as fitting best to explain common linguistic issues like language processing or language as an over-individual system for exchange of information. A probabilistic extension of the Saussurean model of linguistic signs following the great importance of individuals and their perception allows us to appreciate all possible varieties of linguistic objects and to not have to reduce them to only the most common one. Thus, probability distributions and a distinction between a linguistic geno- and phenotype and also recessive information as part of linguistic systems are the terminological foundations of the theory of linguistic recessivity presented in this paper. With its terminology we could examine even special cases of so-called language change like the case of Transylvanian Saxon *auch* showing that the new approach best explains some of the difficulties regarding this case.

With all these insights in mind, the question that was raised in the title and introduction of this paper – whether language change is nothing more than a scientific construct or not – might be answered as follows:

First, we could say: No, there is no such thing as language change in its "classical sense"; language is a system of elements held together by relations between them that are defined by flexible (i.e. unstable) probability distributions, which might be activated differently each time a linguistic object gets phenotypificated. This means that there are no stable linguistic objects – not even for a short time in a certain speech community – and therefore changes are not noticeable (or we have to argue that there are always changes, which would make the concept of language change questionable once more). So, language change in its classical sense is nothing more than a scientific construct referring to states of linguistic objects that do not exist in the way they are said to be. Thus, language change is nothing more than an approach to a reality that is too complex to be explained through the construct of language change.

But secondly one could argue: Yes, language change is not only a scientific construct; it really "happens". Examining examples from language history like the case of Transylvanian Saxon *auch* or the Great Vowel Shift in the history of English, we recognized that there is no complete stability in linguistic systems, but rather stabilizing factors such as macrosynchronizations (in terms of Herrgen and Schmidt 2011) and imaging. Such processes protect linguistic objects from being

changed every time they are phenotypificated, and collective levels thus allow us to communicate based on common "rules". Regarding probabilistic approaches to linguistic systems this means that changes in probability distributions of the possibilities of phenotypification of a linguistic object normally are not significant. So, we could call significant changes in such probability distributions over time language change.

But finally, we have to admit that such changes in probability distributions seem to occur all the time and that we have no numeric criteria that allow us to judge whether a change in probability distribution is significant or not; such a judgement would always be arbitrary to some degree. This means, of course, that we are faced with the same situation here as we were regarding language change in its classical sense: It actually occurs all the time. Therefore, language change is a scientific construct that may be helpful to examine certain linguistic phenomena observed on the "surface", if we keep in mind that we are using a simplification; but it lacks precision regarding the details.

## References

Aitchison, Jean. 2012. *Words in the mind. An introduction to the mental lexicon*. Chichester: Wiley-Blackwell.

Anderson, Stephen R. & David W. Lightfoot. 2000. The human language faculty as an organ. *Annual Review of Physiology* 62. 697–722.

Andor, Daniel, Chris Alberti, Michael Collins, Kuzman Ganchev, Slav Petrov, Alessandor Presta, Aliaksei Severyn & David Weiss. 2016. *Globally normalized transition-based neural networks*. https://arxiv.org/pdf/1603.06042 (accessed 06 June 2018).

Benecke, Georg Friedrich, Wilhelm Müller & Friedrich Zarncke, (eds.). 1990. *Mittelhochdeutsches wörterbuch*. M–R.Stuttgart: Hirzel Bd. 2, Abt. 1.

Bod, Rens, Jennifer Hay & Stefanie Jannedy. 2003. Introduction. In Rens Bod, Jennifer Hay & Stefanie Jannedy, *Probabilistic linguistics*, 1–10. Cambridge (Massachusetts) & London: MIT Press.

Cohen, Ariel. 2009. Probability in semantics. *Language and Linguistics Compass* 3/1. 265–281.

Decker, Eike U. Forthcoming. "Rezessive" information in sprache. Ein probabilistisches informationsmodell als grundlage informationslinguistischer und systemlinguistischer Ansätze / "Recessive" information in language. A probabilistic information model as the basis of informational linguistics and systems linguistics. Heidelberg: University of Heidelberg dissertation.

Euler, Wolfram & Konrad Badenheuer. 2009. *Sprache und herkunft der Germanen. Abriss des protogermanischen vor der ersten lautverschiebung*. Hamburg & London: Inspiration Un Limited.

Fagen, Robert E. & Arthur David Hall. 1956. Definition of system. *General Systems* 1. 18–28.

Gerjuoy, Edward. 1993. Superposition principle. In Sybil P. Parker (ed.), *McGraw-Hill Encyclopedia of Physics*. 2nd edn. 1384. New York et al.: McGraw-Hill.

Hebb, Donald Olding. 1949. *The organization of behavior*. New York: Wiley & Sons.

