Voronoi Languages[†]

By Gerhard Jäger and Lars Koch and Frank Riedel[‡]

Bielefeld University

— this Version: February 20, 2009

Abstract

[†] The research that led to this paper has been supported by project A6 of the SFB 673 "Alignment in Communication" of the German Research Foundation, which is gratefully acknowledged.

[‡] We would like to thank

JEL subject classification.

Key words and phrases.

1 Introduction

In this paper we are going to consider the evolutionary stability conditions of a certain class of cheap talk sender–receiver games. In sender–receiver games, we have two players, the sender and the receiver. The sender has some private information—his type—that is supplied to him by nature. In the first stage of the game, the sender emits a signal that is observable by the receiver. In the second stage, the receiver performs an action. The choice of his action may depend on the sender's signal. The utility of both players depends both on the sender's type and the receiver's action.

It is tempting to assume that in a cheap talk setting, i.e. in a scenario where emitting a signal is costless for the sender, there is an unbounded supply of possible signals. However, there are many real-life situations where this is not the case. To take a simple example, the way a baseball cap is worn serves as a social signal among certain subcultures. There are only finitely many easily distinguishable ways how to wear a baseball cap: the brim facing forward, backward, to the left or to the right. These signals may convey certain information about the type of the wearer of the cap, like age, group membership, musical taste etc. So the space of possible types is virtually unbounded, while the space of signals is small.

To keep things simple, we assume that the purpose of communication is type matching, i.e. the goal of the receiver is to guess the sender's type. We also assume that the interests of sender and receiver are identical. So payoff would be maximal if the receiver would always correctly guess the sender's type.

If the number of types exceeds the number of signals, however, there is no equilibrium where communication is always perfect. Not every imperfect communication is equally severe though. We assume that the space of types is structured by a similarity metric, and the utility is the higher the more similar the receiver's guess is to the true type of the sender. In line with much work in cognitive science, we take it that types are points in a (possibly high-dimensional) Euclidean space. The similarity between two types is inversely related to their distance. Briefly put, the goal of the receiver is to make a guess that is as close as possible to the true type.

A pure sender strategy maps each type to a certain signal. This induces a partition of the type space into disjoint sets, each of which is the inverse images of one signal. This can be interpreted as a **categorization** of the type space. So the evolution of signaling strategies induces an evolution of categories as a side effect.

This co-evolution of signal meanings and categories can frequently be observed in quite different domains. For instance, conspicuous support for a certain political party is, among other things, a social signal that may carry information about political convictions, but also about class membership, level of income and education, religious and regional affiliation etc. The establishment of a new party, or the merging of two parties into a single one, may push a system out of equilibrium and induce an evolutionary dynamics that eventually leads to a new equilibrium with new fault lines dividing the electorate.

Likewise, the grammar of a natural language induces a certain categorization, which

depends on the inventory of grammatical signals.[1] For instance, at some stage the proto-indoeuropean language (which is the ancestor language of most European languages and many western and central Asian languages) had three categories for number: singular, dual and plural. Singular was used for single entities, dual for pairs of objects, and plural for pluralities of at least three objects. In most modern indoeuropean language—as for instance in Modern English—the grammatical category of dual has been lost. As a consequence, plural changed its meaning, and it now also encompasses pairs of objects. Similar correlated shifts in grammatical paradigms and the corresponding semantic categorizations can frequently be observed in language change, for instance regarding grammatical tense or mood.

Empirical research in linguistics has shown that the categorizations that natural languages impose on semantic domains is not arbitrary. Only few partitions of a given domain provide possible meanings of simple natural language expressions. A very specific proposal about what constitutes a possible meaning is due to Gärdenfors (2000). He assumes that semantic domains have a geometrical structure on which a ternary relation of "betweenness" can be defined. (In a Euclidean space, point x is between y and z iff x is on the straight line connecting y and z. Gärdenfors' proposal also applies to non-Euclidean spaces though.) Gärdenfors claims that the extensions of simple natural language expressions (especially of adjectives) must be closed under the betweenness–relation. If an adjective A can be applied to x and y, than it must also be applicable to any object which lies between x and y. The denotation of nouns or verbs are more complex concepts because they usually involve more than one semantic domain. Nevertheless, they are usually closed under the betweenness–relation as well.

There is plenty of empirical support for Gärdenfors' proposal. It evidently holds for spatial expressions like *above the table*, *behind the curtain*, *between Boston and New York* etc. Similarly, it holds if the semantic domain in question has easily identifiable dimensions, like temperature. Even though the boundary between *hot* and *warm* is vague and context-dependent, if you consider 40°C and 60°C to be hot, you will also consider 50°C to be hot.

A less trivial example is derived from a classical study by Labov (1973). He noticed (and confirmed experimentally) that the extension of the words *cup*, *bowl* and *vase* are blurred, and that it is not possible to give a precise definition when an object is a cup or a bowl, say. Rather, categorization is vague and context-dependent. This is illustrated in figure 1 (taken from Löbner (2003), p. 262). While objects 3, 6 and 10 are clear instances of cups, bowls and vases respectively, categorization is not so straightforward for 7 or 9, say. However, if you categorize objects 2 and 4 as cups, you also have to categorize object 3 as a cup, because the shape of object 3 is intuitively between the shapes of objects 2 and 4.

Sets that are closed under the betweenness–relation are called **convex**. So Gärdenfors' proposal amounts to the claim that the denotation of simple natural lan-

[1] An extreme form of this point of view is the so-called "Sapir-Whorf hypothesis" that states that the grammar of our native language shapes the way we perceive the world. We remain neutral on the issue whether and to what degree this claim is correct.

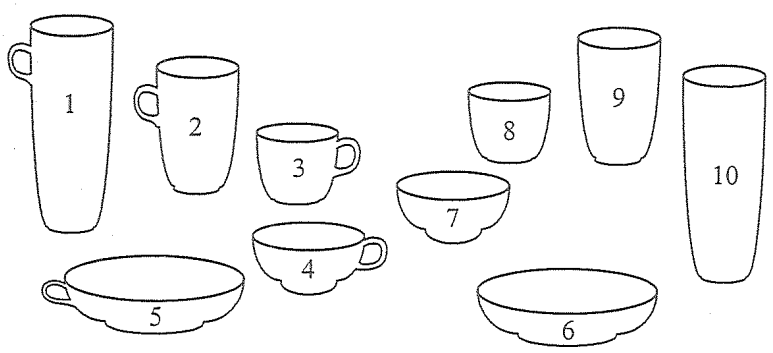


Figure 1: Cups, bowls and vases

guage expressions are convex sets in an appropriate conceptual space.

It seems initially plausible that convexity also holds for categories that are induced by non-linguistic social signals, like the choice of product brands. For instance, if full professors of English literature prefer caffe latte over regular coffee, and assistant professors of English literature prefer caffe latte over regular coffee, the odds are that associate professors of English literature have the very same preference.

In this paper, we will propose an explanation why signals tend to have convex meanings. Briefly put, we will show that in the described setting, evolution will necessarily lead to a partition of the semantic space into convex regions.

In previous work on the evolution of cheap talk sender–receiver games (like Blume, Kim, and Sobel (1993), Wärneryd (1993), or Trapa and Nowak (2000)), it is assumed that there are at least as many signals as types or that the utility is either 1 (success) or 0 (failure). A setting similar to ours was proposed in Jäger and van Rooij (2007) and worked out in some detail in Jäger (2007). In the latter paper it is shown that the replicator dynamics always leads to convex categories if the type space is finite. We generalize the model to continuous type spaces. It turns out that Jäger's results hold there as well. As argued in Oechssler and Riedel (2001), in games with infinite strategy spaces a stronger notion of evolutionary stability than the classical ESS is needed. We will show that not every ESS in the class of games considered here is evolutionarily robust in this sense.

2 Model and Notation

The sender has a type $t \in T$, where T is a convex and compact subset of $\mathbb{R}^L$ for some $L \geq 1$ that has nonempty interior. He chooses a word (signal) $w \in W := \{w_1, \ldots, w_N\}$ from a finite language and sends it to the receiver. The receiver interprets w as some point $i \in T$. Both players want type t and interpretation i to be as similar as possible. We assume that $l(\|t - i\|)$ measures the loss of the players where the function $l : \mathbb{R}_+ \to \mathbb{R}$ is convex and strictly increasing. A natural choice that we consider frequently below is the square Euclidean distance $\|i - t\|^2$. The probability of types is described by a

atomless distribution F on T with strictly positive density $f : T \to \mathbb{R}_+$.

A (pure) strategy for the sender is a measurable function $w : T \to W$. We denote by Σ the set of all sender strategies. A (pure) strategy for the receiver is a vector $i = (i_1, \ldots, i_N) \in T^N$ where i_j denotes the interpretation of the word w_j. The expected loss of players is then

$$L(w, i) = \int_T l\left(\|t - i_{w(t)}\|\right) F(dt)\,.$$

Note that null sets play no role for the expected loss. Hence, we ignore them in the sequel when we characterize strategies.

3 Optimal Languages

To start with, we study what the two players can achieve in cooperation. Ideally, we might think of super–rational players who have a meta–language to communicate with each other; before playing, they meet in an ideal place to discuss their optimal strategy. Formally, we call a language (w, i) optimal if it minimizes the loss $L(w, i) = \int_T l\left(\|t - i_{w(t)}\|\right) F(dt)$.

Before coming to the proofs, let us give a short synopsis of the results. If there were as many words as types, the players would clearly choose a language that distinguishes perfectly all private information (a fully separating equilibrium in the language of game theory). In our situation, this is not feasible as the type space is a continuum and the set of words is finite. Nevertheless, players will use equilibria that are "as separating as possible", i.e. they will use all available words and attach different meanings to them. The sender will thus choose a partition $(W_k)_{k=1,\ldots,n}$ of the space T and say word w_k whenever his type t is in the cell W_k. Given such a partition, the receiver has to choose a "prototype" $i_k \in T$ for each word w_k that describes the average type in W_k optimally (given the environment (or prior) F. An optimal interpretation consists thus of optimal Bayesian estimators for each cell W_k. At first glance, it might seem that we cannot say much more. This is not true, however. Given that the receiver uses prototypes $i_k \in T$, we can ask in turn what the optimal partition is. The sender wants the prototype to be as close to his type as possible. Hence, he will say w_k whenever the prototype i_k is closest to his type t among all prototypes. Such a partition of the type space is called a *Voronoi tessellation* of the space. Summing up, optimal languages consist of what we call *Voronoi languages with full vocabulary*—a Voronoi tessellation of the space T that is induced by points i_k which are at the same time optimal Bayesian estimators for the average type in each cell. Depending on the reader's intuition, you might expect to find a plethora or very few of such optimal languages. We illustrate by examples that there are usually very few Voronoi languages (up to the obvious symmetries, of course).

Let us come to the formal analysis.

Definition 1 *A language (w, i) consists of a measurable mapping $w : T \to W$ (the signaling strategy) and points $i \in T^N$ (the interpretation). A language (w, i) has full vocabulary if* $\operatorname{range} w = W$.

We show now that our problem is well–posed, i.e. that optimal languages (w, i) which minimize $L(w, i)$ exist. A slight problem comes from the fact that the payoff $l\left(\|t - i_{w(t)}\|\right)$ is not continuous in t, in general. We proceed as follows. We first show that one can restrict attention to strategies w that are induced by Voronoi tessellations. As Voronoi tessellations can be described by their center points $i_k \in T$, we can study now an auxiliary payoff function which only depends on N points in T. As T is compact, it is enough to show continuity of this function. This is done by noting that $l\left(\|t - i_{w(t)}\|\right)$ jumps only at the boundaries of Voronoi cells which form a Lebesgue null set. Hence, the auxiliary payoff function is continuous by Lebesgue's theorem. As a consequence,

Lemma 1 *Optimal languages exist.*

The proof of this and all other results can be found in the appendix.

We turn now to an analysis of optimal languages. Let us begin with a detour. So far, we have not even discussed the possibility of mixing, or randomized strategies (and we will not do so later on). For one paragraph, we will allow for this possibility — just to show that mixing is not optimal. This is quite plausible: the players have no reason to introduce randomness in their communication when they cooperate[2]. A mixed strategy for the sender is a measurable mapping $\omega : T \to \Delta W$ where ΔW denotes the set of probability vectors over W. We denote by $\omega_k(t)$ the probability that the sender chooses word w_k if in type t. A mixed strategy for the receiver consists of probability measures $(\mu_k)_{k=1,\ldots,N}$ over T[3]. The generalized payoff function for such strategies is then

$$L(\omega, \mu) = \int_T \sum_{k=1}^{N} \int_T l\left(\|t - i\|\right) \mu_k(di)\omega_k(t)F(dt)\,.$$

Lemma 2 *Allowing for randomized strategies does not lead to better languages. For every language (ω, μ) in randomized strategies, there is a pure strategy language (w, i) which is at least as good.*

From now on, we thus return to pure strategies (w, i).

Our next rather obvious point is that players should use all available words given that there is no cost in using them. The proof uses the fact that F is atomless. When a language does not use one word w_n, say, one can split a used word, w_1, say, in two words, and obtain a better language. It is also clear that the receiver should interpret different words differently (as they represent different areas of the type space T).

Lemma 3 *Optimal languages (w^*, i^*) have full vocabulary and interpretations i_k^* are pairwise distinct.*

[2]There are, of course, mixed, or partially mixed Nash equilibria, and mixing can be a best reply. The convexity of the loss function induces risk aversion for the players. Their payoff is thus not increased by mixing.

[3]Whenever we speak of measurability, probability etc. we think of T as endowed with the Borel σ–field.

We can thus focus on languages in which all interpretations are pairwise distinct. Given that the receiver uses the pairwise distinct points (i_k), what words should the sender choose if in type t? Clearly the word that leads to the interpretation i_k which is closest to t among all interpretations.

Lemma 4 *In optimal languages (w^*, i^*), the sender uses a Voronoi tessellation corresponding to i^*, i.e. F–almost everywhere*

$$w^*(t) = \text{argmin}_{j=1,\ldots,N} \|t - i_j^*\| \,. \tag{1}$$

Note that the above strategy is not uniquely defined at points t that have equal distance to two or more interpretations. As these points form a null set, we can ignore this ambiguity; without loss of generality, we always take the word with smallest index in this case.