Herrgen, Joachim & Jürgen Erich Schmidt. 2011. *Sprachdynamik. Eine einführung in die moderne regionalsprachenforschung*. Berlin: Schmidt.
Jakobson, Roman Osipovich. 1960. Closing statement: Linguistic and poetics. In Thomas Albert Sebeok (ed.), *Style in language*, 350–377. Cambridge (Massachusetts): MIT Press.
Jäncke, Lutz. 2013. *Lehrbuch kognitive neurowissenschaften*. Bern: Hans Huber.
Köbler, Gerhard. 1989. *Gotisches wörterbuch*. Köln & Leiden: Brill.
Kroonen, Guus. 2013. *Etymological dictionary of proto-Germanic*. Boston & Leiden: Brill.
Krug, Manfred. 2017. Chapter 14, The great vowel shift. In Alexander Bergs and Laurel J. Brinton (eds.), *The history of English. Volume 4: Early modern English*, 241–266. Berlin & Boston: de Gruyter Mouton.
Labov, William. 1994. *Principles of Linguistic change. Volume 1: Internal factors*. Oxford & Cambridge (Massachusetts): Blackwell.
Lass, Roger. 1980. *On explaining language change*. Cambridge: Cambridge University Press.
Leopold, Edda. 2005. Das piotrowski-gesetz. In Gabriel Altmann, Reinhard Köhler & Rajmund G. Piotrowski (eds.), *Quantitative Linguistik/Quantitative Linguistics: Ein internationales Handbuch/An International Handbook*, 627–633. Berlin & New York: Walter de Gruyter.
Luhmann, Niklas. 1987. *Soziale Systeme. Grundriß einer allgemeinen theorie*. Frankfurt am Main: Suhrkamp.
McNamara, Timothy P. 2005. *Semantic priming. Perspectives from memory and word recognition*. Hove & New York: Psychology Press.
Mitchell, Bruce & Fred C. Robinson. 2007. *A guide to old English*. Oxford: Blackwell.
Montague, Richard. 1970. Universal grammar. *Theoria* 36(3). 373–398.
Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Jaakko Hintikka, Julius Moravcsik & Patrick Suppes (eds.), *Approaches to natural language. Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, 221–242. Dordrecht: Reidel.
Nübling, Damaris. 2000. *Prinzipien der irregularisierung. Eine kontrastive analyse von zehn verben in zehn germanischen sprachen*. Tübingen: Niemeyer.
Quillian, M. Ross. 1967. Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science* 12. 410–430.
Quillian, M. Ross. 1968. Semantic memory. In Marvin Minsky (ed.), *Semantic information processing*, 227–270. Cambridge (Massachusetts): MIT Press.
Quillian, M. Ross. 1969. The teachable language comprehender. *Communications of the ACM* 12. 459–476.
Rey, Günter Daniel & Karl F. Wender. 2011. *Neuronale netze: Eine einführung in die grundlagen, anwendungen und datenauswertung*. Bern: Hans Huber.
Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Lausanne & Paris: Librairie Payot & Cie.
Schlüter, Julia. 2017. Chapter 3, Phonology. In Alexander Bergs and Laurel J. Brinton (eds.): *The history of English. Volume 4, Early modern English*, 27–46. Berlin & Boston: de Gruyter Mouton.
Schmidt, Jürgen Erich. 2010. Language and space: The linguistic dynamics approach. In Jürgen Erich Schmidt (ed.), *Language and space: An international handbook of Linguistic variation*, 201–225. Berlin & New York: de Gruyter Mouton.
Schrodt, Richard. 2004. *Althochdeutsche Grammatik II: Syntax*. Tübingen: Niemeyer.
Schweizer, Harro. 1979. *Sprache und Systemtheorie. Zur modelltheoretischen Anwendung der kybernetischen Systemtheorie in der Linguistik*. Tübingen:Narr.

Shinohara, Sara. 2016. Der konjunktionale Gebrauch von *auch* im Siebenbürgisch-Sächsischen. In Japanische Gesellschaft für Germanistik (ed.), *Beiträge zur dialektologie*, 58–75. München: Iudicium.

Simmons, Robert. 1972. Some semantic structures for representing English meanings. In John B. Carroll & Roy O. Freedle (eds.), *Language comprehension and the acquisition of knowledge*, 71–97. Washington (D.C.): Winston.

Simmons, Robert. 1973. Semantic networks: Their computation and use for understanding English sentences. In Roger C. Schank & Kenneth Mark Colby (eds.), *Computer models of thought and language*, 63–113. San Francisco: Freeman.

Wirtz, Markus Antonius (ed.). 2017. *Dorsch – Lexikon der Psychologie*. Bern: Hogrefe.

Zoëga, Geir T. 2004. *A concise dictionary of old Icelandic*. Toronto: University of Toronto Press.