It is quite easy to see (cf. for instance Okabe, Boots, and Sugihara (1992) for a proof) that in Euclidean spaces, the interior of each cell of a Voronoi tessellation is a convex set. (To see why, please observe that for each pair of prototypes x and y, the set of points that is closer to x than to y forms an L-dimensional half-space that is bounded by the hyperplane of points that are equidistant to x and y. A half-space is evidently a convex set. A Voronoi cell is an intersection of finitely many half-spaces, and the intersection of convex sets must be convex again.) So we have the

Corollary 1 *In optimal languages (w^*, i^*), the sender uses convex categories, i.e. for each i_j^*, $w^{*-1}(i_j^*)$ is (up to a null set) a convex set, the intersection of a convex polyhedron with the type space T.*

Let us now come to the receiver. Given that the sender uses a Voronoi tessellation (where each cell has positive measure), she has to determine an optimal interpretation. By Bayes' rule, she has to choose an optimal estimator given that she knows the type to be in that cell.

Definition 2 *Let $C \subset T$ be a convex set with positive measure. Call*

$$b(C) = \text{argmin}_{i \in C} \int_C l\left(\|t - i\|\right) F(dt)$$

the optimal Bayesian estimator conditional on C.

Remark 1 *Note that the optimal Bayesian estimator is uniquely determined. This follows from Jensen's inequality. Note that we get the strict inequality because the integrand $l\left(\|t - i\|\right)$ is convex, increasing, and not linear.*

Let us state the two best known estimators for the quadratic and linear loss function.

Example 1 *1. For $l(d) = d^2$, the best estimate is the conditional expectation,*

$$b(C) = \mathbb{E}[t | t \in C] := \frac{1}{F(C)} \int_C t \, F(dt) \,.$$

2. *For $l(d) = d$, the best estimator is the (generalized) conditional median type given the cell C.*

Lemma 5 *In optimal languages (w^*, i^*), the receiver uses the best interpretation of the partition induced by w^*, i.e.*

$$i_k = b(W_k^*)$$

for

$$W_k^* = \{t \in T : w^*(t) = w_k\}.$$

We summarize our findings in a definition.

Definition 3 (Voronoi Language) *A Voronoi language (w, i) consists of a Voronoi tessellation for the sender and an optimal Bayesian estimator interpretation for the receiver. i.e. we have both*

$$w^*(t) = \text{argmin}_{j=1,\ldots,N} \|t - i_j^*\| \tag{2}$$

$$i_k = b(W_k^*) \qquad (\text{ for } W_k^* = \{t \in T : w^*(t) = w_k\}). \tag{3}$$

The new concept allows us to characterize optimal language succinctly.

Theorem 1 *Optimal languages are Voronoi languages with full vocabulary.*

To get a better intuition, we start with two (highly idealized and simple) examples where there are only two words and types are uniformly distributed. On the unit interval $[0, 1]$, there is only one Voronoi language with full vocabulary (which is also the unique optimal language). On the unit square, there are two Voronoi languages (up to symmetries). Interestingly, only one of them is optimal. The converse of the above theorem is thus not valid.

Example 2 *Consider the unit interval $T = [0, 1]$ with the uniform distribution $F(x) = x$, quadratic loss $l(d) = d^2$, and two words $W = \{w_1, w_2\}$. The two words have the obvious everyday meaning of "left" and "right". The optimal Voronoi language has $w^*(t) = w_1$ for $t \leq 1/2$ and $w(t) = w_2$ else. The best interpretation is $i_1^* = 1/4, i_2^* = 3/4$. Let us quickly show that this is the only Voronoi language[4] with full vocabulary here. If we denote by K the threshold that separates the two Voronoi cells, we must have*

$$i_1 = K/2$$
$$i_2 = (1 + K)/2$$

as optimal Bayesian estimators and

$$K = (i_1 + i_2)/2$$

because K must correspond to the Voronoi tessellation induced by i_1 and i_2. Substituting i_1, i_2 in the third equation, we get $K = 1/2(1/2 + K)$, or $K = 1/2$, and then $i_1 = 1/4$ and $i_2 = 3/4$ as desired.

[4] up to the obvious symmetry, of course

In the previous example, optimal and Voronoi languages coincide. This need not be the case, as we now illustrate.

Example 3 (Not all Voronoi languages are optimal) *Consider the unit square* $[0,1]^2$*, with the uniform distribution, quadratic loss* $l(d) = d^2$*, and two words* $W = \{w_1, w_2\}$*. A typical Voronoi tessellation consists here of two points that lead to two trapezoids as illustrated in Figure HIER BILD TRAPEZOID. The two border cases are the* vertical *tessellation of Figure HIER BILD VERTIKAL and the diagonal tessellation of Figure HIER BILD DIAGONAL. One might guess that trapezoid tessellations that are symmetric around the center point* $(0.5, 0.5)$ *are Voronoi languages. This is not true, however because the center of gravity (the Bayesian estimator) of a trapezoid does not coincide with the point that generates the cell (see Figure HIER BILD TRAPEZ IST NICHT VORONOI). A nice exercise in geometry shows that the diagonal and the vertical language are the two unique Voronoi languages. The diagonal is not optimal, however, because it leads to a loss*

$$2 \cdot \int_0^1 \int_0^{1-y} (x - 1/3)^2 + (y - 1/3)^2 dxdy = 1/9 \simeq 0.111,$$

whereas the vertical language has loss

$$2 \cdot \int_0^1 \int_0^{1/2} (x - 1/4)^2 + (y - 1/2)^2 dxdy = 5/48 \simeq 0.104\,.$$

4 Pure Strategy Nash Equilibria

Although cooperative solutions can be achieved in ideal situations where the players have the possibility to use a meta–language for before–play communication, the everyday situation is different. Here, we rather have to guess what our partner might mean with his words—not an easy task. This situation is better modeled as a noncooperative signaling game between the two players. Let us assume rationality of the players for the moment (we turn to the more realistic case of bounded rationality later on).

As in all signaling games, there is a plethora of Nash equilibria. We focus here on strict equilibria[5] where the best replies of both players are unique. In this case, we can make use of our optimality analysis of the preceding section.

First, let us note that strict Nash equilibria share with optimal languages the feature that all words are being used. The argument is different from the one we used for optimal languages, though. For optimality, we use the fact that the average loss can be reduced by using more words. Such a cooperative argument does not work in the game–theoretic setting. Instead, we rely on the uniqueness of best replies to exclude such a phenomenon. If the sender never uses a word, say w_n, then the receiver is indifferent between all interpretations for w_n, and the best reply is not unique.

[5]Strictly speaking, we require only that the sender's best reply is almost surely unique (changes on a null subset of T do not influence the payoff).

As for optimal languages, the sender's best reply to a given interpretation is the corresponding Voronoi tessellation. Similarly, the receiver's best reply to a partition consists of the best Bayesian estimator—here, the arguments are identical to those in Lemmata 4 and 5.

Conversely, every Voronoi language with full vocabulary consists of a pair of mutually best replies, and is thus a Nash equilibrium. The sender's Voronoi tessellation is the (almost sure) unique best reply (there is indifference at the points that are equidistant to two or more interpretations, a null set). The receiver's best reply is unique because of the strict convexity of the loss function $i \mapsto l(\|t - i\|)$, compare Remark 1.

Theorem 2 *Every Voronoi language with full vocabulary is a strict Nash equilibrium and vice versa.*

We thus have a full characterization of strict Nash equilibria. In particular, we see that suboptimal languages can arise even if we impose the relatively strong condition of strictness on the set of Nash equilibria, compare Example 3 above. Rational communication does not necessarily result in optimal signaling systems.

5 Evolution of Voronoi Languages

Our current language is not a fixed system, rather a fluent and flexible body of words and rules that is constantly evolving. As such, it is shaped by the typical forces of selection and mutation that govern evolution. We are thus led to study dynamical systems that describe possible evolutionary dynamics.

On the technical side, we face here a rather complicated dynamical system because a population is described by a probability measure over all strategies—and strategies are pairs of signaling systems, i.e. simple measurable functions on T with values in W and interpretations, points in T^N. For several dynamics, the technical foundations for the study of the replicator (Oechssler and Riedel (2001), Oechssler and Riedel (2002), Cressman, Hofbauer, and Riedel (2006)), payoff–monotone (Heifetz, Shannon, and Spiegel (2007)), and Brown–von–Neumann–Nash dynamics (Hofbauer, Oechssler, and Riedel (2007)) have been worked out. Although our strategy space is slightly more general than in some of the cited papers, the general results of these papers hold true in our setting.

For our dynamical considerations, we consider the symmetrized version of the game. Let us suppose that the players are equally often in the role of receiver and sender, and every player thus chooses both a sender strategy v or $w \in \Sigma$ as well as a receiver strategy i resp. $j \in T^N$. Then the ex ante payoff is

$$\Lambda((v,i),(w,j)) = 1/2(L(v,j) + L(w,i)) .$$

A population of players is described by a probability distribution $P(dw, di)$ over the strategy set $\Gamma := \Sigma \times T^N$ of the symmetrized game. For two such distributions P and

Q, we can extend the symmetrized payoff function in the usual way by setting

$$\Lambda(P,Q) = \int_\Gamma \int_\Gamma \Lambda((v,i),(w,j)P(dv,di)Q(dw,dj)\,.$$

The dynamic analysis is greatly simplified by the fact that average loss is decreasing along the paths of typical selection and innovative dynamics, as usual in common interest games.

Lemma 6 (Fundamental Law of Natural Selection) *The symmetrized payoff function is a Lyapunov function for the replicator (more generally, regular, payoff–monotone) and the Brown–von Neumann–Nash dynamics.*

Technically, it is important to show that the loss function is continuous with respect to the weak topology for probability measures because we can only expect convergence in the weak topology, in general[6].

Lemma 7 *The payoff function is continuous with respect to the weak topology.*

We can now focus on the dynamic analysis. For games with finite strategy spaces or infinite strategy spaces that are equipped with the strong topology, the standard static stability notion is the concept of an evolutionarily stable state (ESS). In the case of asymmetric games (or symmetrizations thereof) as discussed here, this is equivalent to the notion of a strict Nash equilibrium. In our cases, these are precisely the Voronoi languages. The generalization of evolutionarily stable state (ESS) or uniformly superior state to the weak topology has been termed *evolutionarily robust state* (ER) in Oechssler and Riedel (2002). P^* is an ER if for some invasion barrier $\epsilon > 0$ if we have $\Lambda(P^*,Q) < \Lambda(Q,Q)$ for all $Q \neq P^*$ that have distance smaller than ϵ in the Prohorov topology.

It is also shown in Oechssler and Riedel (2002) that every ER is an ESS.

Observation 1 *Every ER is a Voronoi language.*

Lemma 8 *Evolutionarily robust languages are strict local optima.*

Theorem 3 *Locally optimal languages are stable with respect to replicator (more generally, payoff–monotone) and Brown–von Neumann–Nash dynamics.*

Example 4 (Optimal Languages need not be asymptotically stable with respect to replicator dynamics) *Sometimes, ER do not exist. For instance, reconsider example 2 (where we have two words and the unit interval as the type space T, plus a uniform probability F distribution over T and a quadradic loss function). The optimal language here is (w^*,i^*) (disregarding null sets and up to permutation of words) where $w^*(t) = w_1$ if $t \in [0,1/2]$ and $w(t) = w_2$ else, and where $i_1^* = 1/4$ and $i_2^* = 3/4$. Now*

[6]See Oechssler and Riedel (2002) and Hofbauer, Oechssler, and Riedel (2007) for an extended discussion of this point.

consider a mutant language (w', i') *with* $w'^{-1}(w_1) = [0, 1/2 + \epsilon]$, $w'^{-1}(w_1) = (1/2 + \epsilon, 1]$, $i'(w_1) = 1/4 + \epsilon$, *and* $i'(w_2) = 3/4 + \epsilon$ *(for some* $\epsilon \in (0, 1/4)$*). We have*

$$\begin{aligned}
L(w,i) &= \int_0^{1/2} (1/4 - x)^2 dx + \int_{1/2}^1 (3/4 - x)^2 dx \\
&= 1/48 \\
L(w,i') &= \int_0^{1/2} (1/4 + \epsilon - x)^2 dx + \int_{1/2}^1 (3/4 + \epsilon - x)^2 dx \\
&= 1/48 + \epsilon^2 \\
L(w',i) &= \int_0^{1/2+\epsilon} (1/4 - x)^2 dx + \int_{1/2+\epsilon}^1 (3/4 - x)^2 dx \\
&= 1/48 + \epsilon^2/2 \\
L(w',i') &= \int_0^{1/2+\epsilon} (1/4 + \epsilon - x)^2 dx + \int_{1/2+\epsilon}^1 (3/4 + \epsilon - x)^2 dx \\
&= 1/48 + \epsilon^2/2 \\
\Lambda((w,i),(w',i')) &= 1/2(L(w,i') + L(w',i)) \\
&= 1/48 + 3\epsilon^2/4 \\
\Lambda((w',i'),(w',i')) &= L(w',i') \\
&= 1/48 + \epsilon^2/2 < \Lambda((w,i),(w',i'))
\end{aligned}$$

As (w, i) *is optimal, a homogenous population of* (w, i)*-players cannot be invaded by a small fraction of mutants of any sort. However, in the weak topology a homogenous population* (w', i') *is also within the* ϵ*-environment of* (w, i)*. The calculation above shows that a homogenous population of* (w', i')*-players cannot be invaded by a small fraction of* (w, i)*-players either. Considering only these two pure strategies, we are dealing with a* 2×2 *game with the utility matrix*

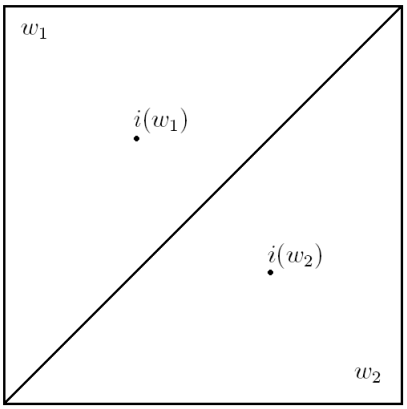
	(w,i)	(w',i')
(w,i)	$-1/48$	$-1/48 - 3\epsilon^2/4$
(w',i')	$-1/48 - 3\epsilon^2/4$	$-1/48 - \epsilon^2/2$

.

In this reduced game both (w, i) *and* (w', i') *are strict equilibria and thus evolutionarily stable. According to a result of Eshel and Sansone (2003), we see that the optimal language is not asymptotically stable, although it is stable.*

Example 5 (An inefficient (subotpimal) Voronoi language is eliminated by evolution) *Consider a language with two signals and types uniformly distibuted on the unitsquare* $[0, 1] \times [0, 1]$.

A strict Nash equilibrium is the "diagonal language" in which the reader sends w_1 *if* $t_2 \geq t_1$ *and* w_2 *if* $t_2 < t_1$. *The receiver interprets* w_1 *as* $i(w_1) = (\frac{1}{3}, \frac{2}{3})$ *and* w_2 *as* $i(w_2) = (\frac{2}{3}, \frac{1}{3})$. *These strategies are mutually unique best replies.*



The diagonal language with to signals

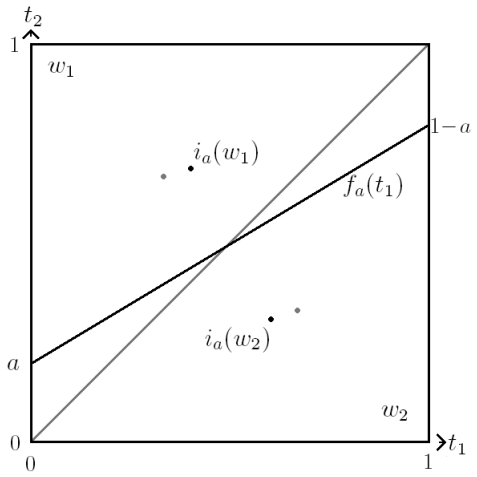
To specify a mutant strategy we parametrize the equilibrium strategy by the nonnegative small real number a. The sender strategy is

$$w_a(t) = \begin{cases} w_1 & \text{if } t_2 \geq a + (1-2a) \cdot t_1 \\ w_2 & \text{if } t_2 < a + (1-2a) \cdot t_1 \end{cases}$$

Defining deviating interpretations as functions of a, we consider best interpretations given w_a: Bayesian estimators $i_a(\hat{w}) = E[t|w_a^{-1}(\hat{w})]$.

$$\begin{aligned} i_a(w_1) &= \left(\frac{1}{3}(1+a), \frac{1}{3}(2+a-a2)\right) = i(w_1) + \frac{a}{3}\,(1, a(1-a)) \\ i_a(w_2) &= \left(\frac{1}{3}(2-a), \frac{1}{3}(1-a+a2)\right) = i(w_2) - \frac{a}{3}\,(1, a(1-a)) \end{aligned}$$

Of course, such a parametrization does not capture all possible deviations. Still, it captures a subset of deviations that can invade a population of agents with the equilibrium strategy in the sense of Apaloo (1997). Further, this parametrization allows us to directly apply Cressman, Hofbauer, and Riedel (2006) who study stability with respect to the replicator equation of equilibria for games with multidimensional continuous strategies.



A deviating strategy parametrized by a.

The payoff of an agent that uses a-deviation (w_a, i_a) *and meets an agent that uses b-deviation* (w_b, i_b) *is then*

$$\begin{aligned}\Lambda(a,b) &= \sum_{\hat{w}\in\{w_1,w_2\}} -\frac{1}{2}E\left[||t-i_a(\hat{w})||^2|t\in w_b^{-1}(\hat{w})\right] - \frac{1}{2}E\left[||t-i_b(\hat{w})||^2|t\in w_a^{-1}(\hat{w})\right] \\ &= \Lambda(0,0) + \frac{1}{18}\left(2a(1-a)b(1-b) - 2\left(b-a\right)2 - \left(b(1-b)-a(1-a)\right)2\right)\end{aligned}$$

with gradient

$$\nabla\Lambda(a,b) = \frac{1}{18}\left[\begin{array}{c} 4(b-a)+2(2b(1-b)-a(1-a))(1-2a) \\ 4(a-b)+2(2a(1-a)-b(1-b))(1-2b)\end{array}\right]$$

and second derivatives

$$\begin{aligned}\frac{\partial 2\Lambda(a,b)}{(\partial a)2} &= -\frac{1}{9}\left((1-2a)2+2(1-2b(1-b)-a(1-a))\right)\Big|_0 = -\frac{1}{3} \\ \frac{\partial 2\Lambda(a,b)}{\partial a\partial b} &= \frac{1}{9}\left(2+2(1-2a)(1-2b)\right)\Big|_0 = \frac{4}{9} \\ \frac{\partial 2\Lambda(a,b)}{(\partial b)2} &= -\frac{1}{9}\left(2+(1-2b)2+2(2a(1-a)-b(1-b))\right)\Big|_0 = -\frac{1}{3}\end{aligned}$$

At the equilibrium $(a,b)=0$ *the gradient is zero and* $\frac{\partial 2\Lambda(a,b)}{(\partial a)2}$ *is negative, which is the analytical counterpart of saying that the diagonal language is a strict Nash equilibrium.*

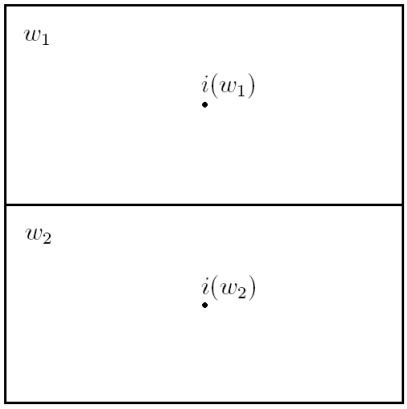
According to Eshel (1983) Theorem 1, a necessary condition for the diagonal language to be continuously stable is that $\frac{\partial 2\Lambda(a,b)}{(\partial a)2} + \frac{\partial 2\Lambda(a,b)}{\partial a\partial b} \leq 0$ *at* $(a,b)=0$*. As* $-\frac{1}{3}+\frac{4}{9}>0$*, the diagonal language is not CSS.*

Applying Theorem 4 of Cressman, Hofbauer, and Riedel (2006) to our setting, $\frac{\partial 2\Lambda(a,b)}{(\partial a)2} + \frac{\partial 2\Lambda(a,b)}{\partial a\partial b} > 0$ *at the diagonal equilibrium* $(a,b)=0$ *implies that the state in*

which each agent of the population chooses (w_0, i_0) is unstable with respect to the replicator equation restricted to normal distributions.

One can check that $\Lambda(0,a) < \Lambda(a,a)$, hence the diagonal language is not NIS (Apaloo (1997)).

Finally, the diagonal language is not efficient: $\Lambda(0,0) = -\frac{1}{3}$, which is less then the expected loss for the remaining strict Nash equilibrium ($-\frac{5}{48}$) which is depicted below:



The (efficient) horizontal two word-language

6 Case Studies

6.1 The line $[0,1]$

We consider now finite languages on real intervals. For simplicity, we look at uniformly distributed types and quadratic loss.

The game has many symmetries. In particular, for every Voronoi language, there exists an isomorphic language in which the words are permuted arbitrarily. Without loss of generality, we thus look at Voronoi languages that consist of points $0 \leq i_1 < i_2 < \ldots < i_K \leq 1$ for the receiver and corresponding Voronoi cells $[b_0, b_1), [b_1, b_2], \ldots, [b_{K-1}, b_K]$ for the sender, with $b_0 = 0$ and $b_K = 1$. $K \leq N$ is the richness of the language.

We claim that we must have

$$b_i = \frac{i}{K},\ i_j = \frac{2j-1}{2K}\,.$$

In other words: *Voronoi languages on $[0,1]$ consist of equidistant partitions and their midpoints.* Up to symmetries, there exists only one Voronoi language of a given richness.
PROOF : As the interpretation is the conditional expected types in a cell, we must have $i_j = \frac{b_{j-1}+b_j}{2}, j = 1, \ldots, K$. On the other hand, the points $b_0, \ldots, b_K$ describe the Voronoi tessellation corresponding to $i_1, \ldots, i_K$. Hence, we must have

$$i_1 = \frac{b_1}{2}, i_2 = \frac{b_2 + b_1}{2}, \ldots, i_K = \frac{b_{K-1}+1}{2}\,. \tag{4}$$

The unique solution of this system of linear equations is

$$b_i = \frac{i}{K},\ i_j = \frac{2j-1}{2K}\,.$$

(It is straightforward to see that this is a solution. Uniqueness may be unclear. Note that the i_j are uniquely determined by the b_l. Replace i_j by $1/2(b_{j-1} + b_j)$ in Eqn. (4). Rearrange these equations and you get sequentially

$$\begin{aligned} b_2 &= 2b_1 \\ b_3 &= 2b_2 - b_1 = 3b_1 \\ b_4 &= 4b_1 \\ &\vdots \end{aligned}$$

and so on until $b_K = Kb_1 = 1$ and you are done.) □

6.2 Two–Word–Languages in the unit square $[0,1]^2$

We omit the discussion of the (trivial) Voronoi language with one word in which no communication occurs. In this case, $i_1 = (1/2, 1/2)$ is the best interpretation.

In the unit square, we can have the following types of Voronoi tessellations (up to symmetry):

figure here

Up to symmetry, there are only two Voronoi languages: "left–right" and the diagonal language. (Beweis ist länglich, aber durchaus intressant).

The diagonal language is not stable, "left–right" is optimal (hence stable).

7 An Algorithm for Computing Voronoi Languages and Further Examples

In this section we examine languages with more than two words. As the computational problem of solving for three Voronoi tiles is demanding the problem becomes ambitious for more than three words. Even more challenging is to analyze stability properties. In the unit square example, a language with three words can be parametrized by a six-dimensional vector (two 'coordinates' for each of the three interpretations). For stability analysis, one needs to calculate the 6 dimensional Hessian matrix of the loss function. At least for more complex languages we expect to loose tractability when following a strictly analytical approach. For this reason we provide a section that relies on simulations. Although not stringent from a mathematical viewpoint, such simulations can well indicate whether a particular Voronoi tessellation renders stable or not. Further, one can extend the algorithm described here easily to settings with a finitely dimensional state space S (departing from the unit square) or to a setting with arbitrary distribution functions (departing from the uniform distribution).

7.1 The Algorithm

We verbally describe the algorithm step by step. The source code can be received upon request (/found in the appendix?).

• *Initialization* $t = 0$*:* $\quad i_1(0), i_2(0), \ldots, i_N(0)$

To start the algorithm, the interpretations receive initial values. These can be sophistically chosen as a particular Voronoi tessellation to test for its robustness in the presence of randomness. Alternatively, they can be randomly assigned to check for path dependence.

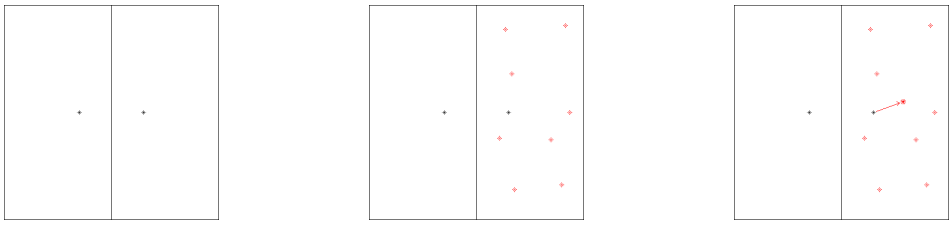
Hereafter, the algorithm finitely often iterates the following two steps:

• *Random Types*

Each iteration begins with randomly drawing finitely many types from T. Each sensation is assigned to its closest interpretation.

• *Tile Adaption*

The new value of the interpretation that represents a tile is the arithmetic mean of the types that are contained in that tile.

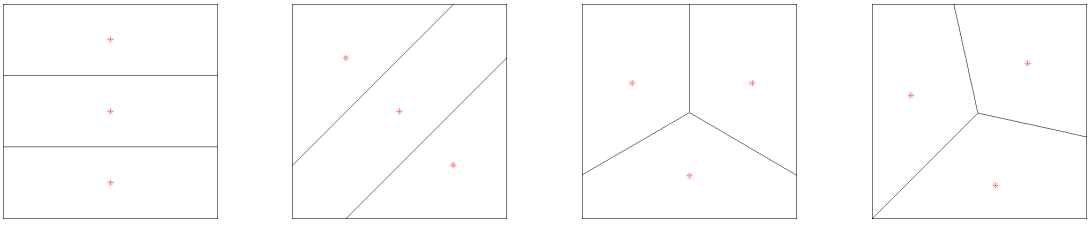


Initialization *Random Types* *Tile Adaption*

This surprisingly simple algorithm robustly selects some particular languages from a variety of Voronoi languages. On the other hand it is easy to show that some candidate languages render unstable in the presence of small deviations. We give some examples:

7.2 Three Words

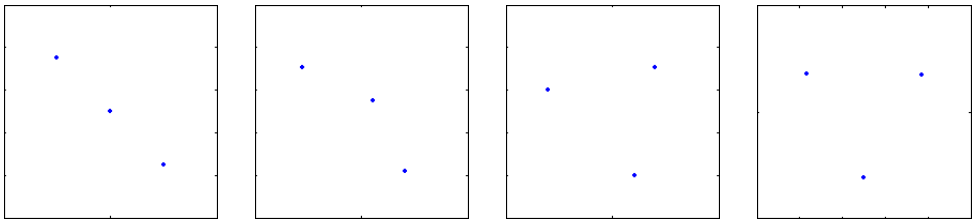
If the language comprises three words, up to symmetry there are two types of Voronoi tessellations which each have a 'horizontal' and a 'diagonal' version:



Ia Ib IIa IIb

The algorithm selects language 'IIa' which also has the property of minimizing the expected loss ← to be verified!. This has been tested for arbitrary initial conditions. The appendix derives the tessellations analytically, nevertheless we need to rely on the simulations for the finding of robustness.

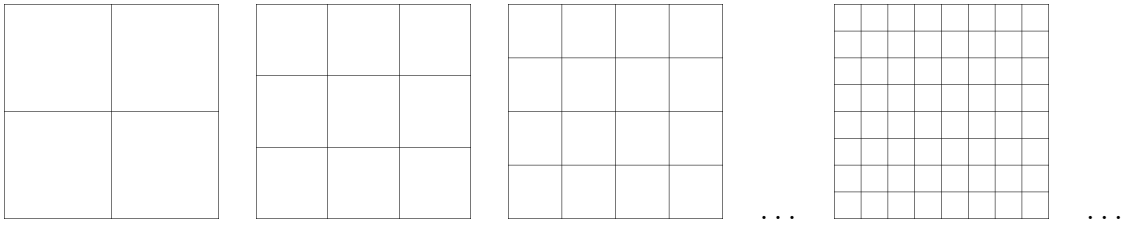
A sample path starting at 'Ib', bypassing 'IIb' and terminating close to 'IIa':



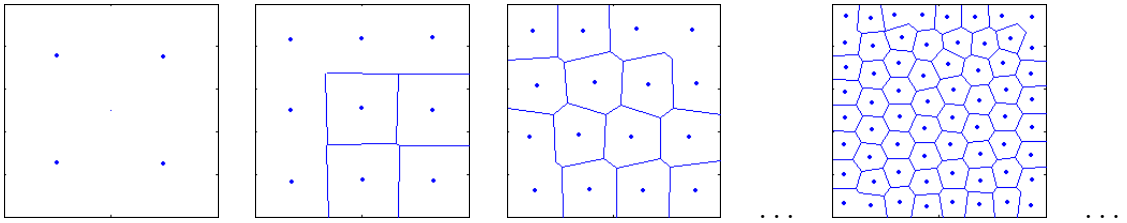
$t = 1$ $t = 50$ $t = 100$ $t = 1000$ ($\circlearrowleft$ by 90°)

7.3 Square Tessels

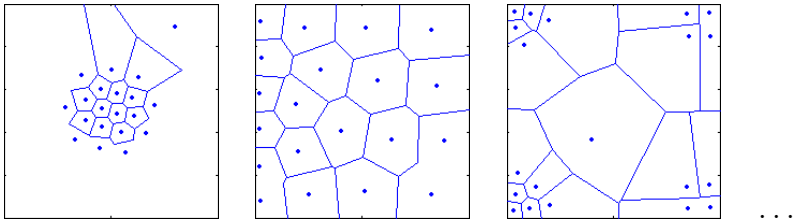
As indicated above, characterizing the set of Voronoi tessellations becomes more complex a problem, the more words the language has at disposal. Still, some tessellations are straightforward to describe. For any $n \in \mathbb{N}$, there is a Voronoi language with $n2$ tessels, as is illustrated below.



For small n, these languages seem stable while for large n they adapt a hexagonal structure known from behives. We depict the tessellations after 1000 iterations.



7.4 Other State Distributions



$\text{normal}_{.5}$ $s2$

7.5 Relation to k-means clustering

The algorithm described above can be seen as a stochastic generalization of the k-means clustering algorithm that is widely used in multivariate statistical data analysis and machine learning (sometimes under the heading of *vector quantization*; see for instance chapter 9.1 in Bishop (2006)). In these applications, we have finitely many observations that are unevenly distributed in an L-dimensional Euclidean space. The goal is to find a partition of the observations into k clusters (for some natural number k that is small in comparison to the number of observations) that minimizes the within-cluster loss (squared distance) and maximizes the between-cluster loss. The standard algorithm to find an optimal clustering of this kind is to start with an arbitrary k-tuple of prototypes, calculate the corresponding Voronoi tessellation, and to update each prototype to the arithmetic mean of the observations within its Voronoi tile. This process is repeated until a fixed point is reached. In the language of game theory, this amounts to an iterated best response computation for a discrete probability distribution of the type space. Our algorithm generalizes this idea to the continuous case.

A Proofs

A.1 Existence of Optimal Languages (Lemma 1)

We can identify strategies $w : T \to W$ for the sender with the corresponding partition $(W_j)_{j=1,\dots,N}$ given by

$$W_j = \{t \in T | w(t) = w_j\} \ .$$

Let $(i_j)_{j=1,\dots,N}$ be a pure strategy for the receiver. Given that strategy, a type t optimally selects an interpretation that is as close as possible to t, i.e. $w(t) \in \operatorname{argmin}_{j=1,\dots,N} \|t - i_j\|$. Note that in general, the interpretations i_j need not be pairwise distinct. In this case, we choose always the index with the smallest subscript, so we set

$$W_k^i = \left\{t \in T | k \text{ is the smallest number in } \operatorname{argmin}_{j=1,\dots,N} \|t - i_j\|\right\} \ .$$

We have thus reduced our optimization to a minimization problem over the compact set T^N, namely

$$\min_{(i_j)_{j=1,\dots,N} \in T^N} \int_T \sum_{k=1}^{N} l\left(\|t - i_k\|\right) 1_{W_k^i}(t) F(dt) \, .$$

For the existence of an optimal policy, it is thus sufficient to prove the continuity of the integral

$$\int_T \sum_{k=1}^{N} l\left(\|t - i_k\|\right) 1_{W_k^i}(t) F(dt)$$

in (i_k). By Lebesgue's theorem of dominated convergence, it is enough to show that the integrand $\sum_{k=1}^{N} l\left(\|t - i_k\|\right) 1_{W_k^i}(t)$ is F–almost everywhere continuous. We can ignore

the boundaries of the sets W_k^i because these boundaries are intersections of hyperplanes with the set T and therefore Lebesgue, hence F–nullsets. Take a type $t \in T$ in the interior of some W_m^i for some $1 \leq m \leq N$. Being in the interior of W_m^i, i_m is the unique interpretation with minimal distance to t. Take a sequence $\left((i_j^n)_{j=1,\ldots,N}\right)_{n\in\mathbb{N}}$ with $i_j^n \to i_j$ as $n \to \infty$ for all $j = 1, \ldots, N$. For n sufficiently large, i_m^n is the unique interpretation among (i_j^n) with minimal distance to t and $i_k^n \in W_k^i$. Therefore, the continuity of l entails

$$\sum_{k=1}^{N} l\left(\|t - i_k\|\right) 1_{W_k^i}(t) = l\left(\|t - i_m\|\right) 1_{W_m^i}(t) = \lim_{n\to\infty} l\left(\|t - i_m^n\|\right) 1_{W_m^i}(t)\,.$$

Thus, the integrand is F-a.e. continuous.

A.2 Mixed Strategies are Never Optimal (Lemma 2)

Fix any $t \in T$. Randomized strategies (ω, μ) lead to a probability distribution $\gamma_t(di) = \sum_{k=1}^{N} \mu_k(di)\omega_k(t)$ over T. Suppose that this measure is not a Dirac measure.

Now denote by $\bar{\gamma} = \sum_{k=1}^{N} \int_T i\ \mu_k(di)\omega_k(t)$ the average outcome of communication in T when (ω, μ) is played. The function $i \mapsto l\left(\|t - i\|\right)$ is strictly convex; by Jensen's inequality,

$$\sum_{k=1}^{N} \int_T l\left(\|t - i_k\|\right) \mu_k(di_k) \leq l\left(\|t - \bar{\gamma}\|\right)\,,$$

and the inequality is strict when γ is not a Dirac measure. This shows that mixing is never optimal.

A.3 Structure of Optimal Languages (Lemma 3, Lemma 4, Lemma 5) and Theorem 1

Let (w, i) be an optimal language. From our analysis in Section A.1, we know that we can identify w without loss of generality with the partition

$$W_k = \left\{t \in T | k \text{ is the smallest number in } \operatorname{argmin}_{j=1,\ldots,N} \|t - i_j\|\right\}\,.$$

Note that these sets W_k are either intersections of convex polyhedra with the type set T or empty, if some word is not used. Suppose that the word w_N is not used, i.e. $W_N = \emptyset$. The idea of the proof is to take a word that is used for a big set of types and to split that set into two smaller sets and to use two words instead of one. This allows to decrease the average loss.

By definition, word w_1 is used with positive probability , i.e. the convex set W_1 has positive mass with respect to F. As F is atomless, we can find two disjoint, convex,

nonnull sets A_1, A_N with $A_1 \cup A_N = A$. Now let $j_1 \in T$ be a minimizer[7] of

$$\int_{A_1} l\left(\|t - j\|\right) F(dt),$$

and similarly, $j_N \in T$ be a minimizer of

$$\int_{A_N} l\left(\|t - j\|\right) F(dt).$$

By strict convexity of l, the minimizers are uniquely determined. Moreover, we have $j_1 \neq j_N$ because the minimizer lies in the interior of A_1 resp. A_N.

Set $j_k = i_k$ for $k = 2, \ldots, N-1$. Moreover, set $v(t) = w_1$ for $t \in A_1$ and $v(t) = w_N$ for $t \in A_N$, and $v(t) = w(t)$ else. We claim that (v, j) is a better language than (w, i):

$$L(v,j) - L(w,i) = \int_{A_1} \left(l\left(\|t - j_1\|\right) - l\left(\|t - i_1\|\right)\right) + \int_{A_N} \left(l\left(\|t - j_N\|\right) - l\left(\|t - i_1\|\right)\right) > 0$$

where the last inequality comes from the fact that j_1 and j_N minimize the loss over the sets A_1 and A_N and either $j_1 \neq i_1$ or $j_N \neq i_1$.

It remains to be shown that all interpretations (i_k) are pairwise distinct. Given that the signaling system is induced by a partition (W_k) of convex sets with positive measure, the optimal interpretation for word w_k is the "prototype" i_k that minimizes

$$\int_{W_k} l\left(\|t - j\|\right) F(dt)$$

for $j \in T$. As W_k is convex and F atomless, the minimizer lies in the interior of the set W_k. In particular, all interpretations (i_k) are pairwise distinct for an optimal language. Moreover, we see that the receiver uses a best estimator in the sense of Definition 2.

A.4 Evolution (Proof of Lemma 6, Theorem 3, Lemma 8)

The proof that average loss is decreasing along payoff–monotone dynamics and BNN dynamics follows well–known lines, see Heifetz, Shannon, and Spiegel (2007) and Hofbauer, Oechssler, and Riedel (2007). The loss function Λ is continuous with respect to the weak topology if the direct loss function for pure strategies $L(w, i)$ is continuous (in the usual norm on T^N and Σ endowed with the supremum–norm) and bounded.

The maximal distance on T is bounded because T is compact, and l is continuous, so L remains bounded.

To see continuity, choose a sequence (w^n) of sender strategies that converge uniformly to w and a sequence (i^n) of receiver strategies that converges to $i \in T^N$. Let $\epsilon > 0$ and $\delta > 0$ such that $|l(d) - l(e)| < \epsilon$ for all $0 \leq d, e \leq \max_{s,t \in T} \|s - t\|$. (Note that l is

[7]The minimum exists because T is compact and the expression is continuous in j, see the proof of Lemma 1.

uniformly continuous on bounded intervals and that the maximum is finite because T is compact.) As sender strategies can assume only finitely many values in W, there exists $N_0 \in \mathbb{N}$ such that $w^n(t) = w(t)$ uniformly in $t \in T$ for $n \geq N_0$. It follows that

$$\|t - i_{w^n(t)}\| = \|t - i_{w(t)}\| \tag{5}$$

uniformly in $t \in T$ for $n \geq N_0$. Now choose $N_1 \geq N_0$ such that for $n \geq N_1$

$$\|i_j^n - i_j\| < \delta$$

for all $j \in \{1, \ldots, N\}$. Then

$$\begin{aligned} |L(w^n, i^n) - L(w, i)| &\leq \int_T \left| l\left(\|t - i^n_{w^n(t)}\|\right) - l\left(\|t - i^n_{w(t)}\|\right)\right| F(dt) \\ \text{Eqn.(5)} \quad &= \int_T \left| l\left(\|t - i^n_{w(t)}\|\right) - l\left(\|t - i^n_{w(t)}\|\right)\right| F(dt) \\ \text{(Def. of } \delta) \quad &< \epsilon\,. \end{aligned}$$

Hence, L is continuous.

Let P^* be evolutionarily robust with invasion barrier $\epsilon > 0$. Then we have for populations $Q \neq P^*$

$$\begin{aligned} \Lambda(P^*, P^*) - \Lambda(Q, Q) &= \Lambda(P^*, P^*) - \Lambda(P^*, Q) + \Lambda(P^*, Q) - \Lambda(Q, Q) \\ &< \Lambda(P^*, P^*) - \Lambda(P^*, Q) \\ &= \Lambda(P^*, P^*) - \Lambda(Q, P^*) \leq 0, \end{aligned}$$

where we use the definition of ER, symmetry of Λ and the fact that (P^*, P^*) is a Nash equilibrium. This shows that P^* is a strict local minimum of $\Lambda(Q, Q)$ and proves Lemma 8.

Let us come to stability questions (Theorem 3). With a Lyapunov function, dynamic stability of local optima follows as usual (see Bhatia and Szegő (1970), Ch. V).

References

Apaloo, J. (1997): "Revisiting Strategic Models of Evolution: The Concept of Neighborhood Invader Strategies," *Theoretical Population Biology*, 52, 71–77.

Bhatia, N. P., and G. P. Szegő (1970): *Stability Theory of Dynamical Systems*. Springer.

Bishop, C. M. (2006): *Pattern Recognition and Machine Learning*. Springer.

Blume, A., Y.-G. Kim, and J. Sobel (1993): "Evolutionary Stability in Games of Communication," *Games and Economic Behavior*, 5, 547–575.

Cressman, R., J. Hofbauer, and F. Riedel (2006): "Stability of the Replicator Equation for a Single-Species with a Multi-Dimensional Continuous Trait Space," *Journal of Theoretical Biology*, 239, 273–288.

Eshel, I. (1983): "Evolutionary and Continuous Stability," *Journal of Theoretical Biology*, 103, 99–111.

Eshel, I., and E. Sansone (2003): "Evolutionary and Dynamic Stability in Continuous Population Games," *Journal of Mathematical Biology*, 46, 445–459.

Gärdenfors, P. (2000): *Conceptual Spaces*. The MIT Press, Cambridge, Mass.

Heifetz, A., C. Shannon, and Y. Spiegel (2007): "What to maximize if you must," *Journal of Economic Theory*, 133(1), 31–57.

Hofbauer, J., J. Oechssler, and F. Riedel (2007): "Brown–von Neumann–Nash Dynamics: The Continuous Strategy Case," Working Paper.

Jäger, G. (2007): "The evolution of convex categories," *Linguistics and Philosophy*, 30(5), 551–564.

Jäger, G., and R. van Rooij (2007): "Language Stucture: Psychological and Social Constraints," *Synthese*, 159(1), 99–130.

Labov, W. (1973): "The boundaries of words and their meanings," in *New Ways of Analysing Variation in English*, ed. by C.-J. N. Bailey, and R. W. Shuy, pp. 340–373. Georgetown University Press, Washington.

Löbner, S. (2003): *Semantik. Eine Einführung*. Walter de Gruyter, Berlin, New York.

Oechssler, J., and F. Riedel (2001): "Evolutionary Dynamics on Infinite Strategy Spaces," *Economic Theory*, 7, 141–162.

——— (2002): "On the Dynamic Foundation of Evolutionary Stability in Continuous Models," *Journal of Economic Theory*, 107, 223–252.

Okabe, A., B. Boots, and K. Sugihara (1992): *Spatial tessellations: concepts and applications of Voronoi diagrams*. Wiley, Chichester.

Trapa, P., and M. Nowak (2000): "Nash equilibria for an evolutionary language game," *Journal of Mathematical Biology*, 41, 172–188.

Wärneryd, K. (1993): "Cheap talk, coordination and evolutionary stability," *Games and Economic Behavior*, 5, 532–546.